Macrophages and CD4 T-cells in Rheumatoid Arthritis and Their Modulation by JAK Inhibitors

Jagtar Singh Nijjar

BSc (Med.Sci.) Clinical Medicine (Hons.), MBChB (Hons.), MRCP

Submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

College of Medical, Veterinary and Life Sciences

Institute of Infection, Inflammation and Immunity

University of Glasgow

December 2015

Abstract

Background

Rheumatoid arthritis (RA) is a chronic inflammatory arthritis that causes significant morbidity and mortality and has no cure. Although early treatment strategies and biologic therapies such as TNFα blocking antibodies have revolutionised treatment, there still remains considerable unmet need. JAK kinase inhibitors, which target multiple inflammatory cytokines, have shown efficacy in treating RA although their exact mechanism of action remains to be determined. Stratified medicine promises to deliver the right drug to the right patient at the right time by using predictive 'omic biomarkers discovered using bioinformatic and "Big Data" techniques. Therefore, knowledge across the realms of clinical rheumatology, applied immunology, bioinformatics and data science is required to realise this goal.

Aim

To use bioinformatic tools to analyse the transcriptome of CD14 macrophages derived from patients with inflammatory arthritis and define a JAK/STAT signature. Thereafter to investigate the role of JAK inhibition on inflammatory cytokine production in a macrophage cell contact activation assay. Finally, to investigate JAK inhibition, following RA synovial fluid stimulation of monocytes.

Methods and Results

Using bioinformatic software such as limma from the Bioconductor repository, I determined that there was a JAK/STAT signature in synovial CD14 macrophages from patients with RA and this differed from psoriatic arthritis samples. JAK inhibition using a JAK1/3 inhibitor tofacitinib reduced TNFα production when macrophages were cell contact activated by cytokine stimulated CD4 T-cells. Other pro-inflammatory cytokines such as IL-6 and chemokines such as IP-10 were also reduced. RA synovial fluid failed to stimulate monocytes to phosphorylate STAT1, 3 or 6 but CD4 T-cells activated STAT3 with this stimulus. RNA sequencing of synovial fluid stimulated CD4 T-cells showed an upregulation of SOCS3, BCL6 and SBNO2, a gene associated with RA but with unknown function and tofacitinib reversed this.

Conclusion

These studies demonstrate that tofacitinib is effective at reducing inflammatory mediator production in a macrophage cell contact assay and also affects soluble factor mediated stimulation of CD4 T-cells. This suggests that the effectiveness of JAK inhibition is due to

inhibition of multiple cytokine pathways such as IL-6, IL-15 and interferon. RNA sequencing is a useful tool to identify non-coding RNA transcripts that are associated with synovial fluid stimulation and JAK inhibition but these require further validation. SBNO2, a gene that is associated with RA, may be biomarker of tofacitinib treatment but requires further investigation and validation in wider disease cohorts.

Table of Contents

List of Tables

List of figures

Acknowledgments

Firstly, I would like to thank my supervisors Professor McInnes and Dr Kurowska-Stolarska for their support, guidance and inspiration through this PhD. Without your belief and patience in me this would not have been possible.

Thanks also to Dr Derek Gilchrist without whom this thesis could not have been completed and written and also for sparking my interest in all that is bioinformatics and Big Data related. You have been a great mentor and an even better friend and I hope to be able to repay the huge debt to you over the coming years.

My place of work, the University of Glasgow, has put me in contact with many friends and colleagues who have provided immense support. Without help from the wider CRD research group, none of this would have been possible. Special thanks to Miss Ashley Gilmour, Dr Moeed Akbar, Mr Jim Reilly and Mrs Shauna Kerr for your assistance with experiments. I have a group of amazing friends from all walks of life and if I learn nothing more from this experience it is that you cannot quantify the support of those around you.

Combining bioinformatics, lab science and clinical rheumatology has been a challenge. Without the help of Professor Rainer Breitling and his wider group this would have been impossible. Thank you to Dr Ronan Daly for putting up with me when my simple clinician mind could not understand the intricacies of R. A big thank you to Glasgow Polyomics and in particular Dr Tanita Casci and Professor Mike Barrett who gave me a home for so long while I learned the intricacies of data analysis. To Jing and Julie, thank you for getting my sequencing done on time and with the utmost of care. Finally to my two good friends Fraser and David, your infectious enthusiasm, hard work and belief in me gives me courage to continue down this path.

This work would have been impossible the patients with rheumatoid arthritis who gave their blood to allow me to gain knowledge. They inspire me every day as they struggle with a disease that affects each and every aspect of their lives and I will to the best of my ability use this work to treat you better in the future.

To Moira McDonald and the staff at the Rheumatology Day Ward of the Glasgow Royal Infirmary thank you for taking time to identify appropriate patients for my studies and sending samples across whenever I needed them.

To my friend Laura, you have done more for me over the last year than anyone will ever know. Your support has been unwavering and you give selflessly. Your dedication and kindness to your patients is truly amazing and an inspiration. Never change.

To Deepankar, you are always there for me whenever I ask despite working an insane lifestyle in emergency medicine. I hope that I can repay your kindness in the future.

To Becca, thanks for keeping me sane when everything seemed to be falling apart around me. You gave me confidence that I could complete this.

I can't put into words the support my parents have given me. You have always been there for me and have allowed me to pursue opportunities knowing that it would impact you. I haven't been the best son I could be but I will try harder in the future. This work is dedicated to you.

To my Bebeji, who passed during my doctoral studies, I hope I have made you proud of me. And don't worry my hands are warm and I'm wearing my gloves.

To my brother and sister, I thank you for putting up with an absent brother who doesn't see you for days on end. I have tried to be there when it counts and will ensure that I am there for you both in the future

To my children, Meher and Prabhleen, I am sorry that I have not been around very much over the last few years and I hope that you will understand and forgive me as you learn more about your Daddy. You two are truly the most amazing people I have ever had the honour of looking after. Make sure you go forth and be excellent.

My wife, Sohini. This has been a struggle for you and I don't think I ever fully explained how difficult this time would be. Your unwavering support and love have meant that I kept going every day. Without you keeping our home together and making sure that the children were looked after, this would simply not have been possible. To paraphrase John Nash - I'm only here because of you. You are the reason I am and all my reasons. This work is also dedicated to you.

Author's Declaration

I declare that this thesis is the result of my own work. Where others have assisted with experiments is indicated at the beginning of the appropriate chapter. No part of this thesis has been submitted for any other degree at The University of Glasgow, or any other institution.

Jagtar Singh Nijjar

Chapter 1 Introduction

1.1 Rheumatoid Arthritis

1.1.1 Introduction

Rheumatoid Arthritis (RA) is a chronic symmetrical inflammatory arthritis with no cure that affects synovial joints but which typically causes painful swelling of the hands, wrists and knees. Patients are often female with a typical onset from 30-50 years old (1)and often have significant disability and depression caused by their symptoms and the inflammatory process(2). Furthermore, RA has systemic manifestations with organs such as the lungs, eyes and skin being involved. Treatment revolves around a multidisciplinary team made up of doctors, nurses, physiotherapists, occupational therapists and other allied health professionals. Pharmacotherapy includes painkillers, conventional Disease Modifying anti-rheumatic drugs (cDMARDS), immunosuppressant drugs, Biologic therapies and kinase inhibitors.

1.1.2 Epidemiology

The incidence of RA in the UK is 30 per 100,000 patients but the prevalence is approximately 1% in the Caucasian population as there is no cure for this chronic condition. The incidence and prevalence does vary worldwide (3)with a significantly higher level in Inuits (4)and a lower incidence in Asian and Indian populations. This is hypothesised to be due to environmental and genetic factors that will be discussed later in this thesis. Females are four times more likely to be affected by RA and this is in contrast with some other arthropathies such as Psoriatic Arthritis (PsA), which has an equal incidence amongst both sexes or indeed with connective tissue diseases such as Systemic Lupus Erythematosus (SLE) which is much more common in females(5). The average age of onset is 45 with a broad range of onset from 15 till elderly years.

1.1.3 Risk factors

1.1.4 Genetics

There is good evidence that there are various genetic traits that predispose the general population to developing RA. Although the recent metanalysis of GWAS in RA (6) will be discussed further in the immunopathogenesis section, there are a few genetic factors which were discovered prior to GWAS. These were the association of RA with HLA-DR3 and 4 (7,8) and the shared epitope(9). In particular there is evidence to suggest that having a genetic predisposition and other environmental factors such as smoking significantly increase your risk of developing RA(10,11).

The risk conferred by variation of the Type II Major Histocompatibility Complex (MHC) have led to the hypothesis that the interaction between the innate and adaptive immune system is critical in the pathogenesis of RA(12). In particular self-antigen presentation to the naïve or primed immune system has been thought to be a critical step in much the same way as mutations of tumour suppressor genes predispose you to cancer. It is this interaction which led to researchers considering treatments which target antigen presenting cells such as macrophages and B-cells as well as the adaptive immune system in the form of mainly CD4+ T-cells.

1.1.5 Environment

Smoking is one of the major acquired risk factors for the development of both autoantibodies and RA(13,14). Smoking as well as having over four hundred toxins that can cause oxidative damage, is known to cause damage to the mucous membranes of the airways as well as the lung parenchyma. It is suggested that damage to the lungs results in an increase in peptidyl arginine deaminase 4 which leads to citrullination of arginine in self proteins such as fibronectin, alpha enolase and vimentin(15).

In addition to smoking, occupational hazards and in particular pulmonary exposure to silica is a risk factor for not just rheumatoid arthritis but also other rheumatic diseases(16). Caplan's syndrome which is rheumatoid arthritis combined with pneumoconiosis (17)can also be caused by asbestos and coal dust with the later likely contributing to the pathogenesis via similar agents as found in cigarette smoke.

Furthermore High Resolution Computerized Tomography (HRCT) of the lungs of patients with early RA has shown that there are inflammatory changes in the lungs include parenchymal oedema. Furthermore, there is evidence of citrullinated peptides in mucosal biopsies and antibodies to citrullinated peptides (ACPA) that are higher in broncheoalveolar lavage fluid than serum suggesting local production(18). Finally smoking has been shown to facilitate resistance to treatment, in particular biologics in the form of anti-TNFα agents(19).

Social deprivation results in more severe rheumatoid arthritis but no increased risk of development(20). It is likely that multiple risk factors contribute to this but even when smoking is removed from prediction models, deprivation still confers added risk(21). This risk factor may however act as a surrogate for other factors such as microbiome differences or even adherence to treatment regimes and medication.

1.2 Clinical Features

Rheumatoid Arthritis causes pain swelling and early morning stiffness of synovial joints. The arthritis is symmetrical and tends to affect the metacarpal phalangeal joints of the hand and also the proximal interphalangeal joints as well as the radiocarpal and radial ulnar joints. Left unchecked the arthritis results in destruction of the cartilage and bone by inflammation and recruitment of osteoclasts to the site of inflammation resulting in bone erosions and therefore joint destruction. There is often a significant element of early morning stiffness alongwith fatigue which results in further disability and impacts of activities of daily living and the ability of the patient to interact with society.

In addition to the articular features, extra-articular manifestations result in significant co-morbidity either due to a direct pathologic effect or a secondary effect of the inflammation or treatments.

1.3 Diagnosis

There is no diagnostic test for RA and diagnosis relies on a referral being made in a timely fashion from primary care or a family physician to a rheumatologist. Diagnosis is made from the clinical history of painful and swollen joints, an examination demonstrating

synovitis and raised inflammatory markers and autoantibodies on blood testing. In addition to these basic criteria, there have been two sets of classification criteria: the ACR 1987 criteria (22) and the newer ACR/EULAR 2010 (23)criteria. These criteria offer a way to classify RA and therefore allow meaningful studies to be carried out in the disease.

1987 American Rheumatism Association criteria for the classification of rheumatoid arthritis – RA is defined by four out of the seven criteria below being present
Morning stiffness around joints lasting for 1 hour
Soft tissue swelling around 3 or more joints as observed by a physician
Arthritis of the proximal interphalangeal, metacarpophalangeal or wrist joint
Symmetrical distribution of arthritis
Rheumatoid nodules
Presence of rheumatoid factor
Radiographic erosions and/or presence of periarticular osteopenia in hands and/or wrists

Table 1-1 1987 ARA criteria for the classification of RA

2010 ACR/EULAR classification criteria for rheumatoid arthritis where the patient has at least one synovitic joint and this cannot be better explained by another disease		
Joint Involvement		
	1 large joint	0
	2-10 large joints	1
	1-3 small joints (with or without large joints)	2
	4-10 small joints (with or without large joints	3
	>10 joints with at least 1 small joint involved	5
Serology		
	Negative RF and negative ACPA	0
	Low positive RF or low positive ACPA	2
	High positive RF or high positive ACPA	3
Acute phase reactants		
	Normal CRP and normal ESR	0
	Abnormal CRP or normal ESR	1
Duration of symptoms		
	< 6 weeks	0
	≥ 6 weeks	1

Table 1-2 2010 ACR/EULAR classification criteria for RA

As well as classifying disease, it is useful to have a way to measure disease activity in an objective fashion in order to guide treatment escalation and also de-escalation if a state of remission has been achieved. There are various tools to carry this out but the most widely used is the Disease Activity Score in 28 joints (24).

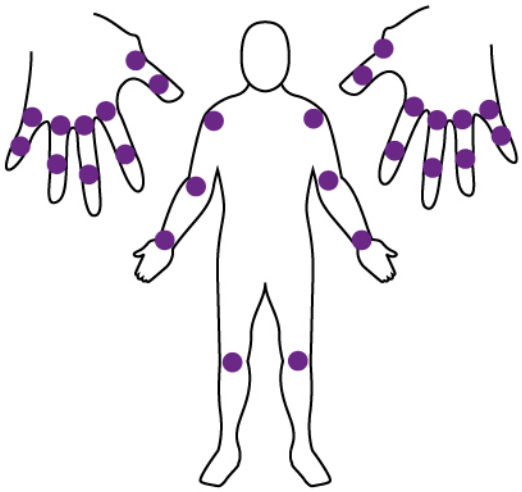


Figure 1-1: Calculation of the DAS 28 score involves assessing each of the joints above for swelling and tenderness. Image is reproduced from http://www.dermatologiahuec.com/index.php?pag=das28

The score is comprised of:

- A tender joint count
- A swollen joint count
- A measure of global health by the patient on a 100mm visual analogue scale
- A measure of inflammation such as ESR or CRP

DAS score corresponding to categories of disease activity			
Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
<2.6	2.6-3.2	3.2-5.1	>5.1

Table 1-3: Categories of Disease Activity Score

However the DAS-28 is limited by the fact that it tends to omit disease activity in the ankles and feet and also the calculation of the score is often delayed because it involves the measurement of an inflammatory marker. Similarly the SDAI score involves calculating the sum of a tender and swollen 28 joint count along with a measure of both a patient and physician global disease activity assessment and an inflammatory marker (25)but the CDAI score omits the inflammatory marker and therefore can be calculated in

clinic(26,27). In much the same was as DAS, there are categories for escalating therapy based on level of disease activity.

As well as measuring the absolute disease activity, there are various markers of response that describe the degree of improvement that needs to be achieved for each level. These are useful because they give criteria that need to be fulfilled in terms of an absolute change to be achieved along with a general disease activity state.

EULAR response criteria (29)			
	Improvement in DAS 28 from Baseline		
	≤ 1.2	**> 0.6 and ≤ 1.2**	**≤ 0.6**
DAS28 at endpoint			
≤ 3.2	Good	Moderate	None
> 3.2 and ≤5.1	Moderate	Moderate	None
> 5.1	Moderate	None	None

Table 1-4: EULAR response criteria

ACR20 response criteria. Achievement of ACR50 or ACR70 requires a percentage improvement of that amount in the criteria below. (28)	
Required criteria	
	≥ 20% improvement in 28 tender joint count
	≥ 20% improvement in 28 swollen joint count
≥ 20% improvement in 3 out of 5	
	Patient assessment of pain
	Patient assessment of global health
	Physician assessment of global health
	Patient self assessment of disability
	An acute phase reactant either ESR or CRP

Table 1-5: ACR20 response criteria

1.4 Investigations

1.4.1 Blood testing

A systemic inflammatory process is often shown in routine bloods such as a full blood count (FBC) and inflammatory markers such as the erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP) level. A thrombocytosis or leucophilia may be present in active RA as well as anaemia that can be due to multiple causes including disease-associated, drug related or a deficiency state. Furthermore due to the systemic nature of the disease, it is necessary to check kidney and liver function as treatments such as non-steroidal anti-inflammatory drugs (NSAIDs) or cDMARDS can cause derangements.

1.4.2 Autoantibodies

Rheumatoid Factor (RF) was the first autoantibody to be described in RA by Waaler. RF is IgG, IgA or IgM to the Fc portion of IgG and therefore it was hypothesised that patients with RF are more likely to form RF and IgG complexes therefore resulting in more tissue damage. The exact mechanism was unclear but was thought to involve complement activation more than antibody dependent cellular cytotoxicity and activation of the innate immune system. Further RF gave further evidence that there is a loss of appreciation of "self" by the immune system and also demonstrates a role for B-cells.

The detection of antibodies to citrullinated peptides (ACPA) led to a further shift in the role of autoantibodies in RA(30). These antibodies are raised to proteins that have undergone citrullination where an arginine residue is converted to citrulline by the enzyme Peptidyl arginine deaminase (PADI). Citrullinated peptides do not exist in humans naturally and therefore this process gives further weight to the theory that the immune systems loses a sense of "self". ACPA have been shown to predate onset of RA (31) (32)and are also more common in smokers(33).

Although RF is thought not to be pathologic, ACPA have been shown to promote and activate osteoclasts therefore giving a mechanism of action for the observation that ACPA are more often found in patients who have erosive disease(34). In addition, citrullinated fibrinogen has been shown via TLR4 to induce TNFα production from macrophages(35).

1.4.3 Imaging

Plain radiographs of the hands, feet and other affected joints typically show evidence of joint space narrowing or periarticular osteopenia in early disease. There may also be evidence of erosion where the inflammatory pannus has caused resorption of the bone in juxtaposition to the joint. Erosions at baseline are a poor prognostic factor for disease severity and response to therapy(36). In addition to erosion, when further joint damage occurs, ulnar deviation, subluxation and dislocation of the joints can be seen especially around the metacarpophalangeal and wrist joints.

However radiographic imaging may lag behind acute inflammatory changes and therefore dynamic imaging techniques such as ultrasound and magnetic resonance imaging (MRI) are employed to detect evidence of synovitis. Furthermore, these techniques image bone and soft tissue and therefore allow assessment of tenosynovitis and bone marrow oedema in the case of MRI.

1.5 Management

The management of RA revolves around three main partners within the larger entity of society. These are namely the patient, their healthcare providers and the patient's family. Interventions in each of these areas make a significant contribution to the wellbeing of the patient and their journey. Prior to effective strategies and drugs to treat synovitis, the hallmark of therapy in RA involved assembling a team of healthcare providers who in partnership with the patient and family could improve the quality of life of the patient and also decrease the impact of their disability interacting with society. The multi-disciplinary team is employed within many other medical specialties but it is crucial in Rheumatology to ensure the safe and effective treatment of our patients(37).

This team is even more important in the era of combination DMARD, biologic and kinase inhibitor treatments as patients, their families and also other healthcare professionals come to grips with the ways in which we can re-train the immune system. A summary of MDT team members is below:

Team member	Role
General Practitioner	Overarching co-ordination of healthcare and wellbeing of patient
Rheumatologist	Diagnosis of rheumatoid arthritis, initiation of therapy and co-ordination of MDT
Physiotherapist	Assessment of joint restriction and rehabilitation. Specialist physiotherapists can give joint injections or perform ultrasound examinations
Occupational Therapist	Determine whether input is required to home or workplace of the patient in order to keep them as active as they wish to be
Rheumatology Specialist Nurse	Initiation of further therapies and a rich source of information and contact for the patient. In the UK, specialist nurses often carry out screening and monitoring for DMARD and biologic drugs
Podiatrist	Specialist input into the assessment and treatment of foot disease which can be neglected by in patients with RA
Social Worker	RA is a risk factor for unemployment and a social worker with knowledge of the varied ways in which the disease can affect a patient is crucial
Psychologist	Diagnosis with a chronic disease is a risk factor for depression and there is evidence that the disease process itself cause changes within the brain

Table 1-6: Multidisciplinary team members crucial in the management of RA

1.5.1 Disease Modifying Anti-Rheumatic Drugs

Once RA is diagnosed, therapy with DMARDs should be started as soon as possible. There is good evidence that treating earlier and in an intensive manner leads to better patient outcomes, less damage and less disability(38). Furthermore combination and step-up therapy is now common and although regimes differ worldwide, the principles remain the same(39,40).

Commonly used DMARDS are shown in the table 1-7 along with their doses and side effect profile. Many DMARDs require blood monitoring although the burden of this lessens after six months of therapy.

Drug	Dose range	Side effects	Monitoring
Methotrexate	10-25mg weekly	Hepatitis. Blood dyscrasias. Pulmonary fibrosis	FBC, UE, LFT, screening chest XR
Sulphasalazine	1.5-3g per day	Hepatitis. Blood dyscrasias. GI side effects.	FBC, UE, LFT
Hydroxychloroquine	200-400 mg per day	GI side effects. Retinopathy	Yearly screening by optician
Leflunomide	10-20mg per day	Hepatitis. Nausea. Hypertension	BP monitoring, FBC, UE, LFT
Ciclosporin	Maximum 4 mg/kg	Hepatitis. Hypertension. Gingival hyperplasia. Renal Toxicity	BP monitoring, FBC, UE, LFT
Penicillamine	Normally 500-750mg daily	Thrombocytopenia, proteinuria	FBC and urine protein testing
Gold	50mg but dosage interval varies	Leucopenia, proteinuria, rash	FBC including differential white cell count. Urine protein testing

Table 1-7: Disease Modifying Drugs in RA

There are a significant number of patients who despite combination DMARD therapy, continue to have high disease activity. Prior to the beginning of this millennium, there were few other treatment options but at that point Infliximab and Etanercept which both target TNFα had just completed two landmark phase III clinical trials in RA(41,42) with excellent results.

1.5.2 Biologics

Biological drugs or biologics are made from living organisms and can be used to prevent, diagnose or treat diseases. Biologics in rheumatology often target either cytokines or their receptors but other classes of biologics include:

- Vaccines
- Synthetic peptides such as synthetic insulin e.g. Lantus
- Growth and haematopoetic factors such as erythropoietin and granulocyte and macrophage colony stimulating factor (GM-CSF)
- Antibodies targeting cell surface receptors therefore facilitating antibody dependent cellular cytotoxicity such as OKT3 against CD3 on T-cells and Rituximab against CD20 on B-cells
- Antibodies targeting free or bound cytokine such as anti-TNFα agents
- Antibodies against receptors such as tocilizumab blocking IL6R
- Fusion proteins such as abatacept that binds to CD80/86

In the 1990s work from the Kennedy Institute in Imperial College London showed that inhibiting Tumour Necrosis Factor Alpha (TNFα), using a monoclonal antibody, in patients with RA led to a significant improvement in both signs and symptoms of their inflammatory arthritis along with a decrease in the level of their disability(43). In 2000, two biologic agents, Infliximab (41)and Etanercept (42)had their Phase III clinical trials which showed good efficacy and relative safety in patients with RA who had a high disease activity despite being treated with methotrexate.

1.5.3 Agents targeting TNFα

There are currently five agents on the market that target this cytokine and their characteristics are shown in the table 1-8.

Drug	Origin	Frequency of dosage	Route
Infliximab	Humanised mouse monoclonal antibody	6 weekly	Intravenous
Etanercept	Human TNF receptor	Weekly	Subcutaneous injection
Adalimumab	Human monoclonal antibody from phage display	Fortnightly	Subcutaneous injection
Golimumab	Human monoclonal antibody	Monthly	Subcutaneous injection
Certolizumab	Human Fab portion of antibody attached to polyethylene glycol	Monthly after loading	Subcutaneous injection

Table 1-8: Currently available anti TNFα agents for RA

These agents are more efficacious if used with methotrexate (44) although the recent RACAT trial (45) demonstrated that DMARD triple therapy was non-inferior to etanercept and methotrexate. These agents have now been used in hundred of thousands of patients in the UK and thanks to registries we have significant amount of safety data with regards to their potential to cause infections (46)or cancers(47).

In general these agents are safe but particular caution has to be taken to screen patients for tuberculosis, as there is a risk of TB reactivation (48)and a disseminated infection due to granuloma breakdown and loss of sentinel immune function.

However there remained a proportion of patients in whom blockade of TNFα does not result in an improvement in signs or symptoms and therefore targeting other key pathways became a necessity.

1.5.4 IL-1 Receptor Antibody

IL-1 levels are increased in patients with RA (49)and therefore it seemed logical to block this cytokine that is produced by inflammatory cells including macrophages and neutrophils.

Anakinra was found to be effective in RA (50)but only attained moderate improvement in ACR responses and high IL-1 levels were thought to be as a result of generalised immune activation as opposed to a pathogenetic problem. However this medication was re-appropriated into the IL-1 and cryopyrin associated conditions and it has also been effective for conditions that affect the NLRP3 inflammasome such as gout(51).

1.5.5 Blockade of the IL-6R

Tocilizumab is used in the treatment of RA and also other inflammatory arthropathies such as juvenile idiopathic arthritis. It was initially available as an intravenous infusion in patients who had failed TNFα blockers (52)but was then found to be efficacious in patients who were unable to tolerate methotrexate(53,54). Finally it is likely that the mechanism of action of tocilizumab is related to blocking the effect of IL-6 released from FLS (55) and also blocking IL-6 mediated B-cell stimulation(56).

1.5.6 Anti CD20 Antibody

Rituximab binds to CD-20 on B-cells but not plasma cells and was originally used as a treatment for non-Hodgkin's lymphoma and also chronic lymphocytic leukaemia(57). It was then repurposed into a treatment for RA although it does not have a license as a first line therapy and is only given following the failure of another biologic agent(58). In patients who respond to Rituximab, they often have a prolonged response following two infusions suggesting that longer term disease modification by B-cells depletion is crucial to disease pathogenesis(59).

1.5.7 Abatacept and blockade of T-cell co-stimulation

Abatacept is a fusion protein that blocks co-stimulation between CD28 on naïve T-cells and CD80/86 on antigen presenting cells. It is as effective as TNFα blockade in improving signs and symptoms of RA and is administered as a subcutaneous injection(60). Abatacept is currently being used in the APPIPRA trial to determine if co-stimulation blockade could prevent the onset of RA in those patients who have arthralgia and have already broken tolerance by developing antibodies to citrullinated peptides (http://www.isrctn.com/ISRCTN46017566).

1.5.8 Other cytokines

Proof of concept and Phase I trials have been performed in patients with RA for IL-15 (61)and IL-21 blockade (https://clinicaltrials.gov/ct2/show/NCT01208506) and although these treatments showed some efficacy, they have not been moved through to further development. GMCSF is targeted by mavrilimumab and this is currently in clinical trails to determine efficacy and safety(62).

Despite combination DMARD therapy and biologic agents, there is considerable unmet need in the treatment of RA. Patients have both primary and secondary non-response to agents and therefore the disease evolves or the immune system develops anti drug antibodies that neutralise the effects of biologic agents. Small molecule kinase inhibitors make attractive targets because they have the potential to target multiple pathways and therefore remain efficacious when the disease evolves. Furthermore small molecules are available as oral preparations and therefore may present a more attractive option to patients. There is evidence for cancer literature that patients value efficacy over ease of dosage and therefore if the agents are equivalent in terms of disease improvement then route of dosage plays a role in compliance(63).

1.5.9 Novel small molecule inhibitors

Inhibition of kinases showed efficacy in the field of oncology and in particular in the treatment of chronic myeloid leukaemia with imatinib mesylate(64). This drug targets the ATP binding site of the BCR-Abl oncogene and became one of the first bench-to-bedside successes for translational medicine.

Specifically in rheumatology, imatinib has been used to treat pigmented villonodular synovitis (PVNS) (65)in which there is a translocation of a collagen promoter juxtaposed to the MCSF gene in a minority of cells which results in a local production of MCSF and expansion of cells expressing MCSF receptor. The reason that a drug which was first developed for CML works in a condition where there is expansion of immune cells in synovitis is that kinase inhibitors tend to have multiple targets. The kinome is known (figure 1-3)(66) (67)and each kinase inhibitor can be mapped onto it (figure 1-2) with examples of commonly used inhibitors shown below where the diameter of the circle relates to the affinity with which a particular molecule binds a target.

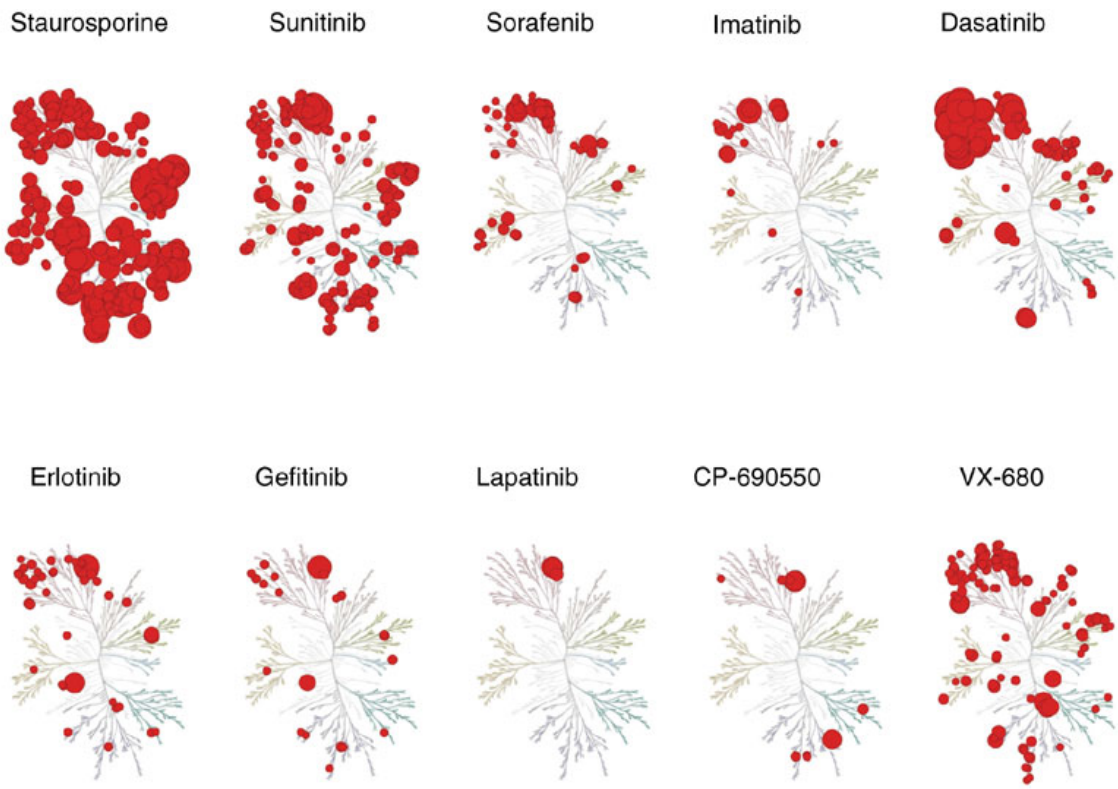


Figure 1-2: Dendrograms of the human kinome depicting target inhibition by various agents including tofacitinib (CP-690550). Kinase inhibitors often target multiple kinase pathways. Staurosporine is used to induced apoptosis of cells and is a pan-kinase inhibitor. Tofacitinib although targeting JAK also targets other kinases. Reproduced from (67)

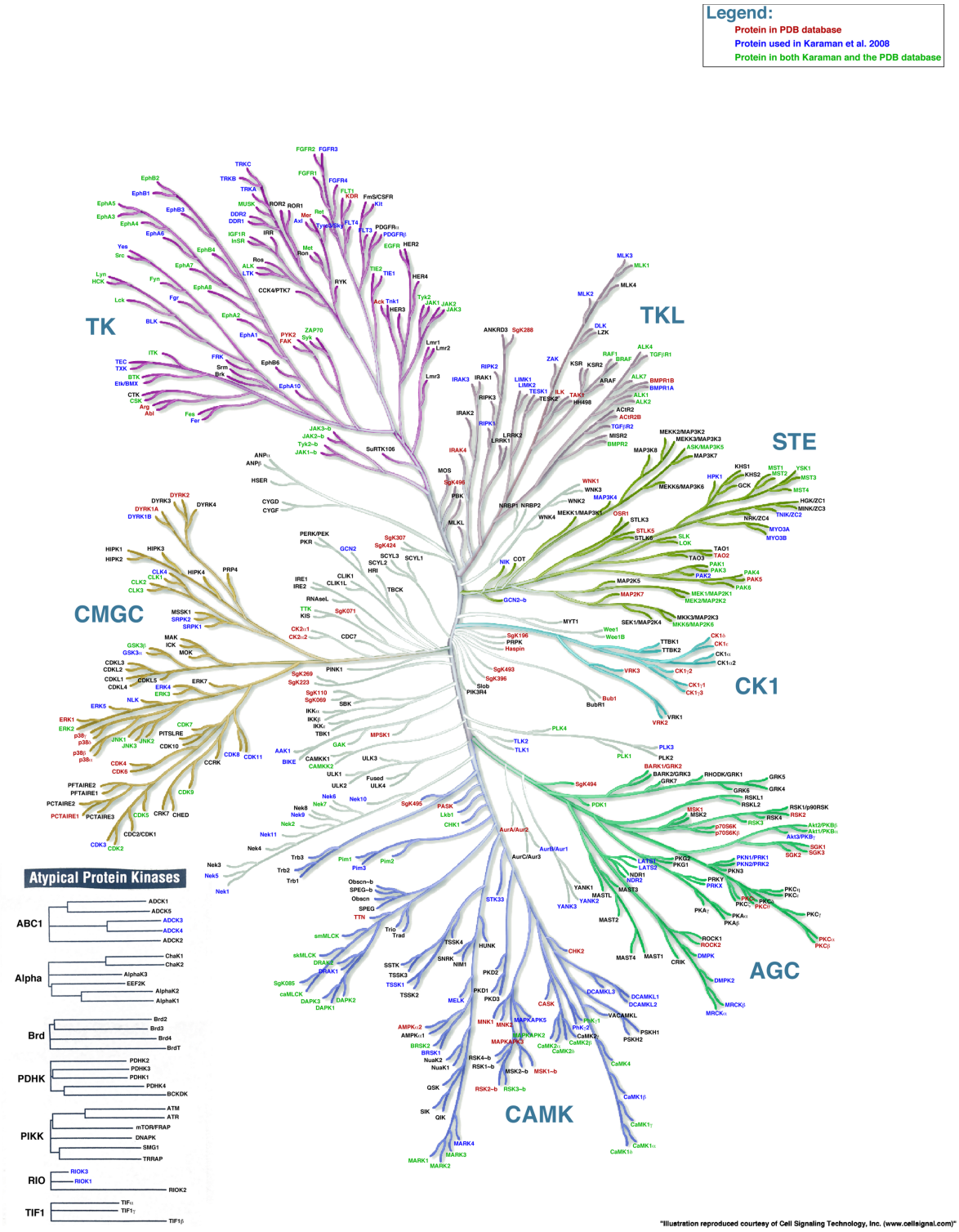


Figure 1-3: The human kinome reproduced from Kinome Render (68). Illustration reproduced courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com)

Although the JAK inhibitor tofacitinib has been successful in the treatment of RA when compared to methotrexate (69)and is useful in patients who are resistant to multiple biologic agents, not all kinase inhibitors have been so successful. Targeting of Syk appeared promising at the pre-clinical and early phases but later trials revealed inadequate efficacy and significant safety signatures(70). In the case of the Syk inhibitor fostamatinib, issues with a lack of efficacy were overshadowed by a poor safety profile with transaminitis, hypertension and hypercholesterolaemia being problematic(71). Some of these effects could be linked to off target effects of the drug and it has been hypothesised that a lack of specificity of target in fostamatinib resulted in failure.

The JAK inhibitor, tofacitinib is approved for the treatment of RA in the USA, South America, Australia and Japan and is available for use off-license in the UK. It was developed in partnership between the O'Shea lab, the NIH and Pfizer and resulted in a small molecule inhibitor with a novel mechanism of action. Tofacitinib was initially thought to target only JAK3 but further investigation showed it to have activity against JAK1/3 (72) and possibly JAK2 but no known activity against TYK2. However although tofacitinib was approved by the FDA, the dosage was limited because of concerns regarding side effects namely reactivation of infections, abnormalities in serum creatinine and also changes to circulating lipids in a patient population with a higher cardiovascular risk.

Furthermore, tofacitinib was thought to have it's main effect in lymphocytes given its ability to inhibit JAK3 however there is now evidence for a wider role in innate immune cells(73), fibroblasts(74), osteoclasts(75) and dendritic cells(76). Furthermore a recent study shows that tofacitinib treatment of white fat results in the generation of more metabolically active brown fat and therefore a possible link to treating obesity and the metabolic syndrome with kinase inhibitors(77).

Another JAK inhibitor, baricitinib, is in clinical trials for the treatment of RA (78)and also for the treatment of psoriasis and diabetic nephropathy. This JAK1/2 inhibitor has shown promise against TNFα blockade (78)and shows that JAK2 inhibition where once thought to be therapeutically impossible due to it's association with colony stimulating and growth factors, may indeed be crucial to treat RA.

Filgotinib is a selective JAK1 inhibitor from Galapagos that has 10-fold selectivity over Tyk2 and 30-50-fold selectivity over other JAK members. Results from the ACR 2015

conference showed promising results in both patients treated with and without methotrexate with ACR70 results at 24 weeks in the region of 30% in the former and 25% in the later. A higher dose of 100mg twice daily resulted in an ACR70 response rate of almost 40% when the drug was used alongside methotrexate. Furthermore, inhibition of JAK1 alone led to little in the form of safety signature with a small decrease in neutrophils and increase in creatinine. Furthermore, lymphocyte counts were not altered and HDL cholesterol improved over LDL suggesting that differential inhibition of JAKs results in significantly different lipid profiles.

Although JAK3 specific inhibition is theoretically desirable, a clinical trial of VX-509 or decernotinib showed significant safety signatures in the form of increased infection, derangements of liver transaminases and hypercholesterolaemia(79). The medication was moderately effective with ACR70 responses at 12 weeks in the region of 20% but only at the 150mg dose where side effects were prohibitive. Following this trial Vertex decided to halt further development.

Other kinase inhibitors which are in trial include Brutons tyrosine kinase which has shown efficacy in pre-clinical models of RA (80)and also the cyclin dependent kinase inhibitor seliciclib (81)may fill the unmet need of patients with RA who have a predominantly fibroid picture and for whom current biologic agents deliver little therapeutic benefit. In addition PI3 Kinase inhibition is being explored with specific inhibitors of the gamma and delta subtypes posing interesting therapeutic targets in the field of autoimmune rheumatic diseases(82,83).

1.6 Immunopathogenesis of RA

1.6.1 Genetics and lessons from GWAS

Early genome wide association studies in RA yielded some useful information regarding the genetic heritability of the condition(84,85). GWAS relies on large numbers to detect a small signal where a single nucleotide polymorphism confers an altered risk of developing a particular condition. Part of the heritability of RA was explored by using initial twin studies where up to 15% of monozygotic twins went on to develop RA when one twin had the condition compared to 4% of non identical twins(86,87). This also gives credence to the role of environmental factors although these could be acting as epigenetic factors such as smoking that alters methylation status of genes.

Early linkage studies showed that the MHC Class II locus was critical in RA with HLA-DR4 being associated with an increased risk of developing the disease. Furthermore this was expanded to include other specific HLA-DR alleles that encode a five amino acid sequence known as QKRAA or the shared epitope(9). Patients who are positive for the shared epitope are more likely to have an aggressive erosive course of disease and also require biologic therapy at an earlier stage(10). Furthermore work from disease registries has shown that the shared epitope and environmental factors such as smoking can act together in a synergistic fashion to further increase the risk of developing RA. This effect is reserved to patients who are seropositive and this makes biological sense where a defect in the antigen presentation apparatus of the adaptive immune system results in autoantibody production and further adaptive and innate activation in the form of CD4 T-cells and macrophages.

Early GWAS studies with small patient numbers confirmed these results but mainly in seropositive patients. Furthermore, the population from which the patients were drawn from tended to be of Caucasian and European ancestry hence conclusions about other populations could not be confidently drawn. In addition small linkage studies were emerging showing that other markers may have been important especially in other populations but these studies could not be replicated.

The first metanalysis of RA GWAS studies increased the power of these studies to detect smaller effects and other defects of the adaptive immune system were discovered. In particular PTPN22, a tyrosine phosphatase, was found to be significantly different in

patients with RA (88)and also components of the JAK/STAT pathway (89)that acts as a signal transduction pathway for many inflammatory cytokines was altered.

A recent metanalysis comprising over a 100000 subjects with almost 30000 cases and the rest controls has shown further insight into the genetics of RA and also the ability of computational biology teams to combine data from various sources and integrate their analysis to provide biological insight(90). This GWAS also included a significant number of individuals from Asian decent although many of these cases were from Japan and China and therefore Southeast Asia was not represented (figure 1-4).

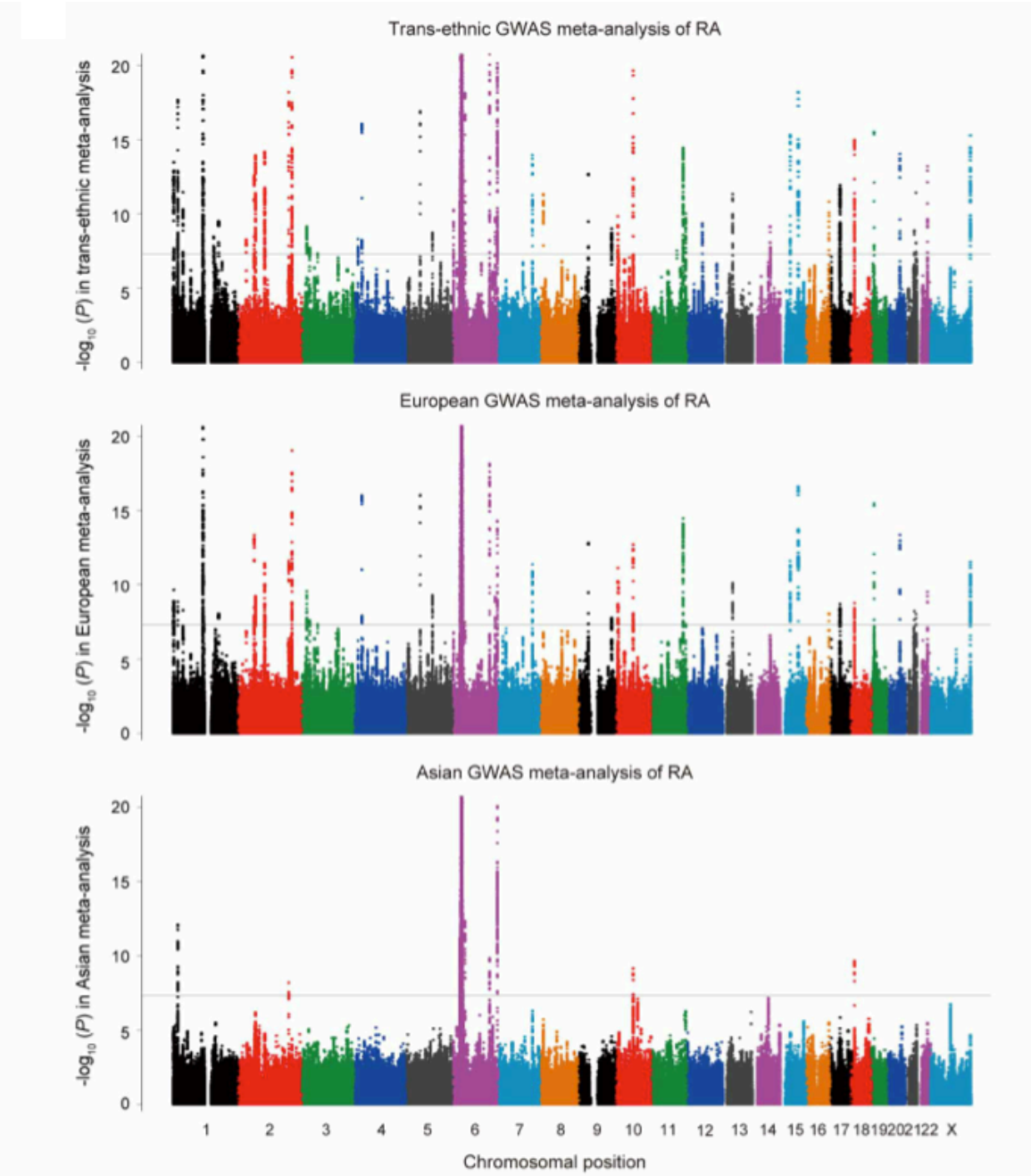


Figure 1-4 Manhattan plot reproduced from (6) showing the larger number of polymorphisms found when a trans-ethnic analysis is employed.

SNP	Chr	Genes	A1/A2 (+)	Trans-ethnic OR (95% CI)	Trans-ethnic P	European OR (95% CI)	European P	Asian OR (95% CI)	Asian P
rs227163	1	*TNFRSF9*	C/T	1.04 (1.02–1.06)	3.9×10^{-4}	1.00 (0.97–1.03)	9.3×10^{-1}	1.11 (1.08–1.16)*	3.1×10^{-9}*
rs28411352	1	*MTF1-INPP5B*	T/C	1.11 (1.08–1.14)*	2.8×10^{-12}*	1.10 (1.07–1.14)*	5.9×10^{-9}*	1.12 (1.06–1.19)	7.8×10^{-5}
rs2105325	1	*LOC100506023*	C/A	1.12 (1.08–1.15)*	6.9×10^{-13}*	1.12 (1.08–1.15)*	3.3×10^{-11}*	1.13 (1.04–1.23)	5.2×10^{-3}
rs10175798	2	*LBH*	A/G	1.08 (1.06–1.11)*	1.1×10^{-9}*	1.09 (1.06–1.12)*	4.2×10^{-8}*	1.07 (1.02–1.13)	6.4×10^{-3}
rs6732565	2	*ACOXL*	A/G	1.07 (1.05–1.10)*	2.7×10^{-8}*	1.10 (1.07–1.14)*	9.4×10^{-9}*	1.04 (1.00–1.08)	4.0×10^{-2}
rs6715284	2	*CFLAR-CASP8*	G/C	1.15 (1.10–1.20)*	1.8×10^{-9}*	1.15 (1.10–1.20)*	2.5×10^{-9}*	-	-
rs4452313	3	*PLCL2*	T/A	1.09 (1.06–1.12)*	1.6×10^{-10}*	1.11 (1.08–1.15)*	5.2×10^{-11}*	1.04 (0.99–1.09)	9.2×10^{-2}
rs3806624	3	*EOMES*	G/A	1.08 (1.05–1.11)*	8.6×10^{-9}*	1.08 (1.05–1.12)*	2.8×10^{-8}*	1.06 (0.99–1.14)	1.0×10^{-1}
rs9826828	3	*IL20RB*	A/G	1.44 (1.28–1.61)*	8.6×10^{-10}*	1.44 (1.28–1.61)*	8.7×10^{-10}*	-	-
rs13142500	4	*CLNK*	C/T	1.10 (1.07–1.13)*	3.0×10^{-9}*	1.10 (1.06–1.15)	2.4×10^{-6}	1.10 (1.04–1.15)	2.8×10^{-4}
rs2664035	4	*TEC*	A/G	1.07 (1.04–1.10)	9.5×10^{-8}	1.08 (1.05–1.11)*	3.3×10^{-8}*	1.03 (0.97–1.08)	3.3×10^{-1}
rs9378815	6	*IRF4*	C/G	1.09 (1.06–1.12)*	1.7×10^{-10}*	1.09 (1.05–1.12)	1.4×10^{-7}	1.10 (1.04–1.15)	2.3×10^{-4}
rs2234067	6	*ETV7*	C/A	1.15 (1.10–1.20)*	1.6×10^{-9}*	1.14 (1.09–1.19)*	4.1×10^{-8}*	1.22 (1.06–1.41)	7.0×10^{-3}
rs9373594	6	*PPIL4*	T/C	1.09 (1.06–1.12)*	3.0×10^{-9}*	1.07 (1.02–1.12)	6.5×10^{-3}	1.11 (1.07–1.15)*	4.8×10^{-8}*
rs67250450	7	*JAZF1*	T/C	1.10 (1.07–1.14)*	3.7×10^{-9}*	1.11 (1.07–1.14)*	2.6×10^{-9}*	1.02 (0.84–1.23)	8.5×10^{-1}
rs4272	7	*CDK6*	G/A	1.10 (1.06–1.13)*	5.0×10^{-9}*	1.10 (1.07–1.14)*	1.2×10^{-8}*	1.06 (0.98–1.15)	1.3×10^{-1}
rs998731	8	*TPD52*	T/C	1.08 (1.05–1.11)*	1.9×10^{-8}*	1.09 (1.06–1.12)*	6.6×10^{-9}*	1.02 (0.96–1.10)	4.9×10^{-1}
rs678347	8	*GRHL2*	G/A	1.08 (1.05–1.11)*	1.6×10^{-8}*	1.10 (1.06–1.13)*	7.3×10^{-9}*	1.03 (0.98–1.10)	2.6×10^{-1}
rs1516971	8	*PVT1*	T/C	1.15 (1.10–1.20)*	1.3×10^{-10}*	1.16 (1.11–1.21)*	3.2×10^{-11}*	-	-
rs12413578	10	10p14	C/T	1.20 (1.13–1.29)*	4.8×10^{-8}*	1.20 (1.12–1.29)	7.5×10^{-8}	-	-
rs793108	10	*ZNF438*	T/C	1.08 (1.05–1.10)*	1.3×10^{-9}*	1.07 (1.04–1.10)	6.1×10^{-7}	1.09 (1.04–1.14)	4.4×10^{-4}
rs2671692	10	*WDFY4*	A/G	1.07 (1.05–1.10)*	2.8×10^{-9}*	1.06 (1.03–1.09)	2.6×10^{-5}	1.10 (1.05–1.14)	9.9×10^{-6}
rs726288	10	*SFTPD*	T/C	1.14 (1.07–1.20)	1.6×10^{-5}	0.96 (0.86–1.06)	4.1×10^{-1}	1.22 (1.14–1.31)*	8.8×10^{-9}*
rs968567	11	*FADS1-FADS2-FADS3*	C/T	1.12 (1.07–1.16)*	1.8×10^{-8}*	1.12 (1.07–1.16)*	1.8×10^{-8}*	-	-
rs4409785	11	*CEP57*	C/T	1.12 (1.09–1.16)*	1.2×10^{-11}*	1.12 (1.08–1.16)*	3.6×10^{-9}*	1.16 (1.07–1.27)	4.3×10^{-4}
chr11:107967350	11	*ATM*	A/G	1.21 (1.13–1.29)*	1.4×10^{-8}*	1.21 (1.13–1.29)*	1.1×10^{-8}*	-	-
rs73013527	11	*ETS1*	C/T	1.09 (1.06–1.12)*	1.2×10^{-10}*	1.08 (1.05–1.11)	1.0×10^{-6}	1.14 (1.08–1.21)	4.1×10^{-6}
rs773125	12	*CDK2*	A/G	1.09 (1.06–1.12)*	1.1×10^{-10}*	1.09 (1.06–1.12)*	2.1×10^{-8}*	1.10 (1.04–1.17)	1.1×10^{-3}
rs10774624	12	*SH2B3-PTPN11*	G/A	1.09 (1.06–1.13)*	6.8×10^{-9}*	1.09 (1.06–1.13)*	6.9×10^{-9}*	-	-
rs9603616	13	*COG6*	C/T	1.10 (1.07–1.13)*	1.6×10^{-12}*	1.11 (1.07–1.14)*	2.8×10^{-11}*	1.08 (1.02–1.14)	1.0×10^{-2}
rs3783782	14	*PRKCH*	A/G	1.14 (1.09–1.18)*	2.2×10^{-9}*	1.12 (0.96–1.31)	1.4×10^{-1}	1.14 (1.09–1.19)*	4.4×10^{-9}*
rs1950897	14	*RAD51B*	T/C	1.10 (1.07–1.13)*	8.2×10^{-11}*	1.09 (1.06–1.12)*	5.0×10^{-8}*	1.16 (1.08–1.25)	1.1×10^{-4}
rs4780401	16	*TXNDC11*	T/G	1.07 (1.05–1.10)*	4.1×10^{-8}*	1.09 (1.06–1.13)*	8.7×10^{-9}*	1.03 (0.98–1.08)	2.5×10^{-1}
rs72634030	17	*C1QBP*	A/C	1.12 (1.08–1.17)*	1.5×10^{-9}*	1.12 (1.06–1.19)	2.9×10^{-5}	1.12 (1.07–1.18)	9.6×10^{-6}
rs1877030	17	*MED1*	C/T	1.09 (1.06–1.12)*	1.9×10^{-8}*	1.09 (1.05–1.13)	1.3×10^{-5}	1.09 (1.04–1.14)	3.2×10^{-4}
rs2469434	18	*CD226*	C/T	1.07 (1.05–1.10)*	8.9×10^{-10}*	1.05 (1.02–1.08)	6.7×10^{-4}	1.11 (1.07–1.15)*	1.2×10^{-8}*
chr19:10771941	19	*ILF3*	C/T	1.47 (1.30–1.67)*	8.6×10^{-10}*	1.47 (1.30–1.67)*	8.8×10^{-10}*	-	-
rs73194058	21	*IFNGR2*	C/A	1.08 (1.05–1.12)	1.2×10^{-6}	1.13 (1.08–1.18)*	2.6×10^{-8}*	1.03 (0.98–1.08)	2.9×10^{-1}
rs1893592	21	*UBASH3A*	A/C	1.11 (1.08–1.14)*	7.2×10^{-12}*	1.11 (1.07–1.15)*	9.8×10^{-9}*	1.11 (1.05–1.18)	1.3×10^{-4}
rs11089637	22	*UBE2L3-YDJC*	C/T	1.08 (1.05–1.11)*	2.1×10^{-9}*	1.10 (1.06–1.15)	2.0×10^{-7}	1.06 (1.02–1.10)	8.9×10^{-4}
rs909685	22	*SYNGR1*	A/T	1.13 (1.10–1.16)*	1.4×10^{-16}*	1.11 (1.08–1.15)*	6.4×10^{-12}*	1.23 (1.14–1.33)	2.0×10^{-7}
chrX:78464616	X	*P2RY10*	A/C	1.11 (1.07–1.15)*	3.5×10^{-8}*	1.16 (0.78–1.75)	4.6×10^{-1}	1.11 (1.07–1.15)*	3.6×10^{-8}*

SNPs newly associated with $P < 5.0 \times 10^{-8}$ in the combined study of the stage 1 GWAS meta-analysis and the stages 2 and 3 replication studies of trans-ethnic (Europeans and Asians), European or Asian ancestry are indicated. SNPs, positions and alleles are based on the positive (+) strand of NCBI build 37. A1 represents an RA risk allele. Chr, chromosome; OR, odds ratio; 95% CI, 95% confidence interval. Full results of the studies are available in Supplementary Table 1. Hyphens between gene names indicate that several candidate RA risk genes were included in the region.
*Association results with $P < 5.0 \times 10^{-8}$.

Table 1-9: Table of SNPs significant from trans-ethnic GWAS analysis. Reproduced from (6)

In total over 10 million single nucleotide were tested and these were combined with other analyses to gain further insight into the pathogenesis of RA. In the first instance, a large trans-ethnic metanalysis was carried out which revealed over forty risk loci just under half of which were novel. This was then replicated and following replication, the total number of risk loci increased to 101 with 100 being outside the MHC locus. In the trans-ethnic analysis, the loci outside MHC explained 5.5% of the heritability in Europeans and 4.7% in Asians. Furthermore the group then went on to use the following methods to gain further insights into the role of the genetic component in the pathogenesis of disease:

1. Enrichment of epigenetic chromatin marks at each loci in 34 different cell types by assessing trimethylation of Histone H3 at lysine 4 (H3K4me3). Methylation of lysine 4 in Histone 3 is associated with gene activation and therefore if chromatin marks are seen at these loci it gives further evidence that these loci are important in pathogenesis. In particular, significant enrichment was seen in the CD4+ Treg samples, a cell which has been linked to failure of resolution of inflammation.
2. Carrying out a *cis*- expression quantitative trait locus analysis using data from PBMC, CD4 T-cells and CD14+ CD16- monocytes. By combining expression data with SNPs you can interrogate whether a particular SNP has an effect on the expression of that particular gene.
3. Overlapping the genes in the data set with other autoimmune conditions, primary immunodeficiencies, cancer somatic mutation genes and knockout mouse phenotypes.
4. Finally pathways enrichment analysis was performed by text mining Pubmed and also determining which genes had a potential protein-protein interaction.

Using a combination of these methods, genes associated with the SNPs were given a score that was higher if the gene was present in multiple analyses outlined above. The approach led to a list of genes that were then further examined to determine whether the gene was currently associated with RA pathogenesis or in the case of genes that were identified in protein interaction networks, whether they were already therapeutic targets.

RA risk SNP (cytoband)	Gene	Score	Biological gene criteria							Nearest gene from RA risk SNP	Overlap with H3K4me3 peaks										Drug target gene	RA drug target gene	PPI with RA drug target gene
			RA risk missense variant	*cis*-eQTL	PubMed text mining	PPI	PID	Haematological cancer	Knockout mouse phenotype	Molecular pathway	T_{reg} primary cells	CD4+ memory primary cells	CD4+ naive primary cells	CD8+ memory primary cells	CD8+ naive primary cells	CD34+ primary cells	CD34+ cultured cells	Mobilized CD34+ primary cells	CD19+ primary cells	CD3+ primary cells			
chr1:2523811 (1p36)	*TNFRSF14*	4			■			■	■		■	■	■	■	■				■	■			■
rs2301888 (1p36)	*PADI4*	2	■							■													
rs2476601 (1p13)	*PTPN22*	5	■	■	■	■			■	■								■					
rs2228145 (1q21)	*IL6R*	5	■	■	■	■				■	■		■	■	■					■	■	■	■
chr1:161644258 (1q23)	*FCGR2B*	5		■	■			■	■	■	■										■		
rs17668708 (1q31)	*PTPRC*	6		■	■	■	■		■	■	■	■	■	■	■		■	■	■	■			■
rs34695944 (2p16-p15)	*REL*	4			■	■		■	■		■	■	■	■	■	■	■	■	■	■			
rs9653442 (2q11)	*AFF3*	4		■	■	■		■															
rs11889341 (2q32)	*STAT4*	3			■				■	■	■	■		■	■								■
rs6715284 (2q33)	*CFLAR*	3			■				■	■													■
rs1980422 (2q33)	*CD28*	4		■	■				■	■	■	■	■	■	■			■		■	■		■
rs3087243 (2q33)	*CTLA4*	4			■	■			■	■	■	■									■		■
rs45475795 (4q26-q27)	*IL2*	5			■	■		■	■	■	■	■	■	■	■	■		■					
rs657075 (5q31)	*IL3*	4			■	■			■	■													■
rs657075 (5q31)	*CSF2*	4			■	■			■	■	■												■
rs2233424 (6p21)	*NFKBIE*	4	■		■	■			■		■	■	■	■	■	■	■	■	■	■			■
rs7752903 (6q23)	*TNFAIP3*	6	■		■	■		■	■	■	■	■	■	■	■	■	■	■	■	■			■
rs1571878 (6q27)	*CCR6*	2			■				■		■	■	■	■					■				
rs4272 (7q21)	*CDK6*	4			■			■	■	■	■	■									■		
chr7:128580042 (7q32)	*IRF5*	4		■	■	■			■		■	■	■	■		■	■	■	■	■			
rs10985070 (9q33)	*TRAF1*	4		■	■				■		■	■	■	■	■	■	■	■	■	■			
rs10985070 (9q33)	*C5*	4		■	■		■			■	■	■	■	■	■	■	■	■	■	■	■		
rs706778 (10p15)	*IL2RA*	5			■	■			■	■	■	■									■		■
rs331463 (11p12)	*TRAF6*	4			■	■			■	■	■	■	■	■	■		■	■	■	■			■
rs331463 (11p12)	*RAG1*	4			■	■	■		■		■	■	■	■	■		■	■	■	■			
rs508970 (11q12)	*CD5*	4	■		■	■			■		■	■	■	■	■			■		■			
chr11:107967350 (11q22)	*ATM*	4					■	■	■	■													■
rs773125 (12q13)	*CDK2*	3				■			■	■	■	■	■	■	■	■	■	■	■	■	■		
rs1633360 (12q13-q14)	*CDK4*	3			■				■	■	■	■	■	■	■	■	■	■	■	■	■		
rs10774624 (12q24)	*SH2B3*	5	■	■	■	■			■		■							■					
chr17:38031857 (17q12-q21)	*IKZF3*	4		■	■				■	■		■	■	■	■	■	■	■	■	■			
chr17:38031857 (17q12-q21)	*CSF3*	4			■	■			■	■		■	■	■	■	■	■	■	■	■			■
rs8083786 (18p11)	*PTPN2*	3				■			■	■	■	■	■						■				■
rs34536443 (19p13)	*ICAM1*	4			■	■			■	■											■		
rs34536443 (19p13)	*TYK2*	6	■		■	■	■		■	■													■
rs4239702 (20q13)	*CD40*	6		■	■	■	■		■	■	■	■	■	■	■	■	■	■	■	■	■		■
rs73194058 (21q22)	*IFNGR2*	6		■	■	■	■		■	■	■	■	■	■	■			■			■		■
rs2236668 (21q22)	*ICOSLG*	5	■	■	■				■	■	■												
rs2236668 (21q22)	*AIRE*	4			■	■	■		■														
rs3218251 (22q12)	*IL2RB*	3			■	■				■	■	■	■	■	■						■		■
rs5987194 (Xq28)	*IRAK1*	3	■		■				■		■	■	■	■	■		■	■	■	■			■

Prioritized biological RA risk genes. Representative biological RA risk genes. We list the summary gene score derived from individual criteria (filled red box indicates criterion satisfied; 98 genes with a score ≥2 out of 377 genes included in the RA risk loci were defined as 'biological candidate genes'; see Extended Data Fig. 6). Filled blue boxes indicate the nearest gene to the RA risk SNP. Filled green boxes indicate overlap with H3K4me3 peaks in immune-related cells. Filled purple boxes indicate overlap with drug target genes. For full results, see Supplementary Table 5.

Figure 1-5 Table of RA risk genes discovered on trans ethnic GWAS and which source of data each satisfies. Reproduced from (6)

From the prioritized genes plot we can see that PTPN22 is present along with members of the JAK and STAT pathway, T-cell co-stimulation proteins CD28 and CTLA-4 as well as IL6R. Furthermore PADI4 has been linked with the process of citrullination of native proteins and therefore the development of ACPA leading to a break in tolerance.

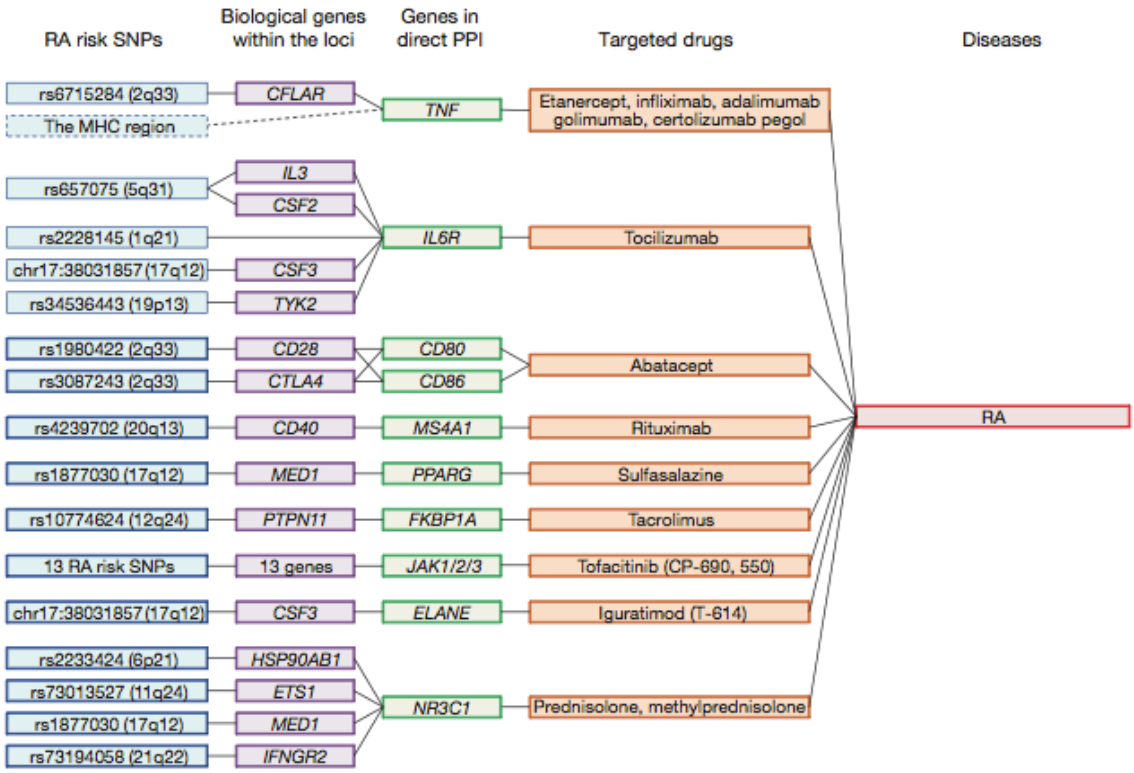


Figure 1-6 SNPs linked via protein-protein interaction networks to current treatments for RA. Only IL6R is linked directly to a treatment. Reproduced from (6)

Notably many of the SNPs could be linked to current therapies in RA but only the SNP in IL6R was directly linked to a treatment namely, tocilizumab. In all other cases, the link was through an extended protein-protein interaction network and may account for one of the reasons why discovering druggable targets via genetic studies has been difficult.

In addition this study led to the discovery of SNPs in cytokine pathways and T and B cell pathways that were not previously shown in earlier GWAS. This has led to the situation where GWAS has validated potential druggable pathways after we have already been using therapies in the field. However it does validate the approach of GWAS and some 40 novel loci were not prioritised as having potential drug targets and further fine mapping is required to determine which genes are potentially responsible for the altered risk.

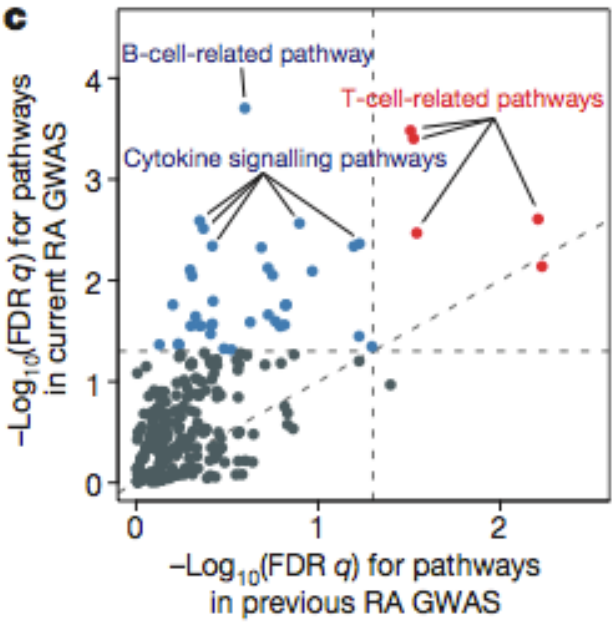


Figure 1-7: Plot of -$\log_{10}$(FDR q values) for pathways in the current versus previous RA GWAS. B cell, T cell and cytokine signalling pathways were significantly changed in the current GWAS. Reproduced from (6)

Therefore this metanalysis of GWAS in RA patients with trans-ethnic origins has shown that genes involved with cytokine signaling, B and T-cell pathways and the JAK/STAT pathway are implicated in the pathogenesis of RA. Some of this information is novel but in other cases therapeutics are already in the field and have been used for decades in the treatment of RA.

1.7 Environmental factors

1.7.1 Smoking

Cohort studies from the Karolinska Institute elegantly demonstrated the interaction between genetics and smoking where having HLA-DR4 positivity and PTPN22 along with being a current smoker increased the risk of developing RA dramatically(10). Furthermore Rheumatoid Factor and ACPA are found more commonly in people who smoke and may predate the onset of arthritis by many years(33).

ACPA are normally raised to structural components such as fibrinogen, fibronectin and α-enolase and studies of the airways in early RA have demonstrated inflammation of the bronchial tree suggesting that the break of tolerance may occur at the mucosal barrier(18).

Other agents such as coal dust and silica have also been implicated in the pathogenesis of RA and the insult of tar, hydrocarbons and free radicals will cause significant disruption to the mucosal barrier and alter self proteins by altering PADI4 an enzyme implicated in the process of citrullination and therefore the formation of ACPA(91). In addition, the discovery of the aryl hydrocarbon receptor has given further credence that cigarette smoke and other sources of hydrocarbons can directly influence RA fibroblast activity(92).

1.7.2 Infectious Agents

Evidence for the involvement of bacterial and viral infectious agents such as Proteus and E.coli as well as cytomegalovirus and Epstein-Barr virus have been in the literature for some time. The hypothesis in these cases is that the immune response against the pathogen results in a break of tolerance to self through molecular mimicry in a fashion similar to Gullian-Barre Syndrome and gastrointestinal bacterial infections.

However alphavirus species such as Chikungunya, which are transmitted by mosquitoes, can cause a debilitating arthritis following the initial infection(93). Furthermore Chikungunya has been spreading from Oceana through Africa and there have been cases in the North of Africa and southern Europe. With air travel becoming more common, this problem is being intensified. The virus itself was thought to be able to cause the development of ACPA but it may be that in those people who are susceptible to developing RA through their genetics and adverse environmental factors that the virus initiates the

onset of synovitis. Without these other risk factors, a synovitis may also occur, hence the name which means "to become contorted" in the Kimakonde language, but is self limiting and resolves with conservative measures.

1.7.3 The microbiome

Although the pathogenesis of RA is linked to genetics and environmental factors, archaeological excavations from mid and southern states of USA showed that bones exhibited pathological changes, which were in keeping with erosions from rheumatoid arthritis(94). Furthermore, these findings pre-dated the description of rheumatoid arthritis in the Europe by thousands of years suggesting that the disease itself may have been communicated from the Americas.

No one infectious agent has been implicated in the pathogenesis of RA but the role of the human microbiome is likely to be significant. We have formed symbiotic relationships with non-pathogenic bacteria and this adds another level of complexity to the pathogenesis of complex diseases. Therefore alterations to the human microbiome may be implicated in disease onset, perpetuation and also failure of resolution.

Molecular mimicry is proposed as one potential mechanism by which bacteria perpetuate the pathogenesis of RA. However organisms such as *P.Gingivalis* that are associated with periodontitis and cause inflammation in the periodontal area can express PAD(95). However the literature surrounding the role of these species in the break of tolerance in RA is still debated.

Evidence from other inflammatory arthropathies has shown the presence of a gut-joint-skin axis that may be influenced by the microbiome. In a rat model of ankylosing spondylitis, administrations of antibiotics to sterilize the GI tract resulted in the prevention of disease pathology(96). This suggests that the microbiome itself may have an influence on the immune system especially through Th_{17} cells that (97)are implicated in spondyloarthropathies and related conditions such as inflammatory bowel disease.

Shotgun sequencing has provided microbiome researchers with the technology to measure the microbiome. The 16s ribosomal RNA gene is conserved between species and therefore by sequencing this gene one can determine the relative abundance of microbial species at a particular site and also their diversity. A recent study has sequenced the oral microbiome

teeth, saliva as well the faecal microbiome in patients with RA(98). Perturbations in the microbiome were found between patients with RA and healthy controls and these changes reversed with treatment. Also, certain species such as *Haemophilus* and *Lactobacilus* were over represented in patients with RA and especially in the case of the latter, were found more often in those with active RA. In conclusion, the microbiome is an area of considerable unmet need and should be considered as another factor in the pathogenesis of RA.

1.8 The synovial microenvironment and lessons from pathology

The future of treatment of RA lies in the realm of stratified medicine and getting the right treatment to the right patient at that right time. In cancer and other conditions such as infection this can be done by biopsy or culture of a specific organism in order to determine which particular mutation is driving pathology or which organism is responsible for infection and also what it is sensitive to in order to determine optimal treatment.

Therefore having a pathologic sample means that the molecular diagnosis is sound especially in diseases which seem to be homogenous e.g. lung cancer and also that there is likely to be a particular response to treatment in the case of determining sensitivities to an infectious organism. The role of the synovial biopsy in RA is more controversial and is dependent on the acceptability of this diagnostic by the patient population, the provision of equipment and resource to carry this out and also whether the information which is gained from this procedure outweighs the risk of the procedure.

The reason for synovial biopsy not being carried out as widely as a biopsy in cancerous conditions is that at a population level, people respond well to our treatment algorithm of early DMARDs, combination therapy and then biologic therapy. However on a patient level, the response rates for the population are not important, you either do or do not respond to a particular treatment. Therefore the goal of stratified medicine is to gain the best treatment outcome for the patient and therefore also improve treatment outcomes for the population and decrease the burden of disease for society.

Many cell types have been implicated in the pathogenesis of RA however there are elements of the innate immune system along with the adaptive immune system which are critical in the pathogenesis of disease and this is demonstrated by the clinical effectiveness of treatments targeting macrophages in the form of anti-TNFα, B-Cells with rituximab and T-Cells with abatacept.

A retrospective study(99) of synovial biopsies have revealed at least four different subtypes of pathology:

1. Lymphoid predominant with evidence of ectopic lymphoid follicles of B and T cells
2. Myeloid predominant

3. Low inflammatory
4. Fibroblast dominant

Furthermore in this retrospective study, the myeloid phenotype was associated with a better response to anti-TNFα therapy. In addition the STRAP trial (http://www.matura-mrc.whri.qmul.ac.uk/) , a part of the MRC-MATURA programme to deliver stratification in RA, is going to correlate patient outcome with biopsy results when they receive rituximab, etanercept or tocilizumab. The outcome of the trial is to determine whether rituximab is a poorer treatment option in those patients who have a B-cell poor biopsy at baseline.

1.8.1 Myeloid lineage cells

Myeloid lineage cells from monocytes to macrophages, dendritic cells and osteoclasts have a central role in the pathogenesis of RA(100,101). These cells provide an important link from the innate to the adaptive immune system with dendritic cells(DCs) and macrophages able to present antigen. They are able to respond to pattern and damage related signals and are able to stimulate T-cells and B-cells.

Furthermore subsets of monocytes with varying levels of CD14 and CD16 expression levels may be associated with an inflammatory or resolving phenotype(102). CD14+ CD16+ monocytes have been shown to be pro-inflammatory but this population has been subdivided into $CD14^{dim}$ CD16+ and $CD14^{bright}$ CD16+ with the bright population being more frequent in the peripheral blood of patients with RA(103,104). Furthermore, the bright population had higher MHC Class II expression, higher levels of CCR5 and also secreted higher concentrations of TNFα when stimulated with T-cells.

Tissue macrophages express the CD68 marker and the presence of CD68 + macrophages has been shown to be a biomarker of therapeutic response(101). Further, there is evidence that these cells are derived from circulating monocytes as part of the mononuclear phagocytic system(105,106). Macrophages have also been subdivided into M1 and M2 type cells with the M1 being more inflammatory and related to an interferon or LPS stimulus and M2 being driven by a variety of stimuli such as IL-4, IL-13 or prostaglandins. Conventionally, macrophages that have been derived from CD14+ monocytes are cultured with MCSF to promote their differentiation and then used in further experiments. These

cells have higher levels of CD64+ and also MHC Class II expression that demonstrates their differentiation. However, these cells are plastic and can be differentiated to an M1 or M2 phenotype dependent on their environmental conditions(107). Therefore, the infiltrating macrophage can change its phenotype based on the phase of the inflammatory process. Macrophages are also a significant source of inflammatory cytokines such as TNFα, IL-1, 6, 12 and 15 as well as matrix metalloproteinases.

However, evidence exists to support the argument that tissue resident cells can arise from the embryonic yolk sac or foetal liver(108,109). In addition to tissue resident macrophages, there is evidence for both conventional and plasmacytoid DCs in the pathogenesis of RA(110-113).

Dendritic cells are described as professional antigen presenting cells and are able to activate naïve T cells and initiate effector T cells responses. However dendritic cells are also critical in the induction and maintenance of peripheral T cell tolerance (114)where autologous T-cells have escaped central tolerance induction from the thymus.

1.8.2 T cells

Abatacept which blocks T-cell co-stimulation is an effective treatment for DMARD resistant RA and therefore demonstrates the role of T-cells in pathogenesis and continued inflammation in RA(115). Furthermore, T-cells are present in both the synovium of patients with active RA and are present in ectopic lymphoid follicles(99). Furthermore, these cells are present in synovial fluid drawn from the joints of patients with active RA (116)and are able to induce inflammatory cytokine production from MCSF macrophages(117).

CD4+ T cells have also been implicated through GWAS with PTPN22, CTLA-4, IL-2, TNFAIP3 and c-REL being over represented in patients with RA(6). A conventional Th_1 CD4+ cell is a significant player in the pathogenesis of RA. IFNγ in response to activation by an APC then goes on to cause further activation of TH_0 cells, in an autocrine loop, to promote their maturation to Th_1 cells. Further, macrophages are stimulated to secrete further pro-inflammatory cytokines, such as TNFα and metalloproteinases that cause cartilage and bone destruction. Th_2 cells are also responsible for initiating an antibody response from B-cells and therefore have a role in the development of RF and ACPA(118).

Th_{17} cells are also implicated in RA with macrophages and dendritic cells providing the necessary transforming growth factor β, IL-1β, IL-6, IL-21 and IL-23(119). Furthermore, regulatory T-cells (Foxp3+) seem to have a limited function in RA patients and therefore suggest that there is an imbalance between Th_{17} and regulatory T-cells that drives RA(120).

As well as implication from GWAS and also histology from synovial biopsy, the CD4 T-cell is crucial in the macrophage: Tck cell contact activation model that demonstrated that an antigen specific response was not crucial in activating a macrophage(121). Although an initial break of tolerance is required in order to develop an antibody response in the form of RF and ACPA, CD4 T-cells, once activated by cytokines such as TNFα, IL-6 and IL-15, can perpetuate the chronic immune response.

The role of CD8+ T cells is not clearly defined in RA although recent evidence suggests that patients who have active RA may have an effector phenotype compared to controls and that this populations secretes more pro-inflammatory cytokines(122). However with treatment and a reduction in disease activity, these cells persist although their cytokine secretion profile changes suggesting that they are being stimulated by external ligands.

1.8.3 Other cells of relevance

1.8.4 B Cells

The presence of RF and ACPA demonstrates the role of B lineage cells in RA. Furthermore, rituximab, which depletes cells expressing CD20, is a licensed and effective therapy for RA(58). Although B cells are crucial in the production of antibodies, they also have a significant role in antigen presentation, activating T cells and are a source of local cytokines. B-cells are also a source of RANKL and may contribute to osteoclastogenesis(123). Finally B_{reg} cells may be important in the pathogenesis of RA but further work needs to be carried out(124).

1.8.5 Neutrophils

Both synovial fluid and the synovial lesion in RA contain large number of neutrophils, which were initially thought to be reactive(125). These cells tend to be involved with acute inflammation and are responsible for the “pus” seen in acute bacterial infections.

Although they are present in the synovial lesion, their presence was thought to be due to chemotactic stimulation as well as activation by integrin families and immunoglobulin. Furthermore, the synovial microenvironment prevents the apoptosis of neutrophils and this has been demonstrated when synovial fluid is used to co-culture neutrophils with fibroblasts and endothelial cells(126,127).

Neturophils can also form NETs or neutrophil extracellular traps and these are comprised of chromatin in a framework that is able to capture and present autoantigens, immunostimulatory molecules as well as components of neutrophil granules. Specifically in RA, NETs have been shown to present autoantigen, in particular, citrullinated peptides and these NETs are present in the synovial lesion(128). Production of NETs was induced by autoantibodies to citrullinated peptides and was augmented by proinflammatory cytokines such as TNFα. Furthermore, NETs from RA neutrophils were able to stimulate synovial fibroblasts to not only secrete pro-inflammatory cytokines such as TNFα but also to produce chemotactic factors such as IL-8. Therefore from a bystander cell, the neutrophil plays an important role where it responds to the altered citrullinated environment and then is able to perpetuate that response by presenting the autoantigen along with stimulatory molecules to other parts of the immune system.

1.8.6 NK Cells

NK cells may be activated in the synovial inflammatory process and there is evidence for a subpopulation present in the synovium of patients with RA(129). They are activated through NK receptor signalling and are a rich source of cytokines including but not limited to IFNγ, TNFα and IL-15 and could therefore contribute to RA pathogenesis through both cytokine secretion and cellular activation.

1.8.7 Stroma and Structure

1.8.8 Fibroblasts

Fibroblasts and specifically RA fibroblast like synoviocytes (FLS) make up an important component of the synovium as well as the pathologic pannus lesion that is the hallmark of RA(130). In a recent study of phenotypic subtypes of disease on synovial biopsy from

patients with RA, a fibroid or fibroblast rich subtype was described(99). This phenotype was less likely to respond to biologic therapies such as TNFα. This subgroup was less likely to show B or T cell aggregates and was also more likely to show a higher angiogenesis score based on a gene profile. These patients are also less likely to be seropositive for rheumatoid factor or ACPA and the findings of this study are echoed in the clinical sphere where we often find that patients who are seronegative tend to cycle through biologics and although there may be some response, there is often continual low-grade inflammation.

However FLS are not just structural components but are a rich source of pro-inflammatory cytokines and also enzymes such as matrix metalloproteinases which degrade surrounding stroma and result in joint destruction. FLS have the ability via expression of adhesion molecules, proteinases and cathepsins to interact with the synovial lining and cartilage to invade and destroy stroma. Furthermore, via TNFα and the production of RANKL(131), they can stimulate osteoclastogenesis thereby further contributing to joint destruction.

The synovial phenotyping study was based on gene set analysis of microarray data and one such gene set was associated with angiogenesis. FLS secrete factors such as VEGF that promote neo-angiogenesis (132)and via adhesion molecule expression can facilitate inflammatory cell infiltrate(133). A recent study from Zurich also demonstrated that FLS could migrate from one joint to another over longer distances thereby demonstrating that RA may begin on one joint and then "spread" to another(134). Finally, FLS can interact with T cells, macrophages, and osteoclasts and therefore co-ordinate the immune response to inflammatory arthritis.

1.8.9 Osteoclasts and Osteoblasts

Osteoclasts are generated from myeloid precursor cells that are either circulating or released from the bone marrow adjacent to sites of inflammation. MCSF and RANKL are found in the inflamed joint and are secreted from synovial fibroblasts and activated T cells and therefore result in increased osteoclastogenesis and erosions. ACPA have been demonstrated to further activate osteoclasts and therefore a mechanism for the clinical finding of a more destructive arthritis in those patients who are seropositive has been proposed(135).

In addition, osteoblasts act in balance with osteoclasts to lay down new bone and therefore keep bones in optimal health. Bone health relies on multiple pathways such as the Wnt/β-catenin pathways involved in osteoblastogenesis and the Dickkopf-1 (136)and sclerostin pathways which are negative regulators. Furthermore other molecules such as LRP-5 and Kremen-1 are involved but their role is still unclear. A recent GWAS of RA patients with erosive progression revealed single nucleotide polymorphisms in Dickkopf-1, sclerostin, LRP-5 and Kremen-1 with serum Dickkopf-1 levels associated with a polymorphism(137).

1.8.10 Chondrocytes

Chondrocytes make up articular cartilage that covers the joint surface of bones. An early change in RA is joint space narrowing on plain radiographs and therefore cartilage and therefore chondrocyte loss is part of the pathogenesis of RA. Chondrocytes produce collagens giving strength to the cartilage and along with proteoglycans such as aggrecan facilitate the retention of hyaluronic acid that gives the cartilage resistance to compression.

The inflammatory and invasive pannus lesion invades cartilage but chondrocytes themselves can be activated by TNFα to produce and secrete MMPs thereby increasing the erosive insult although the majority of erosion is proposed to be due to a combination of fibroblast like synoviocytes, activated macrophages and osteoclasts(138-140).

1.8.11 Cytokines and soluble factors in RA

Cytokines are crucial in the management of a coordinated immune response to pathogens in order to prime the system, execute killing of a pathogen and also resolve the inflammatory response. Deficiency of cytokine regulators such as IL-1RA lead to conditions where inflammation is a hallmark demonstrating the importance of this system(141).

Many cytokines are implicated in the pathogenesis of RA with TNFα and IL-6 being the most widely implicated at a pathologic level and also by the response of patients with active disease to treatments which neutralise these mediators(41,42,53,142,143). TNFα plays a role in perpetuating disease by activating endothelial cells, fibroblasts and other cell types to allow access to the joint by other immune cells and also stimulating the

release of further pro-inflammatory cytokines and mediators to perpetuate the inflammatory response.

TNFα also plays a crucial role in the generation of osteoclasts from precursors and therefore contributes significantly to the formation of the erosive pannus, the hallmark pathology of RA(144). As well as leading to erosions, joint damage and therefore disability, the increased activity of osteoclasts leads to the disease associated periarticular osteopenia and generalized osteoporosis.

TNFα signals via the NF-κB pathway mainly but also has links to the JAK/STAT pathway. However inhibition of the NF-κB pathway to date has not been therapeutically useful perhaps demonstrating the ubiquitous nature of this pathway. Therefore this suggests that TNFα while being crucial in the pathogenesis in the disease of a proportion of patients with RA, is not critical to host defence to bacterial pathogens. A caveat to this is in the case of intracellular mycobacterial infections where early treatment with TNFα blockers was associated with tuberculosis infection due to the breakdown of the granuloma(145).

The role of TNFα inhibition has also been demonstrated in other diseases such as PsA (146)and also Crohn's disease(147). Thus TNFα was thought to be a master cytokine responsible for coordinating aspects of autoimmunity. However although it has a crucial role in autoimmunity, it fails in other conditions such as gout, Familial Mediterranean Fever and other auto inflammatory disorders.

IL-6 receptor blockade is an effective therapeutic option in RA although patients initially had to undergo intravenous infusions of this medication as opposed to subcutaneous injection(53). Therefore this medication tended to be used for those who had failed TNFα inhibition although there is significant evidence that IL-6 blockade is superior to TNFα inhibition in those patients with RA who are unable to tolerate methotrexate.

IL-6 is secreted by macrophages in the synovial environment but is also produced by the fibroblast(74) and has a role in stimulating B-cells to produce antibodies. Therefore this cytokine has a role in linking the stromal component of the disease with the innate and adaptive components to RA.

Both TNFα and IL-6 are also responsible for the systemic components of disease including anaemia, cachexia and raised inflammatory markers. Systemic IL-6 stimulates the liver to

produce CRP and therefore it is most often used when there is a high systemic inflammatory burden.

IL-1 although present in both the synovium and synovial fluid of patients with RA is not a therapeutically useful target with Anakinra failing in patients with RA. Despite this, IL-1 blockade is useful in conditions where the IL-1 and IL-18 axis is deregulated such as in neonatal onset multi-inflammatory disorder or Muckle-Wells Syndrome(148).

Colony stimulating factors are upregulated in macrophages from patients with inflammatory arthritis and in particular targeting GM-CSF is being assessed in a Phase II clinical trial(62). Macrophages treated with GMCSF release more TNFα in response to stimulation and therefore targeting this growth factor was therapeutically viable although concerns regarding cytopenias and pulmonary alveolar proteinosis required significant safety assssments(149).

1.8.12 Implication of the kinases and specifically the JAK/STAT pathway in RA

Kinases phosphorylate proteins that lead to activation of a signal transduction pathway from ligand receptor to transcription factor activation and downstream gene transcription. Many kinases have been implicated in the pathogenesis of RA (150)and are attractive targets because of the ability to target them with small molecules as opposed to biologic therapies.

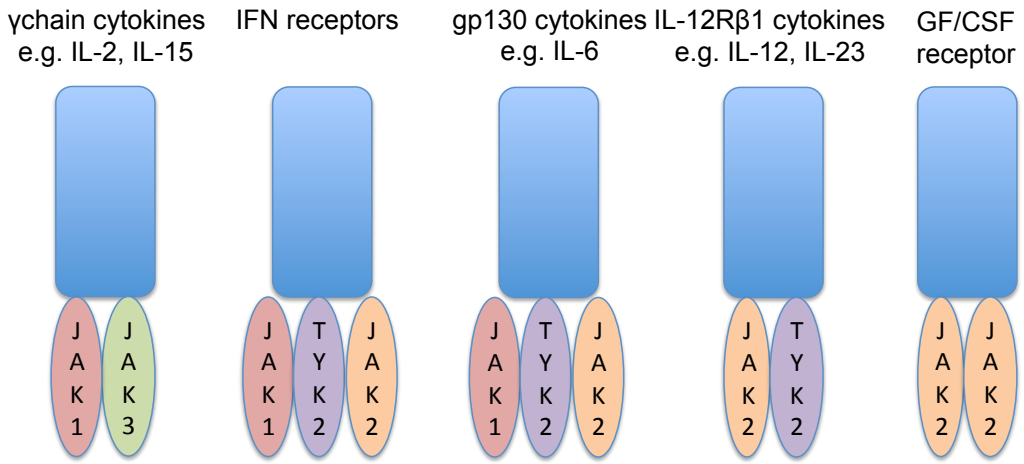
Also, there is often a degree of redundancy in the kinome and therefore inhibiting a class of kinases is theorized to lead to a greater therapeutic potential than just targeting a single cytokine or receptor. Current and previous targets for kinase inhibition in RA include NF-κB, Syk and Btk.

Tofacitinib targets JAK signaling and is the only kinase inhibitor licensed for the treatment of RA(151,152). This medication targets the JAK/STAT pathway which is implicated in RA from genetic studies(90), in vitro studies (153-155)and finally from the clinical effectiveness of the molecule.

The Signal Transduction and activation of Transcription (STAT) proteins transduce signals from various cytokines and growth factors(156,157). In contrast to receptor tyrosine kinases, Janus kinases are associated with the cytokine receptor and on ligand binding become activated by phosphorylation. Thereafter they proceed to phosphorylate the cytokine receptor that leads to the creation of STAT docking sites and phosphorylation of the STAT proteins themselves(158).

STATs then form homo or hetero dimers and translocate to the nucleus and bind to DNA and activate transcription of target genes. STATs are also regulated once activated by protein tyrosine phosphatases such as PTPN22, the suppressor of cytokine signaling (SOCS) group of proteins(159) and also protein inhibitors of activated STATS (PIAS)(160).

The four members of the JAK family are JAK1, JAK2, JAK3 and TYK2 with all being ubiquitously expressed except for JAK3 that is limited to lymphoid cells. Furthermore, each JAK is activated by a particular ligands and also forms homo or heterodimers with particular STAT proteins. In theory this makes studying the JAK/STAT system incredibly complicated due to the number of potential players involved and also redundancy within the system.



	γchain cytokines e.g. IL-2, IL-15	IFN receptors	gp130 cytokines e.g. IL-6	IL-12Rβ1 cytokines e.g. IL-12, IL-23	GF/CSF receptor
JAK1 inhibitor	✔	✔	✔	✗	✗
JAK2 inhibitor	✗	✔	✔	✔	✔
JAK3 inhibitor	✔	✗	✗	✗	✗
TYK2 inhibitor	✗	✔	✔	✔	✗

Figure 1-8 A summary of JAK family members and their association with cytokine receptor families. Gamma chain cytokine, interferon, gp130 associated and growth factor receptor cytokine families have been targeted therapeutically in RA. IL-12Rβ1 cytokines are associated with diseases such as psoriasis, inflammatory bowel disease and seronegative arthropathies. A specific JAK1 inhibitor should target gamma chain cytokines, interferon and gp130 associated cytokines and make an attractive target. JAK2 inhibition would spare gamma chain cytokines but target growth factor receptors and may have unwanted side effects such as anaemia due to inhibition of erythropoietin signaling. JAK3 inhibition will only target gamma chain cytokines and may be too specific for control of inflammatory arthritis. A specific TYK2 inhibitor would spare both gamma chain cytokines and growth factor receptors and may also be therapeutically attractive.

Figure 1-8 outlines in general the JAK architecture and also shows which cytokine families a specific JAK inhibitor would hit. JAK1 sits downstream of interferon receptors although it is also implicated in IL-2 and IL-6 signaling. JAK2 was originally an unattractive target because it was involved with growth factor transduction including Epo but is also downstream of other cytokine receptors including type I and type II interferon, IL12/23 and IL-6(67). Therefore this makes JAK2 an attractive target if a small molecule can be developed with properties that allow inhibition as a heterodimer but relatively sparing inhibition when it is deployed as a homodimer downstream of growth factor receptors. Also the interferon and IL-6 pathways are important in RA but IL-12 and specifically IL-23 have been implicated in psoriasis, spondyloarthropathies and inflammatory bowel disease(161,162).

TYK2 is involved with type I interferon signaling and also downstream of the IL-12 and IL-23 axis and therefore may be a target in spondyloarthropathies or psoriasis. Theoretically, both TYK2 and JAK2 inhibition could lead to a powerful effect on diseases that are driven by Th_{17} cells and abnormal IL-23 signalling although monoclonal antibodies targeting this pathway have already been approved for conditions such as psoriasis and psoriatic arthritis(163).

JAK3 is an attractive target because it sits downstream of the gamma chain cytokines such as IL-2, IL-4, IL-7, IL-9 and IL-15 therefore it would have an effect on both T-cell development as well as maturation of B-cells and therefore would potentiate two important aspects of pathogenesis in RA. Furthermore, JAK3 is limited to lymphoid cells and therefore off target effects would be limited by targeting this kinase alone. Human mutations of JAK3 result in severe combined immunodeficiency and these patients have poor lymphoid development(164).

JAK inhibitors or jakinibs tend to target multiple parts of the kinome with different levels of specificity. Therefore a JAK3 inhibitor alone may be an attractive idea because of target limitation to lymphoid cells but from pathogenesis and current treatments such as relative ineffectiveness of calcineurin inhibitors in RA, we know that other cell types are also crucial to disease pathogenesis. Therefore inhibitors with a relative inhibition of one or two JAKs over others at a particular concentration are the rule rather than the exception.

Although tofacitinib was thought to be a JAK3 specific inhibitor, there is evidence that it inhibits JAK3 and JAK1(73). Therefore this drug targets both gamma chain signaling

cytokines as well as those that signal through gp130 such as IL-6. This combination would appear to be the perfect storm for RA because these cytokines are involved in pathogenesis and there are approved and effective treatments for them in the form of IL-6R blockade(53).

However, although JAK2 inhibition was discounted because of embryonic lethality in murine models and due to it's role in growth factor and colony stimulating factor signaling, relative inhibition of this JAK may lead to an even greater therapeutic effect. JAK2 mutations have been implicated in myelofibrosis (165,166)and ruxolitinib is approved for this indication. The novel JAK1/2 inhibitor baricitinib is showing promise in clinical trials and therefore I believe that JAK2 inhibition is crucial in treating RA given it's widespread use downstream of cytokines which are pathogenic in RA(78).

A clinical trial of tofacitinib versus methotrexate in patients with RA showed good efficacy and tolerability of the drug in patients with relatively short duration of disease(69). Furthermore it is the first drug to demonstrate efficacy beyond methotrexate with almost 40% of patients achieving an ACR70 response by 6 weeks at the 10mg dose compared with 10% achieving the same outcome on methotrexate.

However, jakinibs are associated with an infection signal for both viruses such as herpes zoster and also tuberculosis and this is likely to be drug related in a certain proportion of patients because of interferon disruption(167,168). However the tuberculosis cases tended to occur in those from TB endemic countries and these patients were initially screened as negative.

Furthermore in the case of tofacitinib and baricitinib, abnormalities in serum creatinine and lipids have been demonstrated. The abnormalities in lipids, with a predisposition for a worsening in cardiovascular risk profile, also occurs with IL-6R blockade using tocilizumab and therefore the lipid effects may be due to blockade of IL-6 systemically(169). In addition, lipid abnormalities were improved with concomitant use of medications such as statins although quantifying the long-term effect of jakinibs on cardiovascular outcomes would take Phase IV observational studies.

Although these drugs have been approved and are being prescribed to patients, the exact mechanism of action and mechanism of side effect is theorized as opposed to known. Therefore by using jakinibs you can investigate the effect of relatively inhibiting different

combinations of JAKs and also exploring their effect on primary cells from patients as well as in assay systems such as in macrophage and Tck co-culture.

However, the current model of drug discovery is broken with failure more costly than ever before. Stratified medicine promises to get the right drug to the right patient at the right time but to do this we need to interrogate disease pathology using high-throughput measures. Therefore the use of 'omic technologies and crucially the integration of this data should lead to advances in drug discovery(170,171).

1.9 The role of omic technologies in the investigation, pathogenesis and treatment of chronic diseases

'Omics is a word which is heard often in the current scientific and medical literatures and is meant to herald a new age of stratified medicine and treatment where the right condition is diagnosed so that the right drug can be given to the right patient at the right time. –Ome in this context means all parts considered in their entirety as opposed to alone. There is an assumption that by using a particular 'omic technology that as a researcher we will understand all of the components of that system and furthermore the field of systems biology exists because each particular –ome cannot exist by itself.

However each particular 'omic technique will only allow measurement of the constituent parts and depending on which technology is used this may not comprise of all parts of that particular system. In addition, with increased measurement comes larger amounts of data and therefore conventional methods of visualisation, carrying out statistical tests and interpreting the results become cumbersome and infomatic techniques become a necessity.

The triumvirate of the clinician, the scientist and the bioinformatician is incredibly powerful with each able to offer further insight into a problem which would not have been possible without the other two. Our job as clinical academics is to push the frontiers in the diagnosis and treatment of conditions for our patients. The advances in molecular biology and computing technology have allowed this to happen but the questions, hypothesis generation and interpretation still need to be carried out in the same method as in previous decades. Further with higher fidelity clinical information comes the real ability to integrate clinical information with the results of high throughput techniques and then

interrogate the system based on the results using advanced molecular biological techniques.

1.9.1 Genomics

Watson and Crick published the structure of DNA in 1953 with Rosalind Franklin confirming this with X-ray crystallography(172-175). Molecular biologists became interested in determining the sequence of DNA because Nirenberg and Leder (176)determined that the genetic code was laid out in codons or triplets which coded for different amino acids, the building blocks of proteins.

Early sequencing was called the "plus and minus" method where primers were extended using radio-labelled nucleotides and mixtures of these nucleotides were used alongwith gel electrophoresis to determine the sequence of DNA(177). This method was useful for short lengths of DNA of around 80bp but was incredibly time consuming and also required the use of a radio-label. Following this, Sanger (or shotgun) sequencing was born where read length had increased from 100 to 1000bp and employed a method of early chain termination with fluoro-labeled dideoxynucleotides which are unable to form a phsophodiester bond(178).

This however still meant that longer pieces of DNA could not be sequenced in one step and therefore the DNA had to be sheared into random segments and therefore these were sequenced to give a "read". The reads were then assembled using computational algorithms into the whole DNA sequence based on the overlapping regions and hence the concept of sequencing coverage was also born where a particular base can be "read" multiple times in overlapping "reads" to give confidence that the sequence is correct or that a particular base is different as may be the case if a cancer genome is sequenced.

The first human genome was sequenced in this fashion and Sanger sequencing is still useful for short continuous read lengths such as to confirm the sequence of a PCR product. However this method is not high throughput and therefore methods were developed by companies such as Solexa and Illumina which allowed the DNA shears to be attached to a slide, extended *reversibly* in parallel and at each DNA extension step, an imaging camera captured which fluorescent base had been added. Another sequencing technology called pyrosequencing relied on the measurement of the free hydrogen ion which is released

when DNA extension occurs. In both cases, a collection of reads are built up which would then need to be assembled into a full sequence.

This technology was reliant on appropriate molecular biology enzymes DNA polymerase and indeed this enzyme or variants of it are critical for both genomic and transcriptomic applications from qPCR to microarray and RNA-sequencing. Without the discovery of thermostable versions of the DNA polymerase enzyme and also the reverse transcription enzyme for transcriptomics, high throughput methods would not have been possible.

Early high-throughput sequencers delivered short read lengths which posed a bigger problem when it came to assembly because of the computational challenges. Firstly it is harder to do a puzzle with 100 pieces as opposed to 10 and this is the problem when you have a shorter read. Furthermore the process of producing a read is error prone and therefore assembling reads with multiple errors is extremely challenging.

The process of assembly and genome finishing was made possible by advances in computer processor speed and also reductions in the cost of both RAM and long term storage. Therefore bioinformatics teams were critical to sequencing in order to deal with the problem of having to join up the pieces of the puzzle. Furthermore, because of the earlier sequencing efforts, we now have reference or scaffold genomes for many organisms and therefore we are able to assemble reads against a known sequence akin to making a jigsaw puzzle using the box as a reference guide.

However the early shotgun sequencing approaches required *de novo* assembly where the reads had to be built up without a reference. This is more difficult because it is computationally intensive as each read has to be matched against all others but advances in mathematical algorithms to enable text matching enabled considerable time savings to be made as the Human Genome project continued. Therefore in most cases, reference genome based assembly is used in current genomic or transcriptomic approaches although *de novo* assembly is needed if large mutations are present which result in deletions as opposed to substitutions and also for genomes without a reference.

Following sequencing, the genome itself required annotation for areas which may code for proteins, areas which do not code but may have other functions such as microRNA and long non coding RNA and also the function of these particular areas. These efforts have

been made possible by curated databases such as Ensembl which show regions of the genome in incredible detail and on which multiple sources of information can be overlaid.

In addition to whole genome sequencing which can be resource and computationally intensive, exomes or the exon containing portions of the genome can be sequenced to give a more targeted approach. Furthermore the annotation of the genome and in particular analysis of mutations in patients with diseases allowed companies such as Illumina to make slide or "chip" based techniques which would allow the analysis of genomic or transcriptomic information in parallel. This type of analysis can be used for GWAS studies if known single nucleotide polymorphisms are incorporated onto the chip and this technology will be discussed further in the transcriptomics section.

With the advent of newer sequencers and also advanced sequencing chemistry has come the promise of the £1000 human genome and certainly with the newer Illumina HiSeq X Ten systems this is a possibility. However we now live in an age where patients and the general public wish to have their sequences known even if it is unclear whether the information generated is correct or useful. Companies such as 23andme offer DNA service, which results in a report giving a health overview, a DNA relative function and also a comment on your ancestry. This information is provided not for diagnostic purposes and only for information but it is difficult to see how someone would not be concerned if they were at a high risk of Alzheimers disease. In conclusion, patients and the general public will soon be coming to their doctors and other health care professionals armed with sequencing and other data and therefore a knowledge of how this is generated is critical.

Finally, the epigenome(179) is the collection of modifications which can occur to the genome itself to switch on or off genes or regions of DNA by changing the accessibility to polymerizing enzymes. DNA is normally held in a tight configuration in histones and in this state is not able to be transcribed. Histones have to open to allow portions of DNA to be transcribed and this can be controlled by histone acetylation. Furthermore, methylation is a process where the DNA is altered and normally suppresses gene activity and alterations to the genome sequencing methods have allowed the methylation status of DNA to be determined at a gene and also genomic level.

1.9.2 Transcriptomics

Although the genome for a particular organism is fixed and is stored in the nucleus in each of it's cells, the transcriptome varies. It is the transcriptome which varies between a skin dermal fibroblast and a synovial capillary endothelium. Although the transcriptome itself need to be translated into protein, sequencing of the human genome and also high throughput sequencing technology have allowed researchers to interrogate the RNA species present down to a single cell.

As well as messenger RNA, which is translated into protein, transfer RNA, ribosomal RNA, small RNAs such as microRNA and also long non-coding RNA each have their role in the transcriptome. Furthermore each of these particular species requires modifications to experimental protocols to ensure that enrichment of that particular part of the transcriptome occurs prior to high throughput method deployment.

Prior to discussing both microarray technology and also RNA sequencing, the molecular biology discovery of reverse transcription and also thermostable DNA polymerases made these techniques. Total RNA isolation is possible from many different tissue sources including solid tissue, cultured cells and primary blood cells. This total RNA is then reverse transcribed to give a more stable complimentary DNA or cDNA strand. In the case of a particular cell such as a CD4 T cell from a patient with RA, thousands of cDNA strands will be created.

In the case of RT-PCR and qRT-PCR, specific primers are employed to amplify a particular gene of interest based on a sequence that crosses an exon-exon junction thereby allowing amplification of the mRNA but not genomic DNA. Therefore this process allows quantification of a particular gene under different conditions when compared to a housekeeping gene i.e. a gene that is expressed at a similar level under all experimental conditions. A similar process is carried out in the case of both preparation for DNA microarrays or RNA-Sequencing. As well as the PCR stage, with higher throughput methods, the cDNA is often labeled or barcoded depending of whether microarray or sequencing is being employed.

1.9.3 DNA microarrays

DNA microarray technology was made possible by the annotation of the human genome for both coding and non-coding elements. Companies such as Affymetrix then constructed DNA "chips" where known sequences or "probes" were seeded in a pattern onto a glass

slide. The probes, in the case of Affymetrix, were made of 25 base long lengths of DNA that were specific and complementary to sequences of known genes. In addition mismatched probes were also included as well as probes for non-coding parts of the genome and quality control probes from species other than the one for which the array was intended for.

DNA microarray technology relies on the RNA transcriptome being labeled fluorescently and then hybridised to the array using complementary base pairing. The higher the amount of labeled cDNA hybrisiding to a particular spot, the brighter the fluorescent signal coming from that spot when the array was scanned. Therefore DNA microarrays allowed researchers to relatively quantify the transcriptome between experimental conditions on a grand scale. The Affymetrix U133 Plus 2.0 Human array contains 1.3 million individual oligonucleotide features with around 54,000 probe sets on the array itself. A probe set is typically made up of 20 individual probes for the same gene in the last 3' exon of the gene. Other arrays such as the exon array have probes present in all exons and even the 3' IVT arrays have genes duplicated on the array itself.

A critical factor in both microarray and RNA sequencing is the quality of the RNA because RNA is more susceptible to degradation by natural RNAases when compared to DNA. In particular, if RNA is degraded heavily, the 3' end RNA can be affected significantly and this can be assessed using informatics tools in the array experiment itself. However carrying out an array experiment on poor RNA will lead to poor results and therefore an analysis of the RNA prior to array or sequencing should be carried out. The RNA Integrity Number (RIN) is calculated from an algorithm which analyses the 18S and 28S fragments as well as the regions in between and gives a number for the quality of the RNA based on the peaks which should be present in good quality RNA and also the absence of degradation product bands. RNA with a high RIN (>7) is most useful for high throughput applications although lower RIN values may be used for qPCR experiments.

The employment of a fluorescent measurement to quantify the transcriptome means that quality control becomes more important as batch effects and also scanning artifacts can have a significant effect on an experimental result. The particular quality control measure employed in microarray and sequencing technology are discussed in those particular chapters but assessment of raw and normalised data is crucial prior to statistical tests and making conclusions. Furthermore a microarray or sequencing experiment needs to have its

findings validated by other methods such as qPCR or protein expression of transcriptional changes.

Both array based and sequencing based techniques require bioinformatics input to assist the biologist in processing the data and with initial data analysis. With the use of high throughput techniques comes the problem of multiple testing and the real possibility of false positive results. If we take a stringent p value of <0.01 as the level of confidence we wish to apply to a particular statistical test of a gene between two conditions, we would expect approximately 540 probe sets to be changed by chance alone. This is based on the assumption that on an Affymetrix Human U133 Plus 2.0 array, approximately 54000 genes are represented and with a p value of 0.01, 540 genes may be falsely determined as differentially expressed. False discovery correction(180) is employed to assist researchers in accounting for this but good experimental design in the form of adequate biological replicates and randomisation to prevent batch effects is critical. This problem and others once identified can be taken account for or analysed within bioinformatics packages.

As well as microarrays for determining the transcriptome, the annotation of the human genome allowed manufacturers to produce array based methods to carry out GWAS studies. Basing their chip design on earlier sequencing studies, microarrays such as the Illumina HapMap allow researchers to determine SNPs in common conditions and process the data using common pipelines. Further methylation status can also be determined using an array-based method although sequencing is often used to confirm initial findings.

1.9.4 RNA sequencing

The process of RNA sequencing (181)is similar to DNA sequencing albeit that the cDNA is sequenced as opposed to genomic DNA. The critical step in RNA sequencing is determining which RNA to sequence. Ribosomal RNA makes up over 90% of the Total RNA isolated from a cell. Ribosomal RNA does not pose a problem in microarray based technology because the specific nature of the probes on either the chip or bead, in the case of Affymetrix of Illumina respectively, means that hybridisation to reverse transcribed ribosomal RNA should not occur.

To take account of this, the researcher has to make a decision as to whether they wish to sequence coding RNA ie that which has a poly adenylation signal, coding and non coding RNA which includes long non coding species or microRNA. These decisions occasionally need to be made prior to RNA extraction as small RNA species can be lost and therefore

the input of a bioinformatician versed in sequencing issues is critical at the experimental design stage.

Three techniques exist in general with regards to preparing a sample for sequencing:

- Poly A selection allows RNA which has a poly adenylated signal to be pulled down and sequencing. This signal exists at the 3' end of the RNA and therefore relies on high quality RNA
- Ribosomal reduction is used if long non coding RNA is required and aims to remove the ribosomal RNA. Other species of RNA such as mitochondrial RNA will still be present and these can be removed computationally at the data analysis stage.
- MicroRNAs need to be selected specifically and if a researcher wished to explore this avenue then a separate sequencing experiment is required.

The RNA is then processed in that it is prepared into a library by shearing, reverse transcription and barcoded with specific primers. The process of barcoding enables multiple RNA samples to be loaded onto a particular sequencing chip and therefore allows multiplexing.

Following sequencing, the quality of the data is assessed and then differential gene expression can be carried out using cloud based services, command line tools or graphical interfaces such as CLC genomics workbench or the Galaxy project(182). Each of these techniques has its advantages and disadvantages but the creation of the Illumina Basespace means that biologists do not require an in depth knowledge of the process of sequencing alignment and differential expression in order to get meaningful data in the form of a list of differentially expressed genes.

The Basespace project uses scalable Amazon Web Machines to deploy more computing power when it is required such as during the process of alignment and therefore allows a typical sequencing project to be processed in around 24 hours for a project involving 8-16 samples. Furthermore Illumina sequences upload raw data directly to Basespace and therefore the data itself is protected and stored with each sequencing file often in the region of 4GB. A typical workflow for an RNA sequencing project is shown in figure 1-9 and

although newer and optimised methods for alignment may be developed, the general methodology remains true.

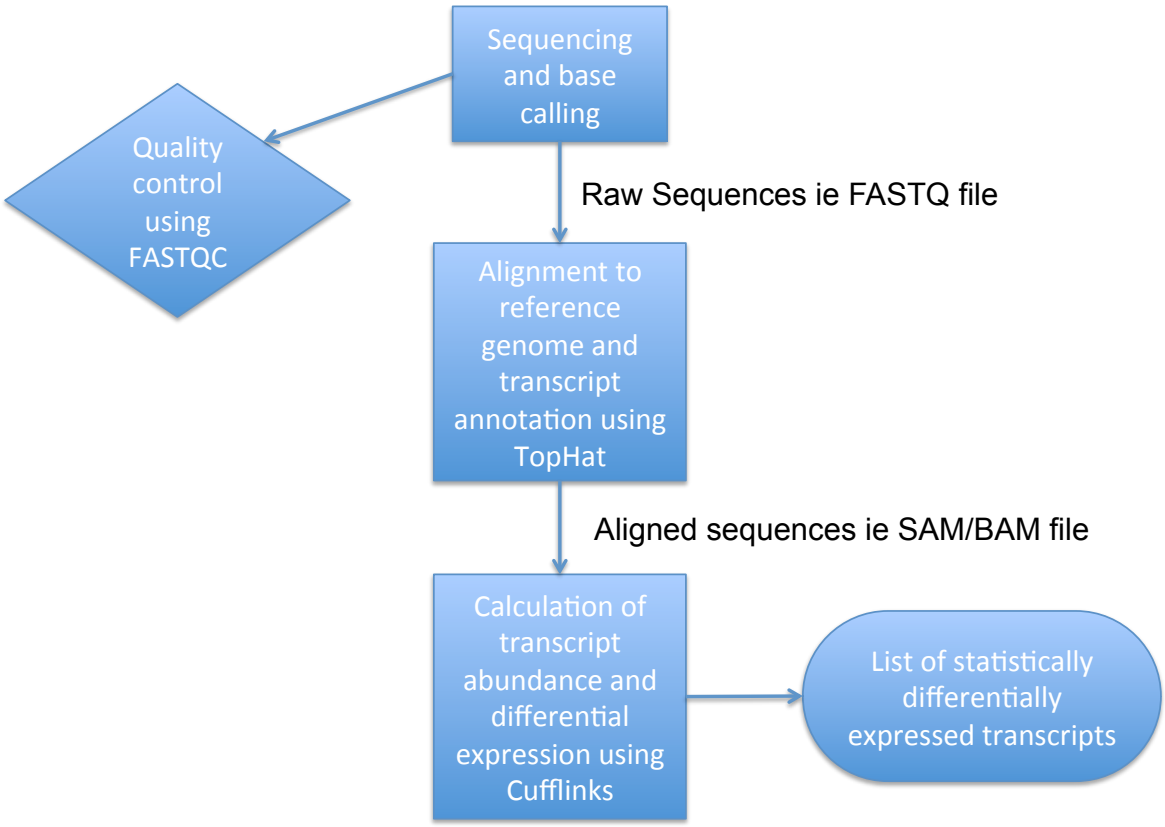


Figure 1-9 An overview of an RNA sequencing project processed in Illumina BaseSpace. All sequencing projects result in base calling with an associated quality score. These sequences are then assessed for their overall quality and high quality sequences are aligned to a reference genome if one is available. Thereafter, differential gene expression is carried out and a list of differentially expressed genes generated.

1.9.5 Bioinformatic analysis of array and high throughput data

Early experiments involving DNA and RNA sequencing involved using visual methods such as gel electrophoresis or blotting. As PCR technology evolved, fluorescent markers made semi quantitation a possibility and this could again be visualised using standard curves and graphing.

However visualising over a million probes led to problems in making the data accessible and meaningful and therefore bioinformatics solutions were required. These came in the form of custom solutions in the first instance and once refined these methods were incorporated into software packages. Common software packages to process and analyse microarray and sequencing data include Genespring and Partek Genomics suite. These tools allow a researcher to take the proprietary data files and derive a differentially expressed gene list in relatively few steps. They however tend to be black box solutions and do not allow much customisation or even an understanding of the processes behind data processing.

The R statistical project grew out of a need to have a simple but powerful command line interface with which to carry out statistical calculations and also data visualisation(183). The R project is open source with a new release every six months and is heavily supported by an online community. Furthermore, when researchers started using R for bioinformatics work, the Bioconductor(184) project was born to enable software developers to create, test and manage packages which would be useful to all bioinformaticians by providing a curated repository with vignettes for each package(184). Bioconductor requires packages to be tested on multiple system environments and therefore when software analysis is done using particular software versions one can be confident that the same result will be found by another researcher using the same data as long as they carry out the same commands as the original researcher. Further the process of scripting, even at the most basic level, means that reproducibility can be ensured and also critique can be applied.

In addition, novel statistical techniques are often implemented in R and similar languages and therefore when new methods are announced, using them in a current structure is often easier than if one was starting from scratch. Also, the power of a command line interface is critical when analyzing large data sets but there is also the reality that biologists may

urinary work better with graphical interfaces and therefore using tools such as Tcl and Shiny(185) have allowed this powerful tool to become more accessible.

Microarray analysis in R is a challenging prospect for the researcher who does not have a computing or mathematics background but with time the same analysis techniques can be applied to microarray, sequencing and qPCR data. Throughout my thesis I will be referring to R and Bioconductor packages that have been used to process and analyse data and also including my scripts in the appendix.

1.9.6 Proteomics

In keeping with the other 'omic fields, the aim of proteomics is to characterise all of the proteins in a particular biofluid or tissue.

The field of proteomics is particularly complex because of the post-translational modifications that can occur to proteins and therefore the alteration of the protein from a simple amino acid sequence to a structured protein. Methods of protein detection are common and the enzyme linked immunosorbant assay, bead multiplex and Western Blotting are methods of measuring proteins that rely on the ability of antibodies to recognise specific epitopes on proteins.

These systems work well if a particular protein or group of proteins is selected a priori such as TNFα or cytokines and growth factors. Discovering novel proteins requires the use of high throughput technology such as mass spectrometry and also curated protein databases again showing the important interface between technology and bioinformatics(186).

Briefly, the proteins need to be separated by mass and charge by gel electrophoresis, and then particular proteins that are changed between experimental conditions are analysed using mass spectrometry. The mass spectrometer takes a sample and then ionises it using electrons or protons. This results in the disruption of chemical bonds between amino acids or in the case of metabolites, within chemicals to produce ionised fragments. These fragments are then accelerated along a voltage gradient until they reach a detector and a highly accurate mass and charge are read from the sample.

Mass spectrometry is often coupled to a method of electrophoresis to separate out complex mixtures and avoid overloading the mass spectrometer resulting in ion suppression and the inability to detect ions at the same mass. Gas chromatography, liquid chromatography or capillary electrophoresis can be used and the method of electrophoresis is dependent on the size, charge of the analyte of interest as well as other factors.

Based on the mass and charge, a predicted protein sequence can be inferred and this is then compared to curated databases of proteins such as UniProt. Furthermore, computational algorithms can now predict protein structure based on amino acid sequences to determine whether particular proteins might have signaling, transmembrane or enzyme domains. X-Ray crystallography and nuclear magnetic resonance (NMR) are then used to confirm structural identities.

However the use of proteomics on a global scale in human biomarker studies has been disappointing. In particular, most proteins are rare and therefore are difficult to detect and furthermore the blood plasma, which is an accessible bio fluid, has high levels of albumin and antibodies which tends to make the detection of other proteins difficult. A solution to this has been the use of urinary proteomics that employs the kidney as a protein sieve that keeps albumin and antibodies in the blood compartment.

Urinary proteomics has been used in small studies to show the difference in the urinary proteome between patients with early RA and healthy controls(187). Further studies are ongoing to determine whether the changes seen are due to joint breakdown products, generalised inflammation or are specific to different inflammatory conditions.

1.9.7 Metabolomics

The metabolome like the transcriptome and proteome is the collection of all small molecules within a biological sample or organism. The study of the metabolome is used widely in clinical medicine on a targeted basis when we measure glucose levels, vitamins and blood creatinine. Furthermore there are some niche applications such as in the diagnosis of inborn errors of metabolism that use mass spectrometry based approaches as opposed to clinical biochemistry or enzymatic methods.

The metabolome of human serum and urine has been studied for decades using two main techniques, mass spectrometry coupled to a separation method and nuclear magnetic

resonance or NMR. The target molecule tends to have a mass less than 1kDa that include amino acids, sugars hormones and lipids. Profiling of bio fluids from patients has involved metabolomics of serum, urine and also synovial fluid (188,189) (190-194).

However as well as fluids, intracellular metabolomics has led to an integration of the metabolomics and immunology. Variation in the concentration of salt and lactate leads to alteration in the activation of immune cells and therefore serum and extracellular concentration of ions and metabolites have effects on the immune system. Furthermore, by measuring the intracellular metabolome we have been able to show that in murine macrophages, LPS activation of TLR4 leads to alterations in succinate and other members of the tricarboxylic acid cycle(195).

This integration of the metabolome with immune cell function gives credence to the idea that the metabolome is the result of or the outcome of the other 'omic levels such as the transcriptome or proteome and therefore should take account of post-transcriptional and translational modification.

However there are significant challenges with mass spectrometry based approaches and in particular when using an untargeted methodology, identification of metabolites can be difficult(196-198). Each metabolite has a particular mass and charge and these are then compared to a database of known metabolites and each potential metabolite then needs to be matched with a standard of that metabolite. In addition, the processing of the sample prior to mass spectrometry needs to be tailored based on expected experimental outcome.

NMR is described as being more quantitative than LC/MS but the number of metabolites which can be detected is less than for LC/MS. Clinically this can be useful where the high throughput nature and non-destructive nature of NMR mean that NMR based tests can be deployed using patient samples(199,200). Furthermore, there are methods that also assess certain metabolite families such as lipids and lipoproteins in a quantitative fashion.

1.9.8 Bringing it all together: The Science of Data

The "Big Data" revolution is truly underway and this is also the case in both medicine and life sciences although the high stakes yields of the financial services or banking industries is not often immediately apparent. Traditional experiments rely on a hypothesis to be tested using a specific experimental setup but the paradigm shift in data science is that the data itself will reveal patterns itself(171,201-204).

On the surface this may seem similar to a hypothesis free approach which was common in the early days of genomics and transcriptomics but in good data science experiments, there is a pre-defined analysis plan with primary data analysis plans being akin to the primary endpoint in a randomized control trial. Thereafter secondary analysis plans can be considered but these need to be pre-defined for each data science experiment.

Big Data has four main properties known as the four Vs:

- Volume
- Variety
- Velocity
- Veracity

The volume of data is measured by the physical storage that it occupies but also can be the number of observations that are held in total. Therefore there is a difference between a trawl of the known World Wide Web with the total number of web pages at a given time with a large number of observations made on a smaller number of individuals such as immunophenotyping combined with RNA sequencing in a RA disease cohort. Many of the statistics which are used to analyse Big Data make the assumption that there are both a large number of observations and a large number of individuals from whom the observations have been made but this is often not the case in life sciences unless we are looking at epidemiological resources or processes such as journeys through Emergency Departments.

Variety in the case of life sciences and medicine could be anything from clinical data, laboratory results to variant calls from genome sequencing. Information from wearable devices and social media also adds to this. Companies such as Hewlett Packard have developed intelligent systems, which can analyse data from hand written records and even video, and therefore allow historical records to be used in Big Data experiments. Imaging such as XR radiographs or MRI studies are often components of clinical trials and cohort

studies and the complete information from these is often not assessed and therefore this is also an area of unmet need.

The Velocity refers to the dynamic nature of method in that more is being collected as each second passes. In systems where analysis of real time data is critical, particular data architectures are more important so that they can deal with the fact that initial data set would have changed by the end of an analysis.

Finally Veracity refers to the uncertainty within the data and refers back to the quote from Donald Rumsfeld where we have known knowns, known unknowns and unknown unknowns. The quality of data going into an experiment is crucial and therefore meta data is extremely precious as any data about the data allows us to make an assessment of quality.

In addition to these properties of Big Data, two other factors are crucial: firstly, data visualisation is important to allow rapid assessment and interpretation of large data sets and secondly machine learning methods have allowed the computers to assess Big Data sets and determine patterns which are not apparent to humans(205).

Finally there are companies who combine both aspects of visualisation and machine learning in the form of topographical data analysis (TDA) that looks at the "shape" of the data(206). Companies such as AYASDI are now helping researchers unlock information within their own data. TDA has been employed in the assessment of Breast Cancer data where a novel subtype of patients who were oestrogen receptor (ER) positive but also had high levels of c-MYB and low levels of immune markers(207). This group was associated with a high survival and low possibility of metastases. Crucially, this group could not be found using traditional clustering methods and therefore TDA offered unique insight.

Recently, publicly available data from brain injury and spinal cord injury experiments were reanalaysed using AYASDI's platform and found that peri-operative hypertension predicted recovery better than any drug(208). The analysis of data that was shelved for decades has now allowed the group to generate new hypotheses that can be tested.

In conclusion, data science should sit at the centre of polyomic approaches to the interrogation of diseases although this is challenging as part of a PhD. However by

employing software packages that are commonly used by bioinformaticians, the data is in a format that is best suited to be further analysed using advanced techniques.

1.10 Psoriatic Arthritis

Although my thesis is focused on the pathogenesis of RA, I will use some microarrays that have been prepared from synovial CD14+ cells from patients with psoriatic arthritis (PsA). Therefore I will provide a brief overview of PsA including some differences with regards to pathogenesis compared to RA.

1.10.1 Epidemiology

PsA has a prevalence of 1% but this may be higher as the disease is further sub classified (209). It affects males and females in equal proportions and occasionally can predate the onset of psoriasis. It is classified with seronegative spondyloarthropathies and spinal disease is linked to the presence of the HLA-B27 haplotype. It occurs in Caucasians more than Asians or Africans.

1.10.2 Classification

PsA can be subdivided further which makes it different to RA and the different types are outlined below:

- Asymmetric oligoarthritis often affecting large joints
- Symmetric polyarthropathy in a distribution similar to RA
- Axial spondyloarthropathy with stiffness and inflammation of the spine and sacroiliac joints
- Distal interphalangeal joint disease in a distribution similar to osteoarthritis but with a younger onset, inflammatory changes and often associated with psoriatic nail pitting
- Arthritis mutilans which is a deforming arthropathy that results in resorption of the terminal ends of the phalanges
- Entheseal disease where tendinopathy is the predominant feature but may also be associated with plantar fasciitis

1.10.3 Diagnosis

A rheumatologist based on a suggestive history, inflammatory symptoms and appropriate investigations makes a diagnosis of PsA. There are different classification criteria and also

CASPAR Criteria – In the context of inflammatory articular disease and ≥ 3 points from following categories(210)		
Category	Description	Points
Current psoriasis or personal or family history of psoriasis	Current psoriasis confirmed by a rheumatologist or dermatologist Personal history of psoriasis from patient or physician Family history of psoriasis from patient in a first or second degree relative	2 – current psoriasis 1 – personal or family history (only one category counted so maximum of 2 points)
Psoriatic nail dystrophy	Onycholysis, pitting or hyperkeratosis	1
Negative Rheumatoid Factor	Preferably ELISA	1
Dactylitis	Current or history from a rheumatologist	1
Radiographic evidence of juxta-articular new bone formation	Ossification near join margins but excluding osteophytes on hand or foot XR	1

Table 1-10 CASPAR criteria for the classification of psoriatic arthritis

disease assessment tools two of which are outlined below. The Classification criteria for Psoriatic Arthritis (CASPAR) (table 1-10) criteria are one of the most commonly used and have a sensitivity of 91.4% and specificity of 98.7%. Disease activity and outcomes are again measured by a variety of tools but require the assessment of 68 joints because of the distribution of disease compared to RA. The Psoriatic Arthritis Response Criteria (PsARC)(146,211) comprises of 66 swollen and 68 joint tender assessments along with both physician and patient global assessment measure on a five point Likert scale. This is used in a similar way to the DAS-28 to measure disease activity and also treat to a target.

1.11 Treatment

The principles of treatment for PsA are similar to RA with involvement of the multidisciplinary team with patient engagement an essential part of management. In addition, dermatology colleagues with their own MDTs will be useful to help control skin disease.

DMARDs are also used in PsA although there is a relative paucity of evidence for medications such as methotrexate(212) and sulphasalazine(211). Although hydroxychloroquine is used, it has been associated with a flare of psoriasis and therefore is not used as commonly.

First line biologic up until recently was only anti-TNF therapy(146) but other treatments that are effective in RA such as tocilizumab, abatacept or rituximab have not been effective in trials of PsA. This may represent inefficacy of the agent but may also reflect that PsA is a heterogeneous disease and effectiveness in subtypes is difficult to assess in “catch-all” clinical trials.

1.11.1 IL12/23 blockade

Ustekinumab is a monoclonal antibody targeting the p40 subunit of IL-12 and IL-23. It has been shown to be effective in controlling plaque psoriasis and is effective in psoriatic arthritis(163).

1.11.2 IL-17 blockade

Secukinumab is a monoclonal antibody to IL17A, a pro-inflammatory cytokine that is implicated in the pathogenesis of PsA. Compared to placebo, secukinumab improved both

the signs and symptoms of disease with 54% of patients treated with secukinumab 300mg achieving ACR20 compared to 15% of those taking placebo(213).

1.11.3 Phosphodiesterase 4 (PDE4) inhibition

Apremilast is a PDE4 inhibitor that has been demonstrated to be efficacious in psoriasis and psoriatic arthritis(214). It breaks down cyclic adenosine monophosphate (cAMP) and this decreases the expression of TNFα, IL17 and IL23 while increasing IL-10 and this may be related to mechanism of action.

The differential epidemiology, clinical features and also the differential response to treatment shows that RA and PsA have a fundamental difference in pathogenesis. They do however have an overlap in that they are both inflammatory arthropathies and can affect similar joints and therefore makes a good comparator for RA.

1.12 Hypotheses and Aims

Many data suggest that the JAK-STAT pathways subserve the effector biology of a range of cytokines in inflammatory arthritis pathogenesis. Herein, I hypothesise that a JAK-STAT transcriptomic signature will be evident in myeloid lineage cells from patients with RA and that this will be distinct form that present in PsA, as a comparator arthropathy, arising because of the rather fundamental differences in the pathogenesis of each disease. On this basis, I predict in consequence that the JAK1/3 inhibitor, tofacitinib, will inhibit critical inflammatory pathways in myeloid lineage cells and, in particular, will inhibit inflammatory cytokine production from macrophages following Tck cell contact activation and in conditions created to mimic the RA synovial microenvironment. Finally using the novel technique of RNA sequencing I will test the hypothesis that novel effector pathways and biomarkers of response in monocytes following tofacitinib treatment will emerge.

To investigate these inter-related hypotheses, I set out with the following aims:

1. Using bioinformatic techniques, to determine whether there is a JAK-STAT signature evidence in a microarray dataset of patient derived monocytes and macrophages and if so is this unique to RA.
2. To establish whether JAK inhibition prevents the production of pro-inflammatory cytokines such as TNFα when macrophages are cell contact activated by cytokine activated T cells.
3. To further determine if JAK inhibition prevents RA patient derived monocytes from being stimulated by soluble factors in RA synovial fluid.
4. To thereafter seek new pathogenetic pathways and biomarkers of treatment using RNA sequencing of RA patient derived monocytes treated with RA synovial fluid and tofacitinib.

Chapter 2 Materials and Methods

2.1 Buffers and Reagents

2.1.1 Complete medium

RPMI – 1640 (Life Technologies), with 10% heat inactivated foetal bovine serum (Invitrogen), 100IU/ml penicillin, 100µg/ml streptomycin and 2mM L-Glutamine. (Gibco)

2.1.2 FACS/MACS wash buffer

Dulbecco's phosphate-buffered saline (DPBS) (Life Technologies) without calcium with 2% heat inactivated foetal bovine serum (Invitrogen)

2.1.3 FACS stain buffer

DPBS with 0.5% bovine specific albumin (Sigma)

2.1.4 PEA buffer

DPBS, 2mM ethylene diamine tetra-acetic acid (EDTA) and 0.5% human serum albumin.

2.1.5 Tofacitinib, Ruxolitinib and Tyrphostin AG-490

Small molecules were obtained from LC-Labs and dissolved in DMSO (Sigma). Serial dilutions were employed to ensure that DMSO concentration did not exceed 0.001% in cell culture. Single use aliquots were prepared and stored at -80C.

2.2 Cell Culture

All tissue culture work was carried out in laminar flow hoods and cells cultured at 37C in a 5% CO_2 atmosphere. Cells were cultured in complete medium.

2.3 Isolation and preparation of cells

2.3.1 Isolation of peripheral blood mononuclear cells (PBMC) from whole blood or buffy coats

Whole blood was obtained from healthy volunteers under University of Glasgow Ethics or from patients with Rheumatoid Arthritis under West of Scotland Research Ethics Committee 4 (14/WS/1035) NHS GG&C additional sample tissue resource to support I3I Research. Buffy coats were obtained from the Scottish National Blood Transfusion Service, Gartnavel Hospital, Glasgow.

Whole blood was diluted 1:2 with DPBS without calcium (Life Technologies) and buffy coats diluted 1:4. Blood was separated by density centrifugation by layering onto Histopaque -1077 (Sigma) and spinning according to manufacturers instructions. The interface layer of PBMCs was isolated and washed twice with wash buffer and once with complete medium. The cell suspension was filtered with a 30µm filter (Miltenyi) prior to downstream applications. Cell viability and number was assessed using Trypan Blue (Sigma) staining and a haemocytometer.

2.3.2 Isolation of mononuclear cells from synovial fluid

Synovial fluid was obtained from patients undergoing joint aspiration in the NHS Greater Glasgow and Clyde area under ethics approval 14/WS/1035. Synovial fluid was spun at 1200g to obtain a cell pellet. The fluid portion was processed into aliquots and frozen at -80C. The cell pellet was washed twice in wash buffer and once with complete medium. The cell suspension was filtered through a 30µm filter (Miltenyi) and then assessed for viability and number as described previously prior to downstream processing.

2.3.3 Positive selection of CD14+ monocytes

PBMCs were processed according to manufacturer instructions with CD14 microbeads, human (Miltenyi). Briefly, cells were suspended in 80µl of buffer per 10^7 cells. An AutoMACS pro was then used for automatic labeling of cells using CD14 human microbeads and the positive fraction collected. Labeled cells were assessed for viability and number as previously described. Purity of cells was assessed using CD14 + mouse anti-human PE conjugated antibody Clone: TÜK4 (Miltenyi).

2.3.4 Positive Selection of CD4+ T cells

PBMCs were processed according to manufacturers instructions using the CD4microbeads, human (Miltenyi). Briefly, cells were suspended in 80µl of buffer per 10^7 total cells. An AutoMACS pro was then used for automatic labeling of cells using CD4 human microbeads and the positive fraction collected. Labeled cells were assessed for viability and number as previously described. Purity of cells was assessed using CD4 + mouse anti-human PE conjugated antibody Clone: M-T466 (Miltenyi).

2.4 Microarray datasets and processing

2.4.1 Microarray Dataset 1

2.4.2 CD14+ peripheral blood monocytes and CD14+ synovial macrophages isolated from patients with inflammatory arthritis

Paired blood and synovial fluid samples were obtained from eight patients with RA and PsA, who met classification criteria, undergoing joint aspiration within the NHS Greater Glasgow and Clyde area. Samples were processed as per protocols for peripheral blood or synovial fluid as described previously. Cells were lysed using Trizol (Invitrogen) at a concentration of $2x10^6$ cells and stored at -80C till processing. An aliquot was reserved for purity staining by flow cytometry and cells were >95% pure.

RNA was isolated according to the protocol below and then processed according to manufacturers instructions for an Affymetrix Genechip U133 2.0 plus array. Briefly this process involved cDNA first strand and second strand synthesis. Purification, fragmentation and biotin labeling are carried out prior to fluorescent labeling and array hybridisation. Chips were scanned using an Affymetrix Genechip® Scanner 3000 and data acquired using Affymetrix GeneChip Command Console software. Each array was processed into a .CEL file.

2.4.3 CD14+ PBMC isolated from healthy volunteers, differentiated into MCSF macrophages and cell-contact activated

CD14+ monocytes were isolated by positive magnetic selection following density centrifugation of blood from healthy volunteers. Cells were assessed for viability and number as described previously. A proportion of monocytes were lysed immediately in Trizol (Invitrogen) with the remainder differentiated into macrophages and cell contact activated as previously described. An aliquot was reserved for purity staining by flow cytometry and cells were >95% pure.

To produce macrophages, CD14+ monocytes were stimulated with MCSF (50ng/ml) (Biosource) and cultured at 37C 6 days. In this case, to produce Tck, CD3+ positive magnetic selection was used and these cells were stimulated with IL-2 (25ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) for 6 days at 37C. These Tck were then washed with

PEA and fixed using DPBS and 2% paraformaldehyde on ice for two hours. Cells were then washed with PEA thoroughly, stained with CFSE and added to mature MCSF macrophages in 12 well plates at a ratio of 1 macrophage to 4 Tck for two hours. All cells were lifted using non enzymatic cell dissociation solution (Sigma) and then FACS sorted using a FACS Aria I with CFSE negative cells representing cell contact activated macrophages and CFSE positive cells representing fixed Tck. Both MCSF macrophages and cell contact activated MCSF macrophages were lysed in Trizol (Invitrogen) at a concentration of $2x10^6$ cells/ml. RNA was processed and arrays carried out as previously described.

2.4.4 Microarray Dataset 2

2.4.5 Wild type mice versus microRNA 155 knockout mice fed on a high fat diet

Briefly, Male C57 BL/6 wildtype (WT) and miR-155-/- (Jackson Laboratories) mice were fed a high fat diet (HFD) (0.15% cholesterol, 21% lard, Special Diet Services) for six weeks in house in a pathogen free facility. Total RNA was prepared from frozen livers by homogenisation and an Affymetrix GeneChip Mouse Gene 1.0ST array was performed. These arrays were uploaded to the ArrayExpress database under accession no. E-MEXP-3932.

2.4.6 Microarray Dataset 3

2.4.7 PBMC from healthy volunteers and patients with rheumatoid arthritis and systemic lupus erythematosus

This microarray dataset was downloaded from the public GEO data repository under accession number GSE 38351. This data set comprises of PBMC obtained from healthy volunteers and patient. A special study module medical student who I was supervising analysed this data set as part of his project but I used two of the figures that he produced to illustrate batch effect in Chapter 3. Cells were processed in a similar fashion as previously described and Affymetrix Human Genome U133A and U133 Plus 2.0 arrays were carried out.

2.4.8 Processing of microarray data

Microarray data was analysed using the R project for statistical computing and packages from the Bioconductor repository. Quality control of data was performed using the arrayQualityMetrics, affyQCreport and affyAnalysisQC script. These packages and scripts generate the various QC plots which have been used in Chapter 3.

Following QC, U133A and U133 Plus 2.0 arrays were processed using the affy package and the Mouse Gene 1.0ST array with the oligo package. Normalisation of array data was carried out using RMA, GC-RMA or fRMA as described in the text. Following normalisation, data was processed for differential gene expression using limma is a parametric test was required or RankProducts if a non-parametric test was employed. In the case of limma, genes with a false discovery rate (fdr) corrected p value of <0.05 were deemed to be differentially expressed unless otherwise indicated in the text. RankProducts generates a predictor of false positive (pfp) results that has inherent multiple testing correction and this was set at <0.05.

Scripts outlining example processes and also scripts used to generate Venn diagrams from Chapter 3 are included in the appendix.

2.4.9 Pathway Analysis

Differentially expressed genes were uploaded to Ingenuity Pathway Analysis (Qiagen) and analysed using Core Analysis or other options. For a core analysis, the significance of association between the differentially expressed genes and canonical pathways is calculated using a Fisher's exact test which is multiple testing corrected using a Benjamini-Hochberg correction.

2.5 RNA sequencing

2.5.1 Experiment outline

CD4+ T cells were obtained from blood of patients who had ACPA+ RA density centrifugation and positive magnetic selection using previously described protocols.

T cells were cultured in complete medium supplemented with 10% pooled synovial fluid. Synovial fluid was obtained from the Centre for Rheumatic Diseases biobank and the pool was made from two donors who were seropositive for antibodies against citrullinated peptides. Tofacitinib dissolved in dimethyl sulfoxide (DMSO) was added to cultures at a concentration of 1000nM. Control experiments contained the same concentration of DMSO as those containing tofacitinib.

1.25×10^6 cells were used in each condition and cells were cultured in 12 well plates for 24 hours. Cells were then harvested, washed and lysed using QIAzol (Qiagen). Total RNA was prepared as described below and samples were analysed by Nanodrop and Agilent Bioanalyzer for RNA quality. RNA Integrity Number (RIN) was greater than 7 for each sample put foreword for sequencing.

An aliquot of CD4+ cells was taken at baseline for evaluation of purity and viability with cells being >95% pure and >99% viable. An aliquot of PBMC was also cultured for 24 hours and then stimulated and stained as per intracellular FACS protocol.

2.5.2 Sequencing preparation and protocol

Glasgow Polyomics carried out all RNA sequencing. Sequencing libraries were prepared using TruSeq Stranded Total RNA with Ribo-Zero Human kit (Illumina) as per manufacturers instructions for control and synovial fluid samples. Sequencing libraries were prepared using TruSeq Stranded Total RNA with Ribo-Zero Gold kit (Illumina) as per manufacturers instructions for tofacitinib and synovial fluid and tofacitinib samples because of an issue with initial sequencing.

Samples were barcode labeled and run on a NextSeq 500 sequencer (Illumina) with 20-25M reads with 75bp paired end reads. Sample barcodes were demuxed on machine and data uploaded to BaseSpace for further processing and analysis.

2.5.3 Data processing using Illumina BaseSpace

Following sequencing, data in the form of FASTQ files were uploaded directly to Illumina BaseSpace and backed up on a local server. Files were assessed for quality using FASTQC and Phred scores analysed. All data had a Phred score >28 and tended to have a score of >30 for most bases. A Phred score of 30 corresponds to a probability that a base is called as incorrect of 1 in a 1000 giving 99.9% accuracy in the base call.

The data was aligned using TopHat2 and differential gene expression was performed using the Cufflinks pipeline to determine Fragments per kilobase of transcript per million mapped reads (FPKM) of each gene. This gave a list of genes that were differentially expressed between two conditions, as multi condition analysis is not possible in BaseSpace. The raw data from candidate differentially expressed genes was then visually inspected using the Broad's Integrative Genome Viewer (IGV).

2.5.4 Data visualisation using CummeRbund

Sequencing data tables were read into R and analysed using the CummeRbund package. Graphs of raw FPKM values were generated within this package to show replicates. The "findSimilar" function was used to discover genes that were differentially expressed in a particular pattern exhibited by an exemplar gene. This function uses the Jensen-Shannon distance to determine the similarity of the probability distribution of each gene across different conditions(215).

2.5.5 Pathway analysis

FPKM tables were uploaded to Ingenuity Pathway Analysis (Qiagen) and analysed using both Core and exploratory analyses. Genes that were differentially expressed (fdr corrected p value <0.05) and linked to the JAK/STAT pathway were validated using

qPCR. Genes that were highly differentially expressed (fdr corrected p value <0.01) between conditions consistently were also validated.

StringDB was also used to explore potential protein-protein interaction networks and clustering of genes.

2.5.6 qRT-PCR validation of RNA sequencing data

cDNA was prepared as described previously. Taqman Low Density Array plates (TLDA, Life Technologies) were designed based on differentially expressed genes. Probes spanned exon-exon junctions and were designed in conjunction with Life Technology technical support. TLDA cards were designed to run 96 genes in singlet with 4 samples per card.

0.7ng of input cDNA was used for each qPCR reaction in the 384 well TLDA plate made up as follows:

cDNA	75µl
Taqman Universal Master Mix II	115µl
dH_2O	30µl

100µl was loaded into each of two injection ports per sample. The cards were analysed using a 7500 Fast Real-Time PCR System (Applied Biosystems). PCR protocol: 95C incubation for 10 minutes followed by 40 cycles of 15 seconds at 95C followed by an incubation at 60C for one minute.

Data was analysed in the R using the HTqPCR package. Samples were normalised to 18S genes and presented using the Relative Quantification method.

2.6 In vitro assays

2.6.1 CD4+ T cell stimulation and differentiation into cytokine activated T cells (Tck)

CD4+ T cells were isolated from blood by density centrifugation and positive magnetic selection using the method described above. They were cultured at $1x10^6$ cells/ml with IL-2 (25ng/ml), IL-6 (100ng/ml) and TNFα (25 ng/ml) (Peprotech) for 6 days. In subsequent experiments, IL-2 was substituted with IL-15 (100ng/ml).

2.6.2 CD14+ monocyte stimulation and differentiation into macrophages

CD14+ monocytes were isolated from blood by density centrifugation and positive magnetic selection using the method described above. Monocytes were cultured in 24 or 96 well plates with MCSF (50ng/ml) (Peprotech) for 3 days. Cells were detached using non-enzymatic dissociation solution (Sigma), washed in DPBS and replated in 96 well plates at a density of $5x10^5$ cells/ml with MCSF (50ng/ml) for 3 further days and then used in subsequent experiments.

2.6.3 Macrophage cell contact activation by Tck

Tck were washed thoroughly with wash buffer and added to macrophages at a defined concentration. When JAK inhibitors were used in experiments, Tck were added to macrophages at a concentration of 4:1. Inhibitors or 0.001% DMSO were added to both Tck and macrophages 30 minutes prior to co-culture. A positive control of LPS (1ng/ml, Sigma) was used in each experiment. Co-culture occurred for 24 hours and supernatants were collected for ELISA or Luminex analysis. Transwell inserts (Corning) with pore size of 0.4μm were used to prevent cell contact between macrophages and Tck but allow soluble factors to permeate where appropriate.

2.7 Flow cytometric analysis of cells

2.7.1 PhosphoFACS stimulation and staining protocol for THP-1 cells

Cytokine stimulants were added to FACS tubes at concentrations described below:

Cytokine	Final concentration	Manufacturer
IFNγ	25ng/ml	R&D Systems
IL-6	20ng/ml	Peprotech
GM-CSF	200ng/ml	R&D Systems
IL-4	20ng/ml	R&D Systems

0.5 ml of THP-1 cells (ATCC) were added (0.5 x 10^6 cells per ml) and incubated in a water bath at 37C for 15 minutes. Cells were washed in ice cold DPBS and spun at 350g for 5min at 4C. Supernatants were removed after all wash steps. Cells were fixed using 500µl Cytofix (BD) and incubated for 15 minutes at room temperature and protected from light. Cells were washed twice with CellWash (BD) and then permeabilised using 1ml Perm III buffer (BD) for 30 minutes on ice and protected from light. Cells were washed twice with CellWash and once with FACS stain buffer.

Intracellular phosphoSTAT staining was carried out using the following antibodies:

Marker	Clone	Isotype	Source	Conc.	Conjugate
pSTAT1	4a	IgG_{2a}	BD	1:10	PerCP-Cy 5.5
pSTAT3	4-pSTAT3	IgG_{2a} κ	BD	1:20	AF488
pSTAT6	18/P-STAT6	IgG_{2a}	BD	1:10	PE

Cells were stained for 30 minutes on ice and protected from light. Cells were washed once in FACS stain buffer and resuspended in 200µl of FACS stain buffer and acquired using a BD LSR II. Experimental anaylsis was performed in FlowJo version v9.7.6 (TreeStar Software) or in Cytobank (www.cytobank.org).

2.7.2 PhosphoFACS stimulation and staining protocol for human PBMC using SNAP cocktail

Stimulants and surface stains were added to FACS tubes at concentrations described below:

Stimulant	Final concentration (in 200ul cell suspension)	Manufacturer
PMA	50nM	Sigma
Ionomycin	1μM	Sigma
Anti-CD3	1μg/ml	BD
Anti-CD28	1μg/ml	BD
IFNγ	25ng/ml	R&D Systems
IL-6	20ng/ml	Peprotech
GM-CSF	200ng/ml	R&D Systems
IL-4	20ng/ml	R&D Systems

Marker	Clone	Isotype	Source	Conc.	Conjugate
CD4	SK3	$IgG1_{\kappa}$	BD	1:40	PE-Cy7
CD5	UCHT2	$IgG1_{\kappa}$	BD	1:40	V450
CD8	SK1	$IgG1_{\kappa}$	BD	1:40	APCH7
CD14	MØP9	$IgG2b_{\kappa}$	BD	1:40	PE-CF594
CD19	HIB19	$IgG1_{\kappa}$	BD	1:40	AF700

200μl of PBMC ($1x10^6$ /ml) were added and incubated in a water bath at 37C for 13 minutes. Cells were washed in ice cold DPBS and spun at 350g for 5min at 4C. Supernatants were removed after all wash steps. Cells were fixed using 500μl Cytofix (BD) and incubated for 15 minutes at room temperature and protected from light. Cells were washed twice with CellWash (BD) and then permeabilised using 1ml Perm III buffer (BD) for 30 minutes on ice and protected from light. Cells were washed twice with CellWash and once with FACS stain buffer.

Intracellular phosphoSTAT staining was carried out using the following antibodies:

Marker	Clone	Isotype	Source	Conc.	Conjugate
pSTAT1	4a	IgG_{2a}	BD	1:10	PerCP-Cy 5.5
pSTAT3	4-pSTAT3	IgG_{2a} κ	BD	1:20	AF488
pSTAT6	18/P-STAT6	IgG_{2a}	BD	1:10	PE

Cells were stained for 30 minutes on ice and protected from light. Cells were washed once in FACS stain buffer and resuspended in 200µl of FACS stain buffer and acquired using a BD LSR II. Experimental anaylsis was performed in FlowJo version v9.7.6 (TreeStar Software) or in Cytobank (www.cytobank.org).

Cytokine stimulation alone was used in subsequent experiments where PMA, Ionomycin, anti-CD3 and anti-CD28 antibody were omitted from stimulation cocktail.

2.8 Enzyme linked immunosorbent assay (ELISA)

Supernatants from cell culture experiments were analysed with TNFα ELISA (Life Technologies) prior to Luminex analysis. Capture antibody was diluted in PBS according to the product data sheet and plated onto high capture 96 well plates (ThermoFisher) and incubated at 4C overnight. Plates were then washed with DPBS/0.05% Tween and blocked with blocking buffer (DPBS/0.5% BSA) to prevent non-specific binding for one hour at room temperature. Standards were reconstituted in RPMI + 5% FCS to represent the sample matrix with the top standard of 2000pg/ml and an eight point standard curve constructed by serial dilution. Cell supernatants were diluted 1:4 with complete medium and 100µl added in duplicate and incubated for 2 hours at room temperature with detection antibody. Plates were thoroughly washed with PBS/Tween and thereafter Streptavidin-HRP was reconstituted according to the product data sheet and was incubated for 30 minutes at room temperature. Plates were thoroughly washed with PBS/Tween and 100µl of TMB chromagen (ThermoFisher) was added and incubated at room temperature while protected from light until standard separation had occurred. 100µl of stop solution (ThermoFisher) was added to each well and the plate read within 30 minutes using a 450nm plate reader (Dynex Technologies).

2.9 Luminex

Following ELISA, a Human Cytokine 30-plex panel (Invitrogen) was carried out on supernatants and synovial fluid samples. Samples were thawed to room temperature and spun at 12000g for 1 minute to remove cellular debris. They were then diluted 1:2 with complete medium. Briefly, standards were prepared according to product data sheets. Assay wells were wet with 200µl of working wash solution and aspirated using a vacuum aspirator. Beads were then resuspended by vortexing and sonication and 12.5µl added to each well. Beads were washed with 200µl of wash solution and this was aspirated as before and washing repeated. 50µl of Incubation buffer was added to each well and 100µl of standard added or 50µl of sample and 50µl assay diluent to each well as appropriate. Plates were then wrapped in aluminium foil, protected from light and incubated at 4C on a rocker overnight. Liquid from wells was aspirated and wells washed twice as before. 100µl of 1x Biotinylated Detector Antibody was added to each well and the plate incubated for one hour at room temperature on a shaker. Liquid from wells was aspirated and wells washed twice as before. 100µl of 1x Streptavidin-RPE was added to each well and the

plate incubated for 30 minutes at room temperature on a shaker. Liquid from wells was aspirated and wells washed three times as before. 100µl working wash solution was added to each well and the plate shaken on an orbital shaker for 3 minutes. Plate was then read using a Luminex 100 instrument and concentration of samples extracted from standard curves.

2.10 RNA isolation

Cells were lysed in QIAzol reagent (Qiagen) as per manufacturer instructions and stored at -80C. RNA was extracted using the miRNeasy mini kit (Qiagen). Briefly, the sample and QIAzol was allowed to thaw and then left at room temperature for 5 minute. 140µl of choloroform (Sigma) was added to each sample, vigorously shaken for 15 seconds and then left at room temperature for 3 minutes. Samples were spun at 12000g for 15 minutes at 4C. The aqueous phase was transferred to a new tube and 525µl 100% ethanol (Sigma) added and thoroughly mixed. 700µl of sample was then transferred to a column in a collection tube and spun at >8000g for 15 seconds. This process was repeated until all the sample had been processed. 350µl of buffer RWT was added to each column and spun at >8000g for 15 seconds and the follow through discarded. A DNAse digest step was carried out using RNase-Free DNase set (Qiagen). 10µl of DNase and 70µl of buffer RDD was added directly to each column and incubated for 15 minutes at room temperature. 350µl of buffer RWT was added to the column and spun at >8000g for 15 seconds and the follow through discarded. 500µl of buffer RPE was added to each column and spun at >8000g for 15 seconds and the follow through discarded. 500µl of buffer RPE was added to each column and spun at >8000g for 2 minutes and the follow through discarded. The collection tube was replaced and the column spun at >8000g for 1 minutes and the collection tube replaced with a 1.5ml tube. 30µl of RNase-free water was added to each column and spun at >8000g for one minute. RNA was analysed using Nanodrop (ThermoFisher) and RNA concentration and 260/280 and 260/230 OD ratios recorded.

2.11 First strand cDNA synthesis

cDNA was synthesised from RNA using the AffinityScript Multiple Temperature cDNA synthesis kit (Agilent). Briefly, 100ng of RNA was added to a microcentrifuge reaction tube and 3µl of random primers added. RNase free water was used to take total reaction volume to 15.7µl. Reaction was initiated by incubating at 65C for 5 minutes and cooling

to 20C for 10 minutes. 2µl of 10x RT buffer, 0.8µl of dNTP mix, 0.5µl RNase block and 1µl Multitemp Reverse Transcriptase was added to each tube. The reaction was then cycled at 25C for 10 minutes, 55C for 60 minutes and 70C for 15 minutes. 80µl of RNase free water was added to take final concentration of mixture to 1ng/µl of cDNA.

2.12 Quantitative RT-PCR

Primer sequences are outlined below:

Gene	Primer Forward Sequence	Primer Reverse Sequence	Product Size
Bcl-6	CCCTATCCCTGTGAAATCTGTG	TCTCACAATGGTAAGGTTTCTCTC	100bp
Bcl2L1	GGCGGCTGGGATACTTT	TCATTTCCGACTGAAGAGTGAG	147bp
Myc	GACTCTGAGGAGGAACAAGAAG	CAGCAGAAGGTGATCCAGAC	103bp
Bcl2	GATAACGGAGGCTGGGATG	GAGACAGCCAGGAGAAATCAA	79bp
SOCS3	GGAGTTCCTGGACCAGTACG	TTCTTGTGCTTGTGCCATGT	116bp
18S	GGCCCTGTAATTGGAATGAGTC	CCAAGATCCAACTACGAGCTT	146bp

PCR reactions were carried out in a 96 well optical plate (Applied Biosystems). PCR thermocycling was carried out in a One-Step Plus Applied Biosystems PCR platform with settings Fast cycling settings:

50C for two minutes, 95C for two minutes and then 40 cycles of: 95C for 3 seconds and 60C for 30 seconds.

PCR reactions were as follows: 5µl SYBR Select, 1µl cDNA template (1ng), 0.5µl of forward primer (500nM), 0.5µl reverse primer (500nM), 3µl dH_20.

Data was analysed in R using the HTqPCR package.

2.13 Statistical Analysis

Statistical analyses were carried out in the R project for statistical computing or using Prism 6 for Mac OS X (Graphpad). The exact statistical test employed is described in the appropriate section.

Chapter 3 Developing a transcriptomic pipeline to enable analysis of a microarray dataset of myeloid lineage cells from patients with inflammatory arthritis

3.1 Introduction

In order to address the key research objectives of my thesis I decided to employ a systematic approach that involved re-analysing a microarray data set that had been generated in our lab. This microarray set was generated to explore the contribution of peripheral and synovial CD14+ cells in the pathogenesis of RA and PsA and I decided to employ command line tools to process this data as opposed to commercial technology in order to fully exploit the data. Programmes such as R with the limma package(216) use hierarchical models to determine differentially expressed genes that borrow data from the complete data set as opposed to treating each gene on it's own. This approach is an accepted method where there are a large number of observations on a small number of samples, the so called "large p, small n" problem(217).

DNA microarray technology has allowed researchers to interrogate the transcriptome of cells, which are thought to be involved in pathogenic processes in RA. The technology can be employed on primary cells, cell lines, cultured cells and even whole blood. The microarray or "chip" is a physical slide on which gene specific anti-sense oligonucleotides are spotted.

Briefly RNA is extracted from a cell of interest and this RNA is then converted to cDNA in a reverse transcription step and labeled with a fluorescent marker. Affymetrix technology is one channel, that is to say that only one fluorescent marker is employed but the principle is similar for two channel array systems or in the beadarray system employed by Illumina.

Fluorescently labelled cDNA is hybridised to the array and then read by an array scanner. The raw images are processed by software and then in the case of Affymetrix, a .CEL file is the output. The .CEL file is useful because it gives a standard format for the way an array is scanned and also contains useful information such as date of scanning and other metadata. The processes employed in both wet and dry lab steps are summarised in figure 3-1.

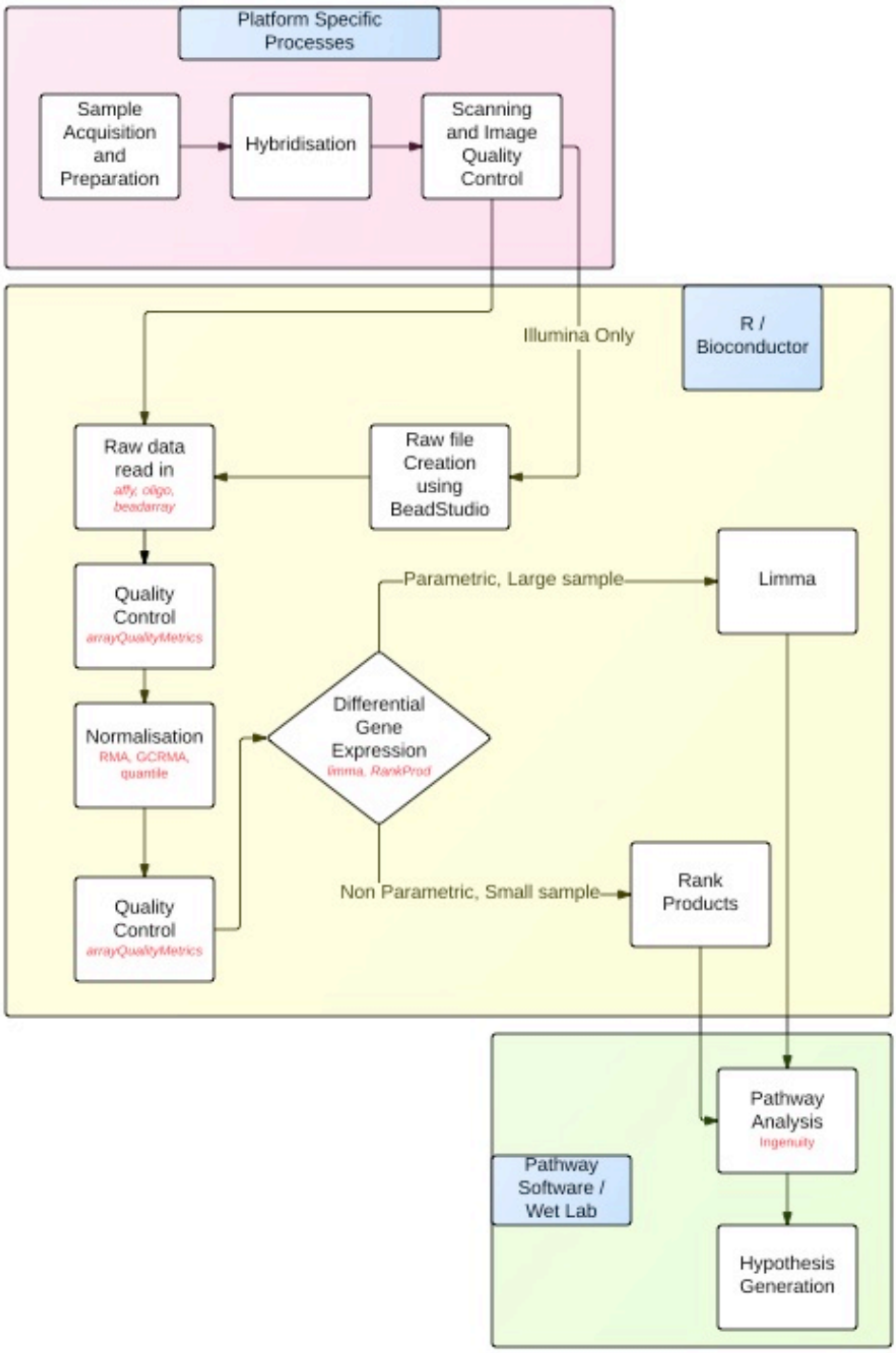


Figure 3-1 Overview of wet and dry lab processes that are required to perform a microarray experiment. Initial steps involve acquiring samples and processing these to generate array files. Thereafter, specific processing occurs computationally with the processes of quality control, normalisation and differential gene expression by different methods occurring according to experimental design. Finally pathway analysis helps researchers analyse results and therefore generate further hypotheses that will be tested using specific follow-on experiments.

Array analysis comes in two forms: end user graphical products or command line tools used by bioinformaticians. Each has advantages and disadvantages that are summarised in table 3-1.

Advantages	Disadvantages
End User Product	
• Out of box solution • Guided analysis • Results often formatted and easily displayed • Customer support	• Black box – little customization possible • Delay in updates to annotation files
Command Line Tools	
• Inherently more powerful for big data sets • Latest analysis packages which are literature driven • Open source • Customisable • Transferrable skills • Community support	• Steep learning curve • Data initially may appear inaccessible • Community support rather than customer support

Table 3-1 Comparison of pertinent features of commercial software packages versus command line tools used to process microarray experiments

The steps taken in the dry lab are summarised in figure 1 and these include initial quality control steps, initial unsupervised analysis and then a form of differential expression analysis. Finally a form of pathway or covariance analysis is often used to make sense of the large amounts of data that are generated.

Therefore in this chapter I set out the following objectives:

1. Investigate how microarray dataset quality can be determined using suitable packages, when these measures can influence downstream analysis and investigate what steps can be taken if data quality is not optimal

2. Develop an analysis pipeline for microarrays and investigate whether batch effects, methods of normalisation and differential expression statistics affect analysis outcome

3. Using the outcome of the above objectives, determine whether a JAK/STAT signature is present in myeloid lineage cells from patients with RA and if this differs from PsA

3.2 Quality control measurements of array files

During RNA extraction the quality of the RNA is measured to determine degradation during sample processing. If RNA is significantly degraded, it should not be processed further.

In addition, there are methods to carry out quality control tests on the array files themselves. The measures shown below look at both raw and normalised data and tend to use unsupervised analyses.

- RNA quality control
- Hybridisation and signal controls
- Overall signal distribution
- Spatial biases and array layout issues
- Probe set analysis
- Correlation between datasets

3.2.1 RNA quality control

There are two methods to assess RNA quality on the Affymetrix array platform: firstly, the degradation from the 5' to 3' end of the beta-actin and GAPDH genes can be shown and finally an overall degradation plot analysed. This is possible because several probes along the terminal and penultimate 3' exon represent each gene and therefore degradation can be calculated as this often occurs from the 5' end. The Human Genome U133 Plus 2.0 is a 3' IVT array and it should be noted that newer arrays such as the Gene 1.0ST and Exon 1.0ST have probes along the length of the gene and therefore allow the assessment of alternative splice variants. Tiling arrays give even more detail and have probes spaces along the whole genome as per figure 3.

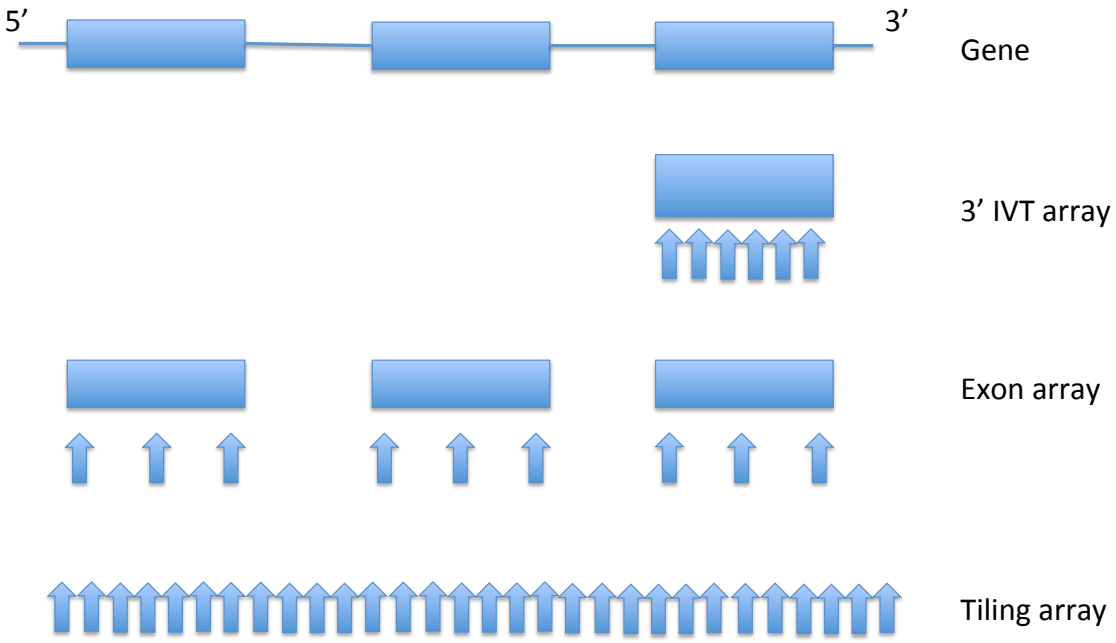


Figure 3-2 Visual demonstration of the differences in probe distribution on 3' IVT, exon and tiling arrays. Older 3' IVT arrays had probes aligned along the terminal exon whereas exon arrays include probes on each exon and therefore when give information regarding alternative splicing. Tiling arrays are probe intensive along the length of the genome.

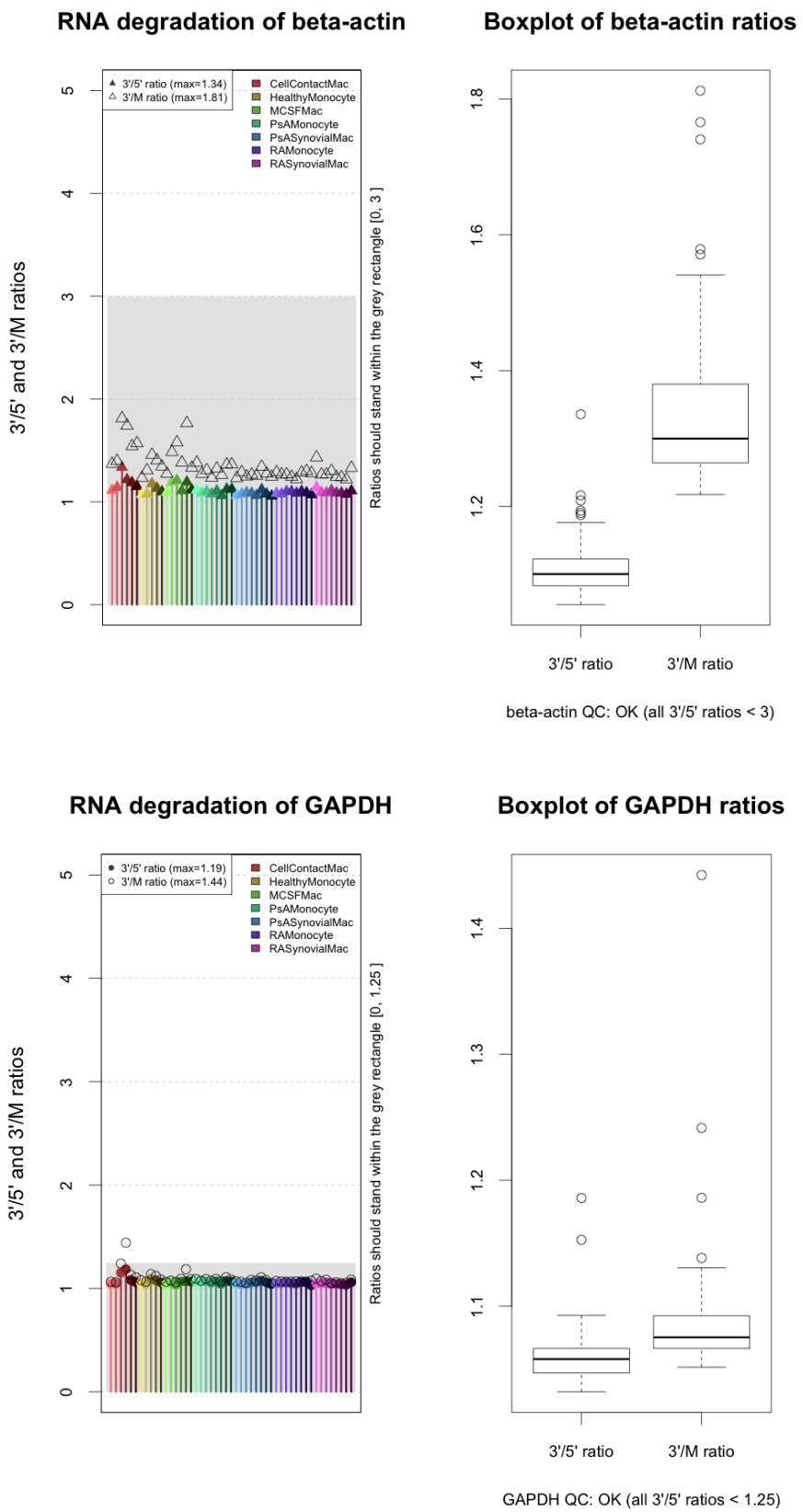


Figure 3-3 RNA degradation plots of beta actin and GAPDH genes shows good RNA quality for primary cells. RNA degradation summary values generated using affyAnalysisQC. Array files were read into R and processed to determine the ratio of intensity of the 3' and 5' probes in both the beta actin and GAPDH genes. In the boxplot ratios of both 3' and 5' probes and also 3' to a midpoint(M) are shown. Although all values are within limits for this analysis, the beta actin gene is showing that in primary culture cells ie MCSF macrophages and also cell contact activated macrophages, there has been degradation of the RNA which is more than other cell types.

Although all of the arrays have passed the QC measure for this test, a few of the cell contact and also MCSF macrophage arrays have poorer RNA quality than the others. It is suggested that a degradation value of 3 be used for beta actin and 1.25 for GAPDH. These are outlined in the plots above by the grey shaded area. The M probe in the boxplot is a probe taken at the midpoint between the 5' and 3' probe in the set in case there has been extreme degradation.

The RNA degradation plot in figure 4 demonstrates that cell contact activated macrophages and some of the MCSF samples have degraded more than primary cells that were separated following blood or synovial fluid draw.

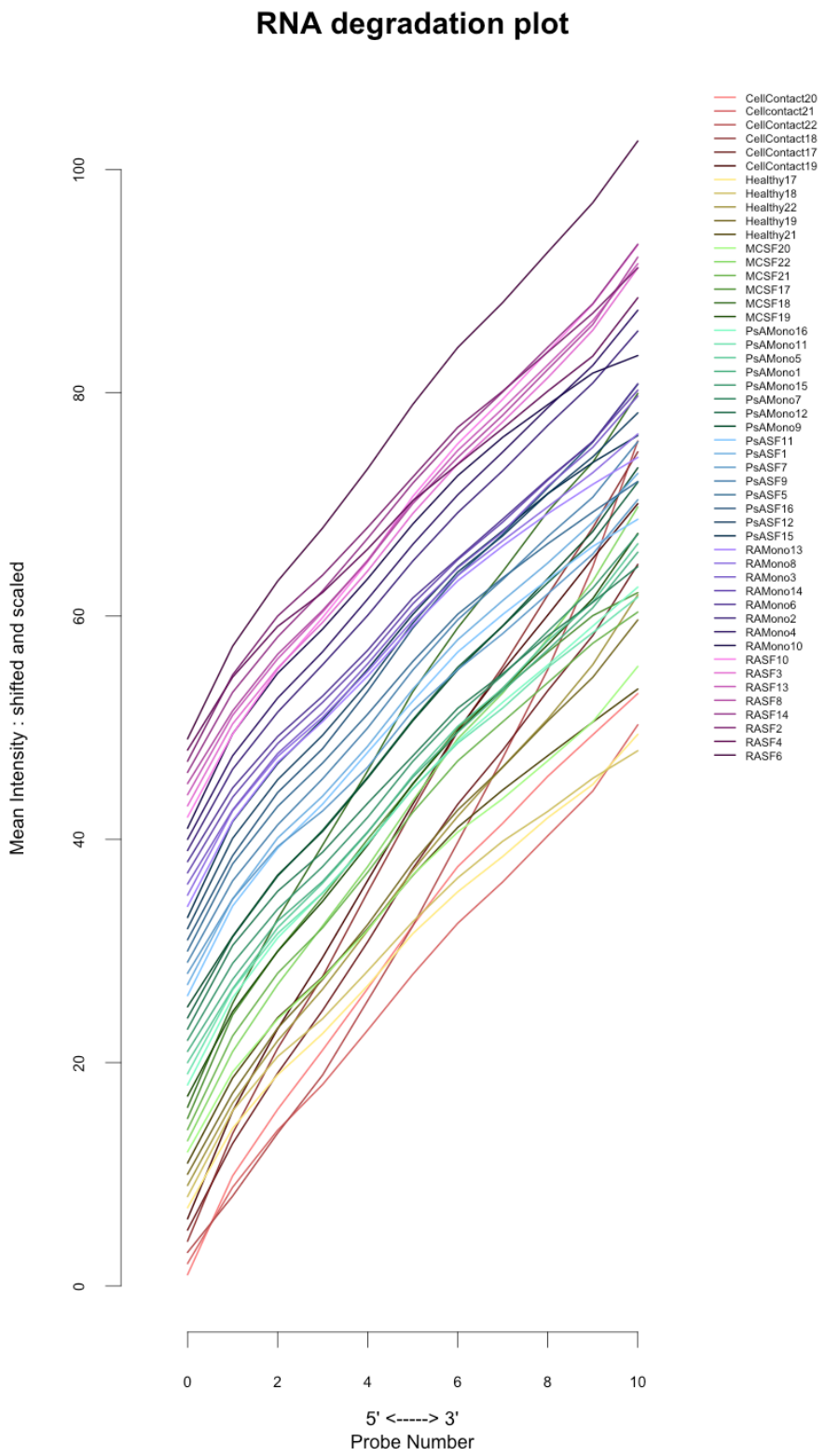


Figure 3-4 RNA degradation plot of probe intensities of summarised genes shows more degradation in cultured cells. RNA degradation plot generated by reading array files into R and processed using affyAnalysisQC. This plot demonstrates the intensity of signal from the 5' to the 3' probes of each gene in the array. Therefore as RNA degradation occurs we expect that there will be a lower signal intensity from 5' probes and therefore the line should slope from bottom left to upper right. Each sample should be parallel showing that RNA degradation is similar in each sample and arrays from the cell contact activated macrophages and MCSF macrophages are showing more degradation.

3.2.2 Hybridisation controls and Signal Quality controls

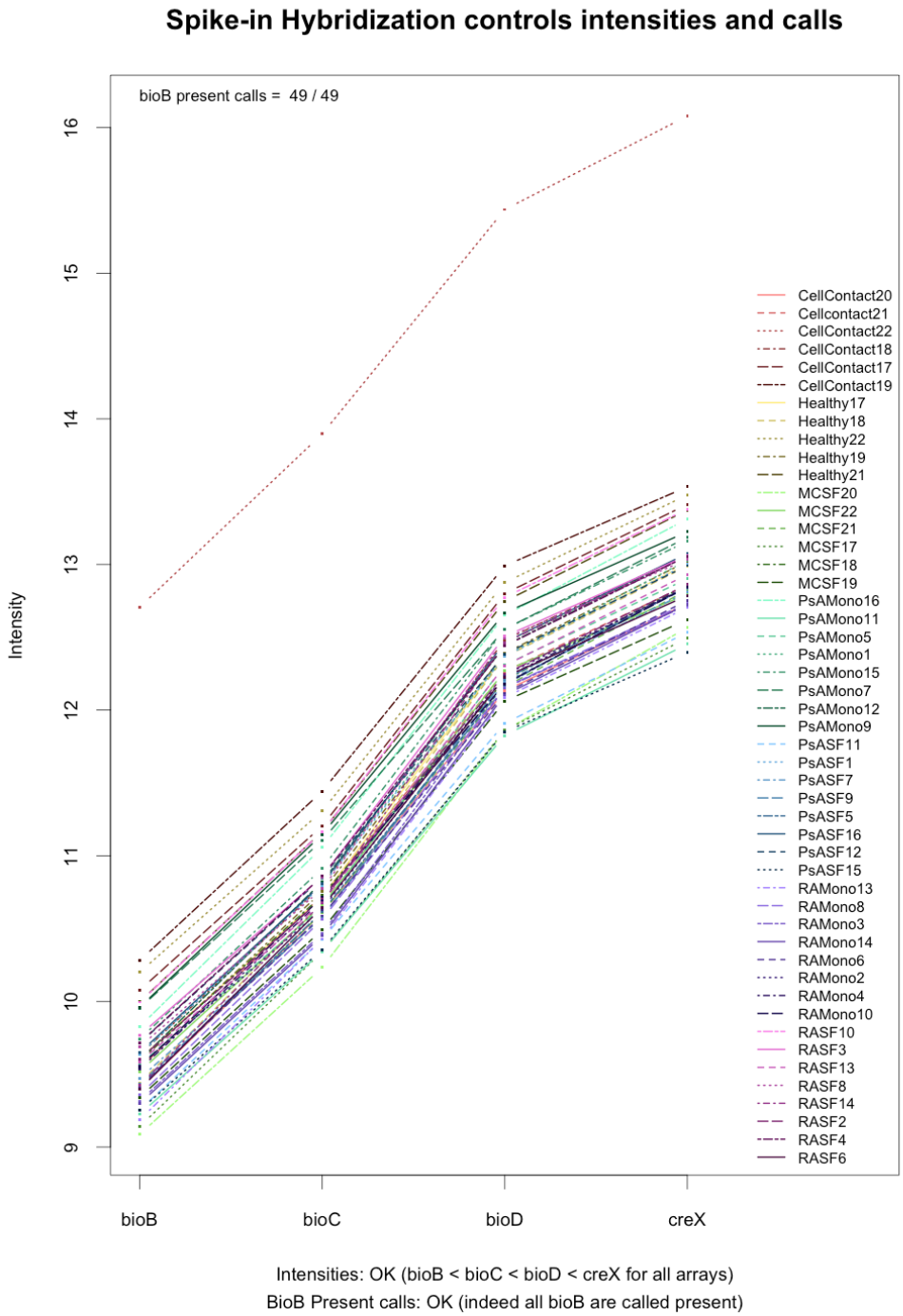


Figure 3-5 Array hybridisation plot demonstrates over hybridisation of a cell contact activated sample. Hybridisation control plot generated by reading array files into R and processed using affyAnalysisQC. This shows the expected distribution of signal intensity across probes although array Cell Contact 22 is showing a much higher intensity suggesting that over hybridisation of this array has occurred and that this array should be removed from further analysis.

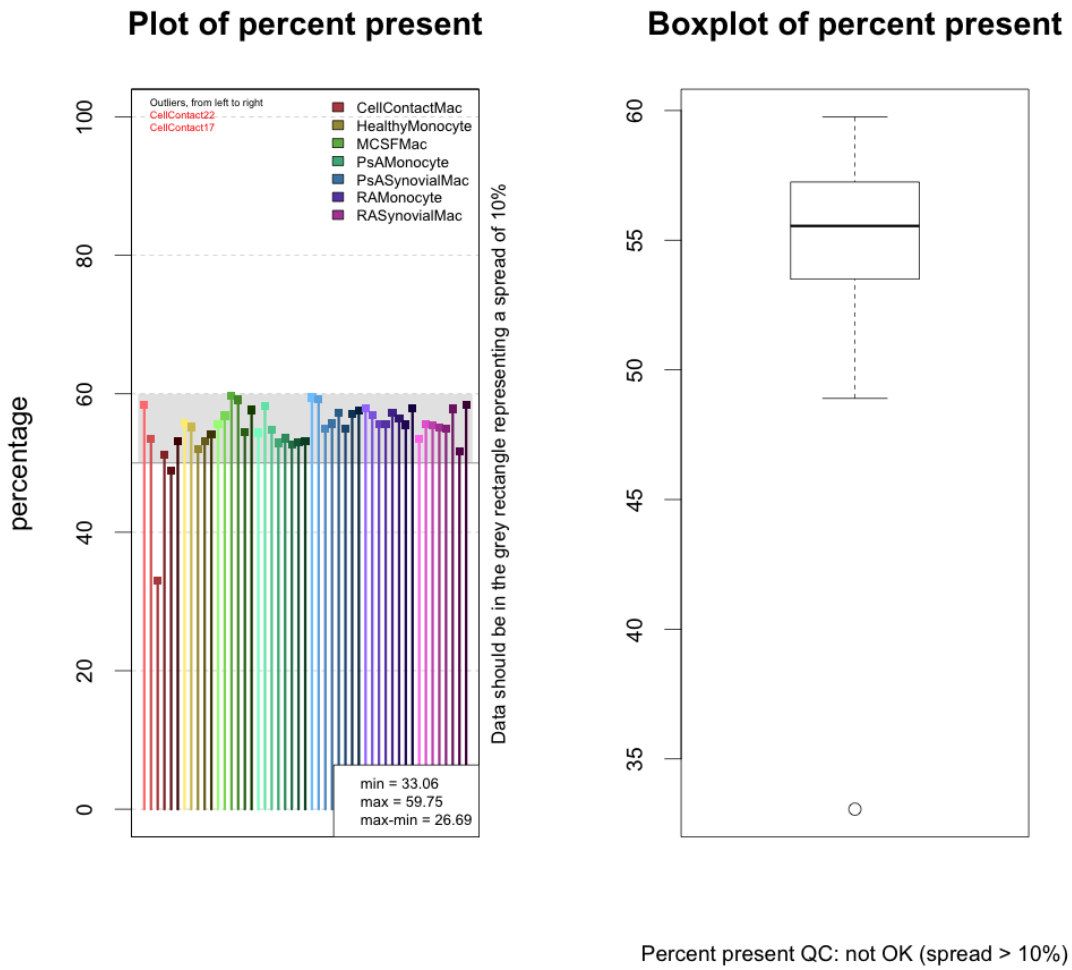


Figure 3-6 Cell contact activated arrays have fewer percentage present genes. A plot of percentage present genes was generated by reading array files into R and processed using the affyAnalysisQC package. It is expected that on any given array that between 50 and 60% of genes will be detectable and therefore significant deviation from this number suggest either technical issues if this is dissimilar to comparable samples. Both Cell Contact 17 and 22 have fewer percent present genes demonstrating by a different method that there is a technical problem with these samples.

The hybridisation control plot demonstrates that the cell contact 22 array is over hybridised and therefore suggests that it should be removed from further analysis (figure 3-5). Additionally the number of genes called as present and detectable is lower in the array cell contact 22 and cell contact 17 suggesting that there may have been a technical problem with these arrays (figure 3-6). In general, the number of present genes should be within 10% between the arrays and anything out with this may be suggestive of a technical failure or batch effect. If they were taken forward to normalisation they may have an adverse effect on experimental analysis because most normalisation methods take account of the variability of genes on all arrays.

3.2.3 Overall signal distribution

Raw data from microarray experiments undergoes a process of normalization in order to remove technical noise. RMA or Robust Multi-array Average is one such method and consists of three steps:

1. Background correction
2. Quantile Normalisation
3. Summarisation and median polish

The process of normalisation log transforms the data and removes positive skew. Normalisation makes an assumption that most of the genes are unchanged and therefore arrays within an experiment should have similar mean signal intensities. This can be visualised in a box plot of raw intensities for each array (figure 3-7) and a density histogram of log intensities (figure 3-8).

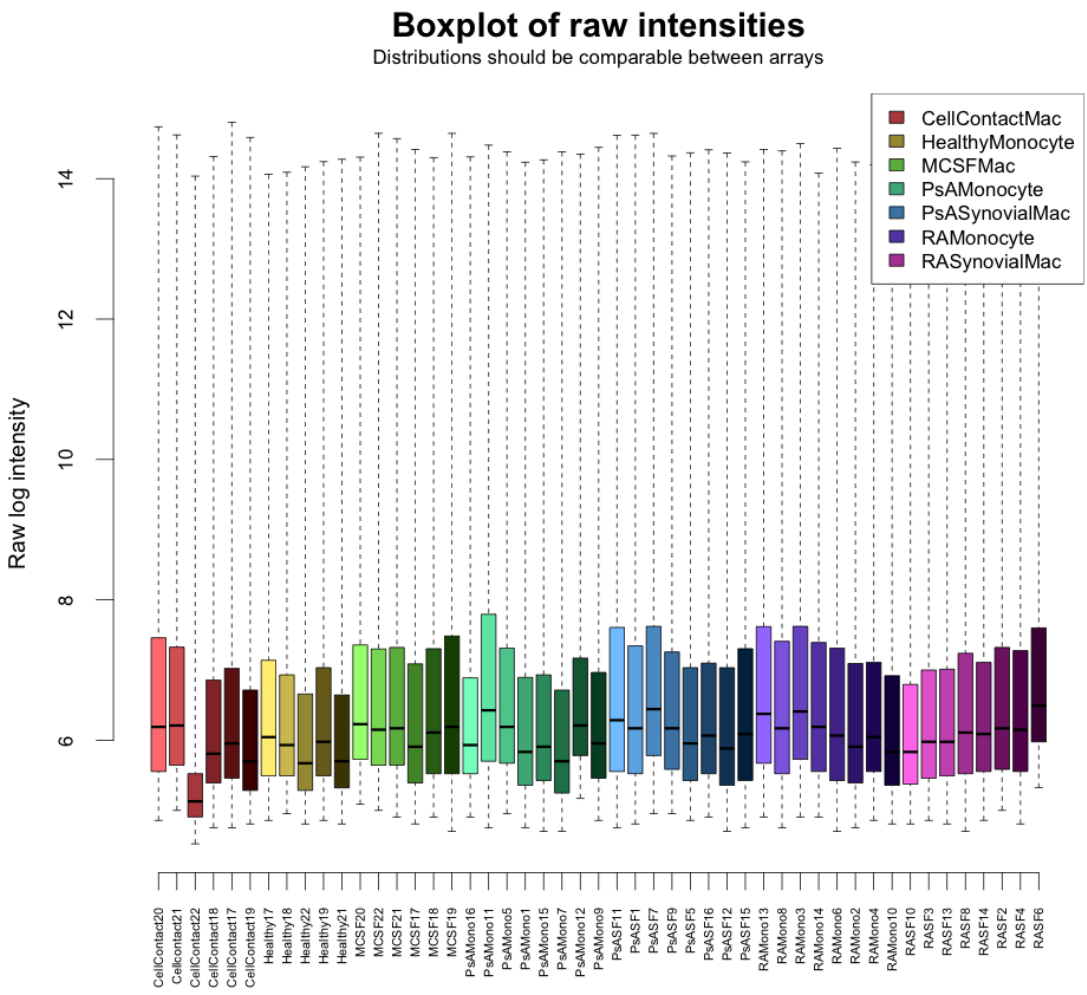


Figure 3-7 Arrays from cell contact experiment have lower raw signal intensities. Boxplots raw log intensity values were generated by reading the array files into R and processed using affyAnalysisQC. The raw log intensity median value is similar for each array except arrays from the cell contact group. There is variation in other array samples but this can be assessed using other methods to decide whether samples should be removed

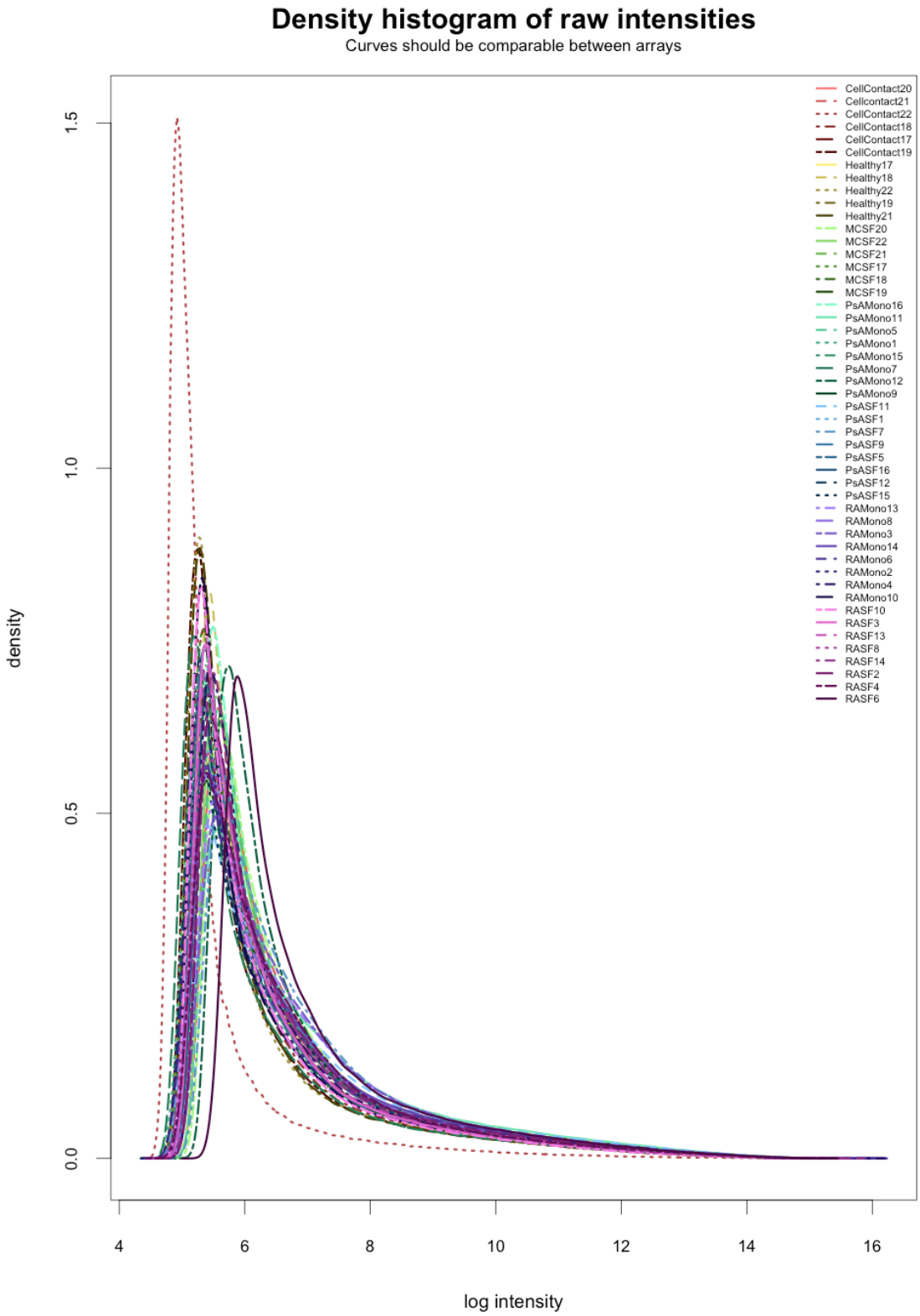


Figure 3-8 A density histogram of raw log intensity shows that cell contact 22 should be removed from further analysis. The density histogram was generated by reading the array files into R and processed using affyAnalysisQC. The signal intensity distribution should be similar for each array but cell contact 22 is an obvious outlier using this QC measure

Quality control after RMA normalisation shows that the technical bias in cell contact 22 cannot be corrected and it is still an outlier (figure 3-9).

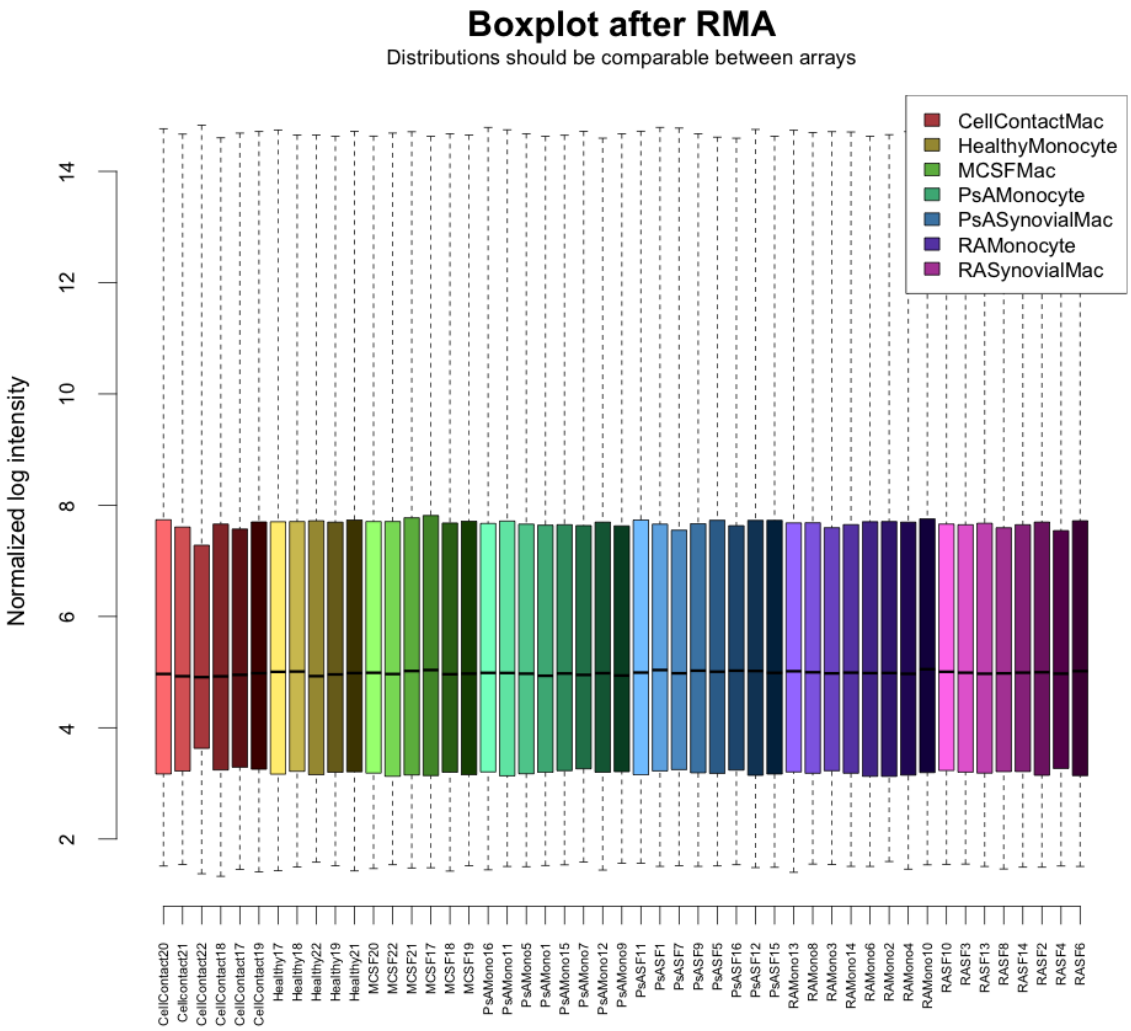


Figure 3-9 RMA normalised intensity boxplot shows that normalisation does not remove abnormal signal distribution of cell contact arrays. RMA normalised intensity boxplots were generated by reading array files into R and processed using affyAnalysisQC. The process of normalisation cannot remove the effects of technical bias and although subtle, some of the cell contact arrays have a narrower signal spread compared to others in the dataset

3.2.4 Assessing Spatial and Probe-set issues within a microarray experiment

The Affymetrix .CEL file holds all of the raw data from a microarray chip once it has been scanned. Visualisation of the chip itself can reveal issues with hybridisation, spatial artifacts or array defects.

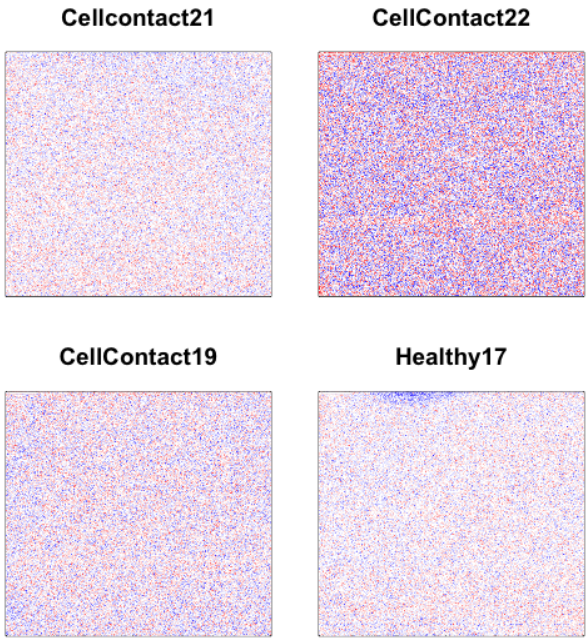


Figure 3-10 Visualisation of raw probe intensities on the array reveals over hybridisation. False colour raw probe intensities were generated by reading array files into R and then processed using arrayQualityMetrics. Visualisation of the raw array intensities shows that cell contact 22 is much brighter than other arrays. Furthermore some arrays show edge effects such as healthy17 but are taken account of by the design where probe sets, which are up to ten probes per gene, are distributed across the array and outliers are taken account of during RMA processing. The colourisation is based on the signal distribution of all probes on that array

Therefore array cell contact 22 has much higher signal intensity compared to other arrays thereby suggesting that it should be removed from analysis (figure 3-10). Furthermore edge effects are demonstrated on array healthy 17 although this has not affected other parts of the array and due to probe redundancy, these effects can be mitigated by taking median or mean values of a complete probe-set. Finally unsupervised correlation methods are used to determine outliers.

3.2.5 Correlation of arrays using unsupervised methods

I used three methods to assess the arrays in an unsupervised manner:

1. A correlation plot that assesses the raw correlation of signal intensity between arrays
2. Principal Component Analysis of PC1 and PC2 of signal intensity in each array
3. Hierarchical clustering of signal intensity in each array

Figure 3-11 shows the correlation between each array with blue demonstrating similarity. In this case, only array 3, which corresponds to cell contact 22 is abnormal. Furthermore, the PCA plot of PC1 and 2 (figure 3-12) show that this array is an outlier and this is confirmed with hierarchical clustering (figure 3-13). The PCA and hierarchical clustering also show that there is an underlying structure to the data with monocytes, synovial macrophages and monocyte-derived macrophages clustering in separate groups.

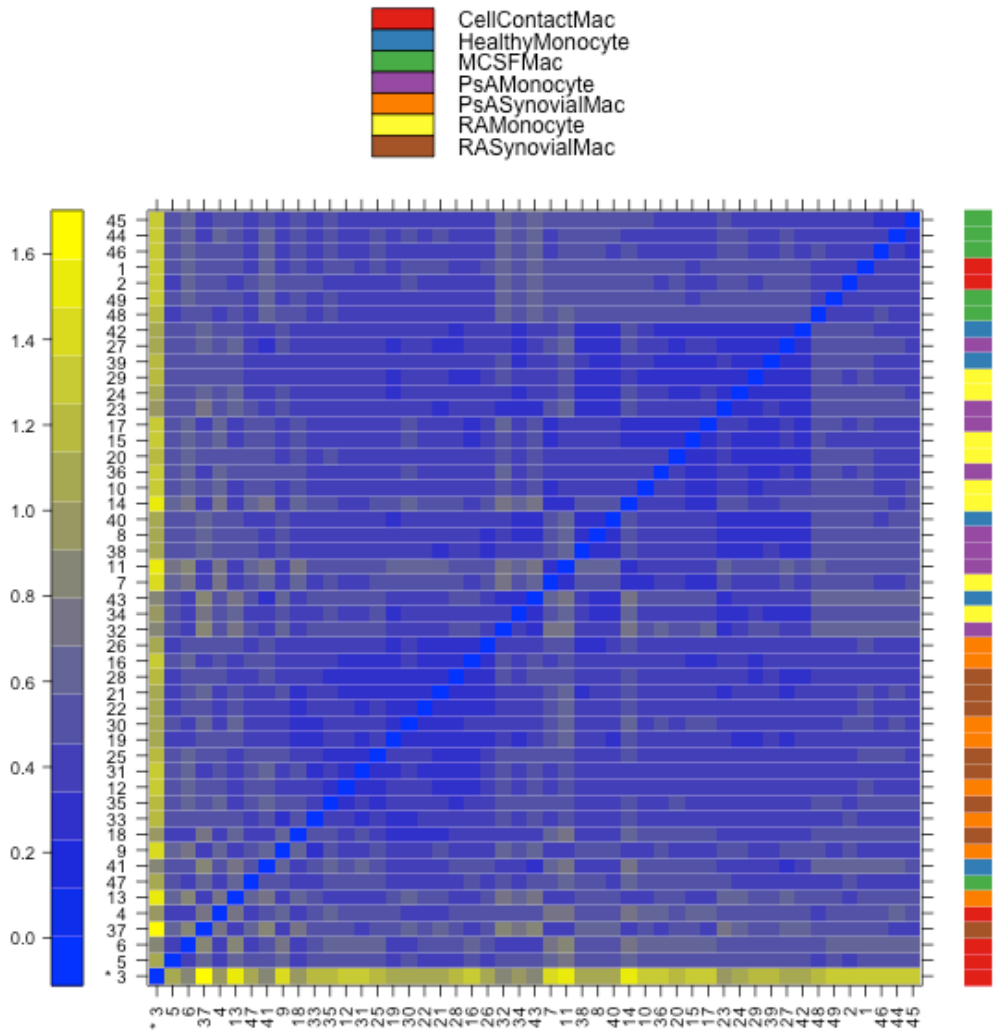


Figure 3-11 A correlation plot confirms that the signal intensity of cell contact 22 is unlike other arrays. A correlation plot was generated by reading array files into R and processed using arrayQualityMetrics. The plot represents a measure of the total distance between arrays where the distance is calculated by all of the probe intensities on the array. Array 3 which corresponds to cell contact 22 is dissimilar to all other arrays in the experiment and is therefore an outlier by this test

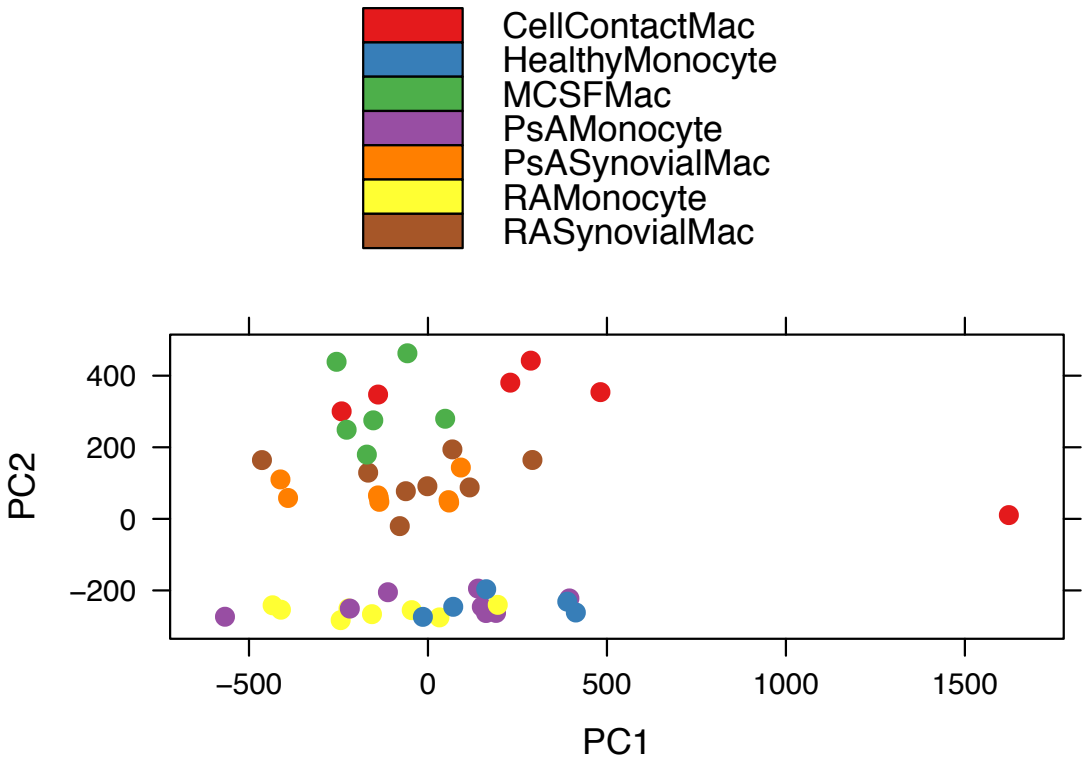


Figure 3-12 PCA of array signal intensities demonstrates that cell contact 22 is an outlier. PCA plot of array signal intensity were generated by reading array files into R and processed using arrayQualityMetrics. The PCA demonstrates that there is one outlier array that corresponds to cell contact 22 although it is difficult to assess other arrays due to the large variability that is contributed by that one array. However, in this plot, the monocyte samples are clustering together as are the synovial macrophages and cultured macrophages.

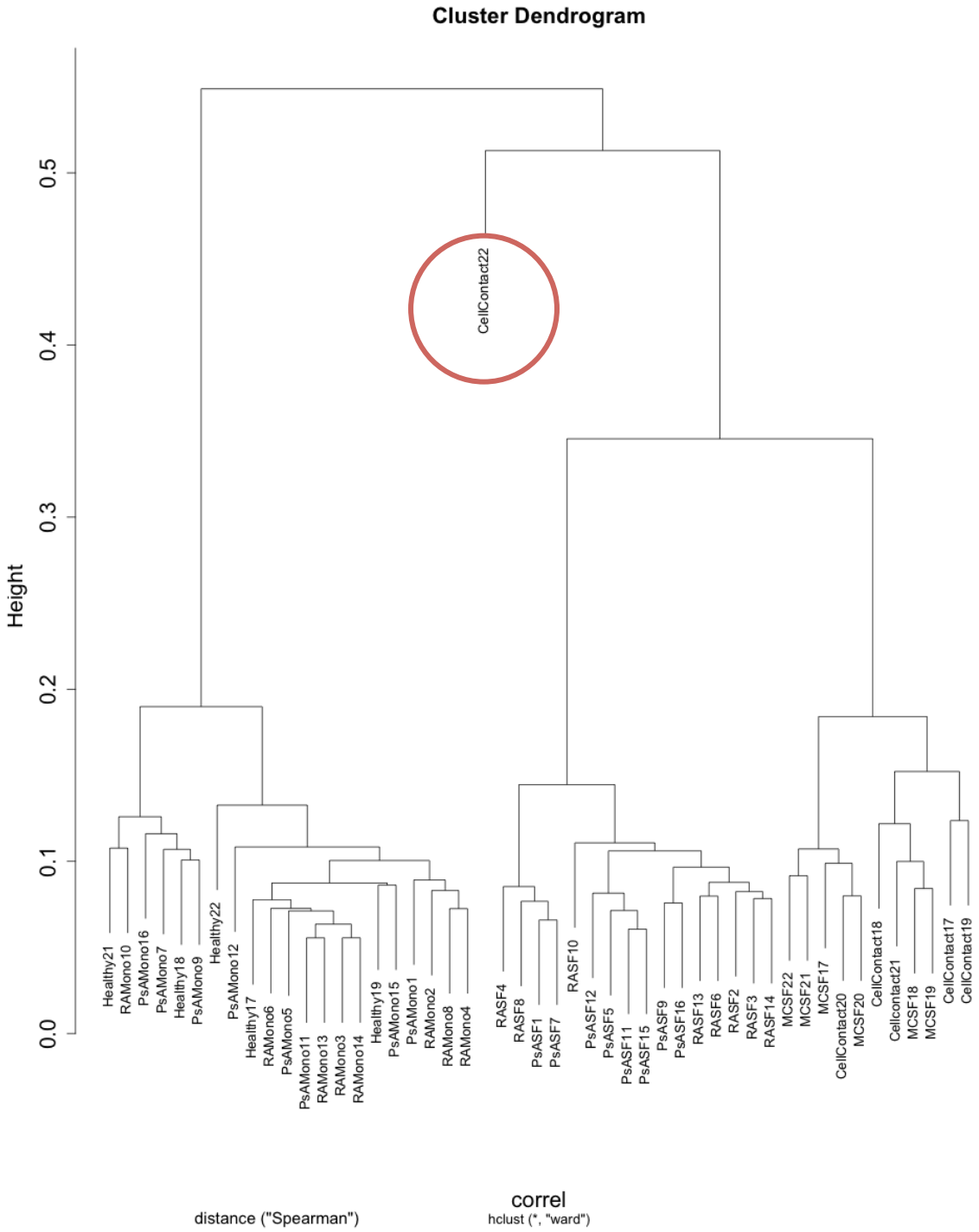


Figure 3-13 Hierarchical clustering demonstrates that monocytes, primary macrophages and cultured macrophages cluster in separate groups. A hierarchical clustering plot was generated by reading array files into R and processed with affyAnalysisQC. Clustering demonstrates that monocytes and macrophages cluster together. Within the macrophages, diseased macrophages form a sub-group as do the MCSF and cell contact activated macrophages. Cell contact 22 is shown as an outlier

3.2.6 Conclusion from quality analysis of microarray files

In conclusion, following extensive quality analysis of the microarray data, I have decided to concentrate my further investigation of this dataset to the cells that have been derived from patients and also healthy monocytes. The quality control of cultured cells revealed problems with RNA quality of various cell contact activated macrophage samples and some of the MCSF macrophage samples especially in the RNA degradation plot. The implications of this could mean that the gene lists that are derived from these arrays are subject to more bias and therefore lead to an unreliable result. This is particularly true of the cell contact activated macrophages and it is likely that because these cells had to undergo cell sorting after co-culture, that the RNA from these samples is unreliable for microarray purposes.

Furthermore, although analysis of the transcriptome of MCSF and cell contact activated macrophages compared to monocytes would be useful, it is not aligned to the core objectives of this chapter: to determine whether a JAK/STAT signal is present in RA myeloid lineage cells and if this differs from that found in PsA.

Therefore I decided to proceed with the analysis of only the blood monocytes and synovial macrophages derived from RA and PsA patients and blood monocytes from healthy donors to achieve my next objectives namely: developing an analysis pipeline for microarrays taking account of methods of normalisation and differential expression and determining whether a JAK/STAT signature is evident in RA myeloid lineage cells.

3.3 Developing and optimizing an analysis pipeline for microarrays by investigating methods of normalisation, differential gene expression and batch effects

3.3.1 Introduction

Once raw data is acquired, converted to .CEL files and been subject to quality control it is normalised. There are many methods to normalise microarray data from simple log2 transformation of the raw data to correcting for background and mismatched probes.

The Affymetrix microarray platform has gene probes that have both a perfect match (PM) and a mismatch (MM). In theory this could be used to account for non-specific binding and therefore give a clearer signal for the perfectly matched probe. However the MAS5.0 algorithm from Affymetrix, which takes account of the mismatched signal, occasionally resulted in a negative intensity for the perfect matched probe once the mismatched one had been taken away especially at low intensities(218). The mismatched probe exhibited higher signal intensity than the perfect-match probe. Possible explanations maybe that a mismatch for one gene turns into a perfect match for another. Furthermore RNA hybridization to a chip depends on more than the raw sequence of the probe and binding affinities vary depending on GC content and three-dimensional structure of the RNA in question.

Newer methods of normalization have been developed which do not take account of the mismatched probes but instead normalised across chips. The first of these methods was developed in the Irizarry and Bolstad lab and was called RMA(219), Robust Multiarray Averaging. RMA carries out three processes:

- Background signal correction
- Quantile normalisation of the data
- Probe level summarisation and median polish

Briefly this means that RMA takes account of the background fluorescence of the chip but not the mismatch probes. It then normalises the distribution of the signal across all arrays so that it is comparable in each quantile. It is important to note that if the data is an outlier and the signal intensity distribution is very different such as in cell contact 22, the quantile

normalisation process will "squeeze" the data to make it fit into the distribution of the other arrays. This will lead to false positive results downstream and a failure to validate analysis results.

Probe intensity is measured individually and then summarised so that a single value is derived for each gene. This takes account of the fact that each gene is represented by up to 20 probes. There is variation between measurements and so the process of median polish takes the median intensity of the chip and also the probe set for that gene into account during normalisation.

RMA is a standard method for normalisation of Affymetrix microarrays however it does not take account of the mismatch data. Although this can be a problem if simply subtracted, a newer method of normalisation called gene-chip RMA or GC-RMA was developed. This method took account of the probe sequence and calculated an altered value of intensity based on predicted probe affinity. Therefore GC-RMA had the advantages of RMA but was still able to take account of mismatched data. In reality GC-RMA may lead to differences in probe intensities that often resulted in false positive results and was also computationally expensive. This is not such an issue now with the advent of faster CPUs and the decreasing price of RAM however the need for a validation cohort was highlighted by the fact that the normalisation method alone could result in a difference in the number and intensity of differentially expressed genes.

I also assessed a final method of normalisation, built on RMA, called Frozen RMA(220) or fRMA. This method is particularly useful if batch effects are discovered in datasets or when samples need to be collected incrementally such as in clinical trials. The probe effects and variances are calculated from a large archive of publicly available data and then "frozen" so that they can be applied to as few as one array. This is in contrast to RMA and GC-RMA where these factors are worked out on the complete array data that has been read in.

Prior to investigating methods of normalisation, I will briefly describe the experimental setup, assess if there has been a batch effect and give an example of a microarray data set where there clearly has been an issue with batch. The reason for this is because if a batch effect is present, it would change my choice of normalisation and I would also need to take account of the effect in the differential expression model.

3.3.2 Probe intensity boxplots and PCA of myeloid lineage cells from healthy volunteers and patients reveal no batch effect

Briefly, the microarray data was generated by obtaining paired blood and synovial fluid samples from patients with RA or PsA and obtaining CD14+ cells by density centrifugation and magnetic bead selection. Healthy volunteers had only blood CD14+ cells with no corresponding synovial sample because of absence of disease. A table of patient characteristics is included in the appendix.

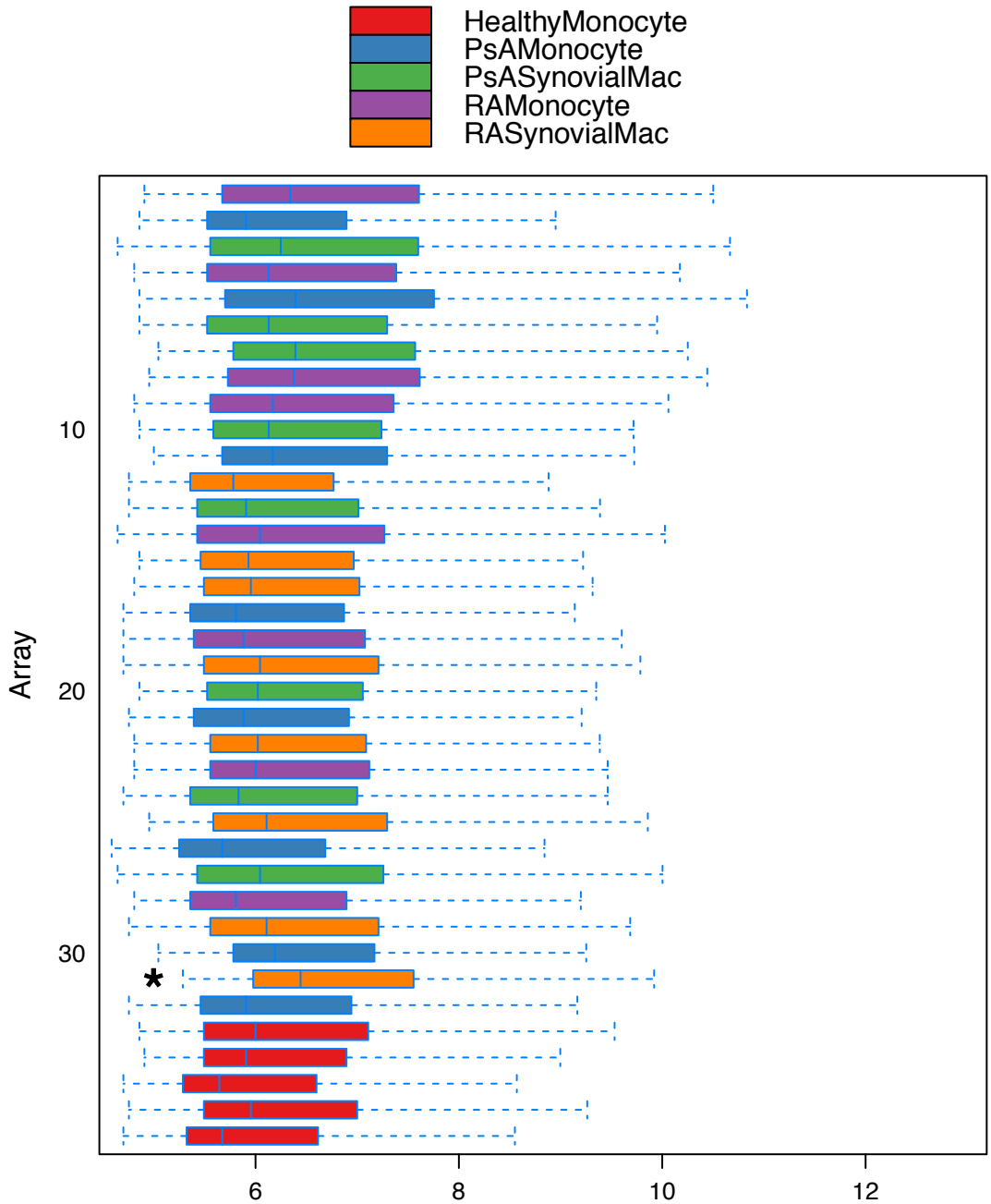


Figure 3-14 Raw intensity boxplots do not reveal a significant batch effect. Microarray files were read into R and processed using arrayQualityMetrics. The log transformed raw intensity boxplots have medians centred at 6. * denotes an array that has been determined as an outlier using the Kolmogorov-Smirnov statistic between that array's distribution and the distribution of the pooled data.

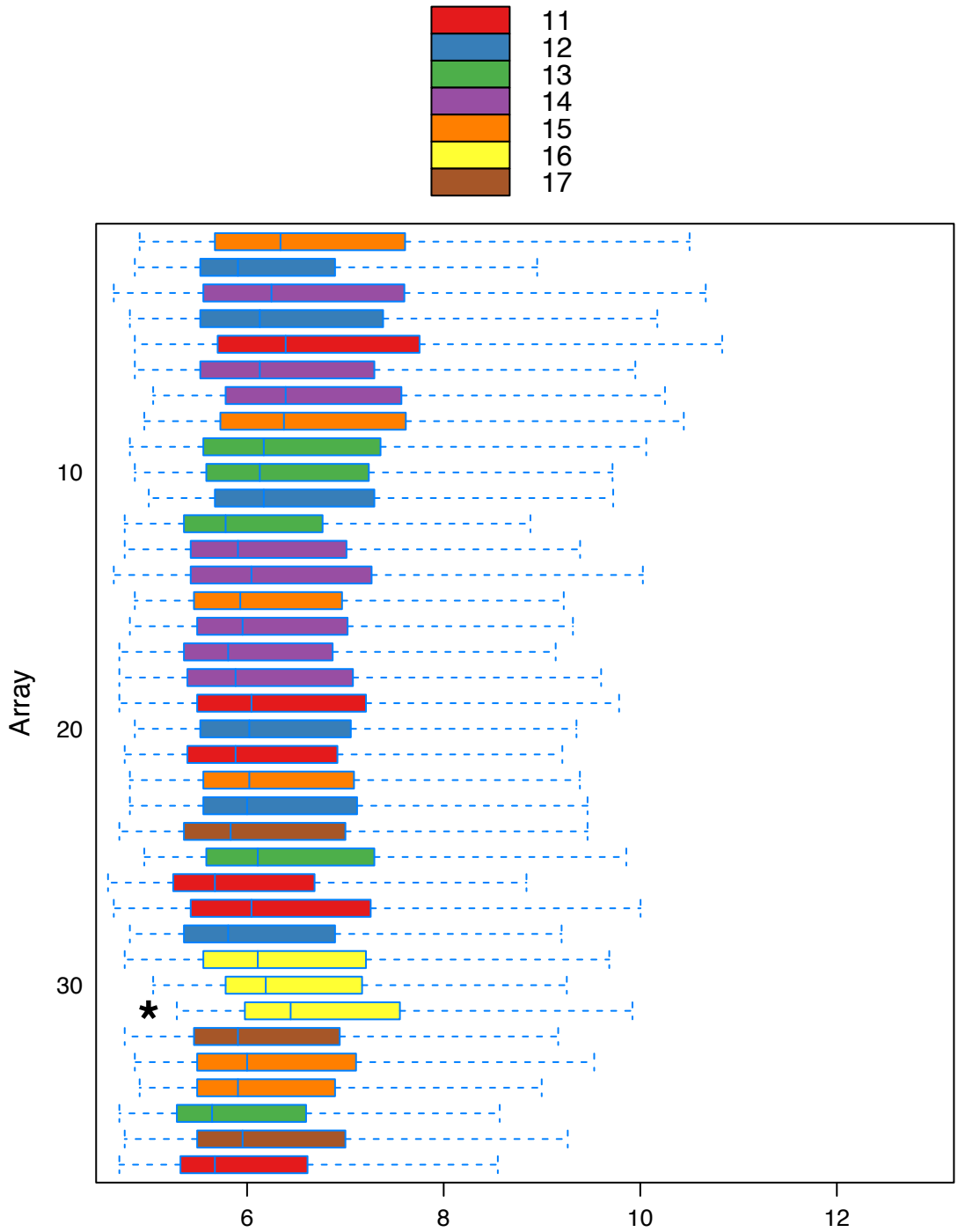


Figure 3-15 Raw intensity boxplot coloured by scanning time. Microarray files were read into R and processed using arrayQualityMetrics. The boxplots have been coloured by the hour of scanning using a 24 hour clock. This reveals no obvious batch effect.

On assessment of the raw intensity boxplots (figure 3-14 and 15) and PCA I conclude that there is no significant batch effect detectable and therefore there is no need to account for a batch effect in the differential expression model (figure 3-16). The lack of batch effect is likely due to the experimental design in that one research collected the samples and extracted the RNA. They RNA was shipped frozen to a core genomics facility and the array processing carried out in one sitting. The date of scanning for each of the arrays was

the same but in figure 3-15 I have coloured the boxplot according to the hour of scanning and no obvious batch effect can be seen. Therefore, frozen RMA is unlikely to confer benefit in this analysis and may over fit the data in this experiment due to the fixed array correction values that are generated from public data although I will still test this.

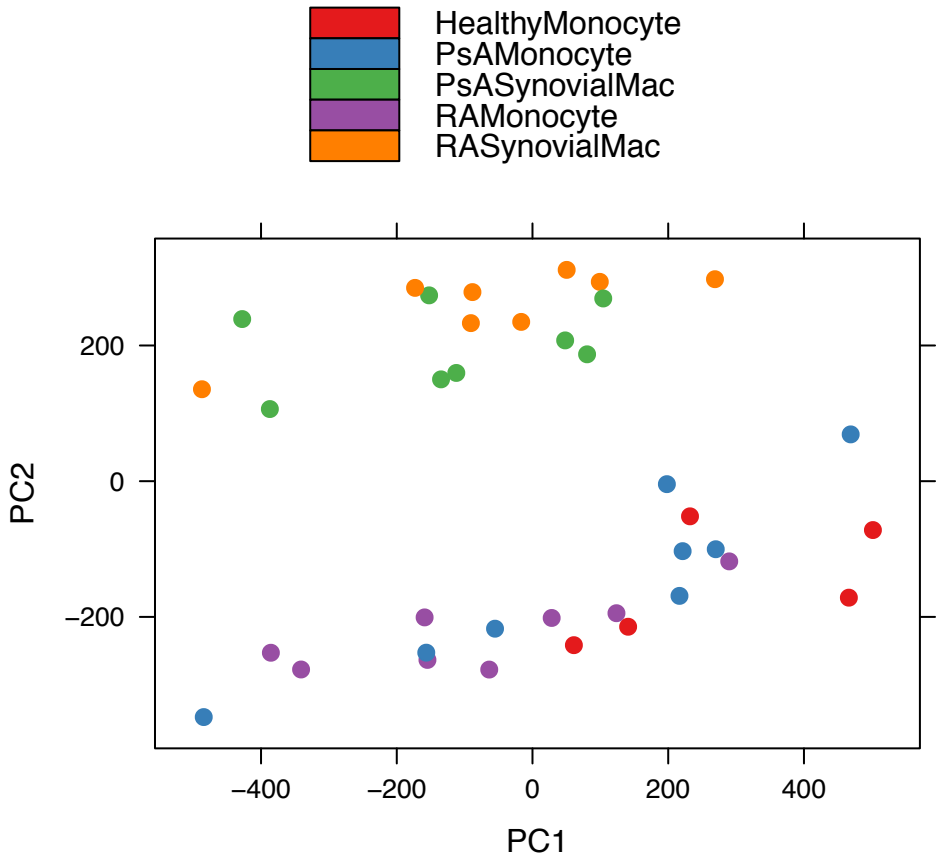


Figure 3-16 PCA plot of microarray data shows grouping by cell type and not disease but does not reveal a batch effect. Microarray files were read into R and processed using arrayQualityMetrics. Synovial macrophages and blood monocytes are clustering in two groups which are separated along the PC2 axis in the raw data. There is no evidence of a batch effect.

Although batch effect is not evident in my dataset, data generated by a student whom I supervised did show this (figure 3-16 and 3-17). He analysed a publicly available microarray study of PBMC from healthy volunteers and patients with RA and Systemic Lupus Erythematosus (SLE), a connective tissue disease.

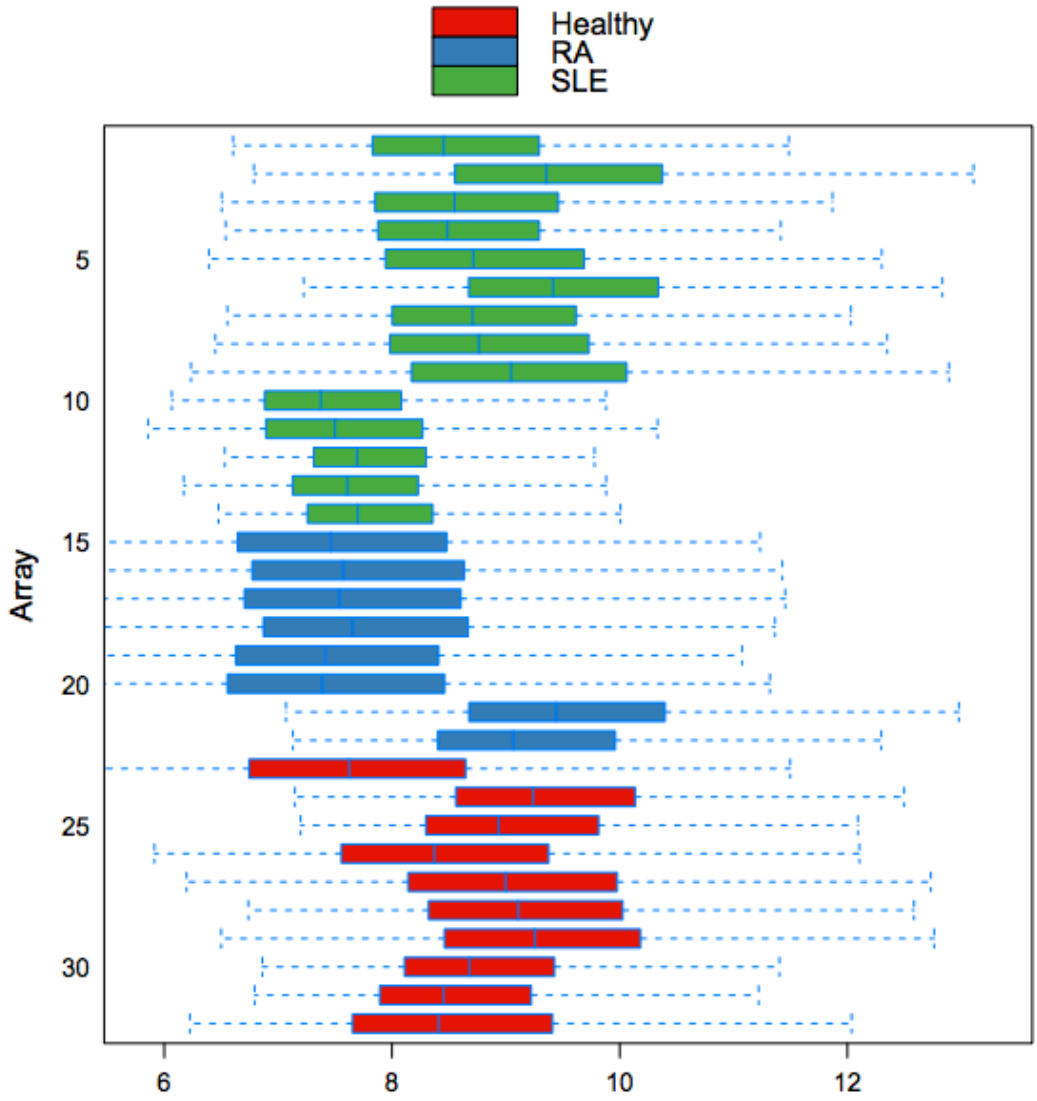


Figure 3-17 Raw intensity boxplot of microarray data from healthy volunteers, RA and SLE patients reveals multiple subgroups within disease. Microarray data was obtained from the ArrayExpress public archive and read into R and processed using arrayQualityMetrics. Within all disease conditions there is variation in the median signal intensity suggesting that a batch effect is present. On further analysis of this dataset very few differentially expressed genes were found between diseases following differential expression analysis

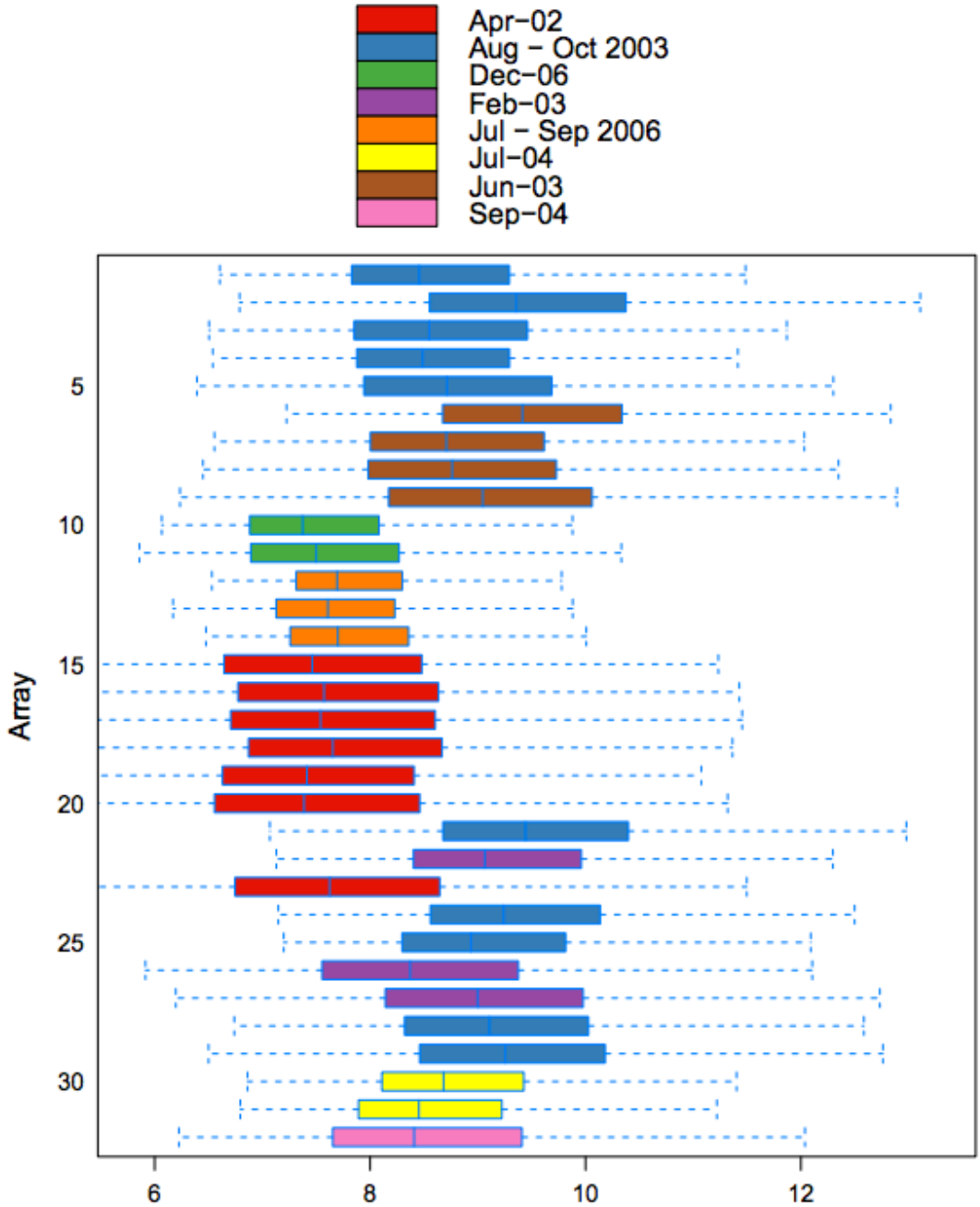


Figure 3-18 Raw intensity boxplot of microarray data from healthy volunteers, RA and SLE patients coloured by date of reading of microarray reveals a cause of the batch effect. Microarray data was obtained from the ArrayExpress public archive and read into R and processed using arrayQualityMetrics. The intensity boxplots are the same as in figure 16 but are coloured by the date of reading of the microarray chip which reveals that the date of reading is a significant effect.

This data demonstrates a batch effect with the time of reading of a chip contributing the to the effect. This suggests that samples from each of the conditions were collected over a period of years and run in small groups. Although this is occasionally unavoidable in studies, experimental planning from the outset may have been able to avoid this problem. However in this case, use of the ComBat (221)batch correction package prior to differential expression resulted in more differentially expressed genes for exploration although they would need to be carefully validated in subsequent experiments.

To assess whether the method of normalisation had an effect on the number of differentially expressed genes, I proceeded to take another data set and analyse it using a standard method in limma
(216,222). Limma is an R package that takes a normalised microarray dataset and then carries out differential expression analysis by using Bayes moderated t-tests. Furthermore it can also correct for multiple testing because in a typical microarray dataset, if it is unfiltered, we will be carrying out over 50,000 t-tests.

I used the murine data set described in the next section because it is small, with few replicates and therefore was ideal to investigate whether a non parametric or a parametric statistic for differential expression is useful.

3.3.3 Comparing parametric with non parametric testing for generation of differential gene lists

Briefly, in this microarray data set, male C57BL/6 wild type and microRNA 155 knockout mice were fed on a high fat diet from 6 weeks old. Total RNA was extracted from homogenised liver and a mouse microarray performed to determine genes that may be differentially expressed in the case of microRNA 155 knockouts. Raw and RMA normalised quality control plots are in Figures 3-18 and 3-19.

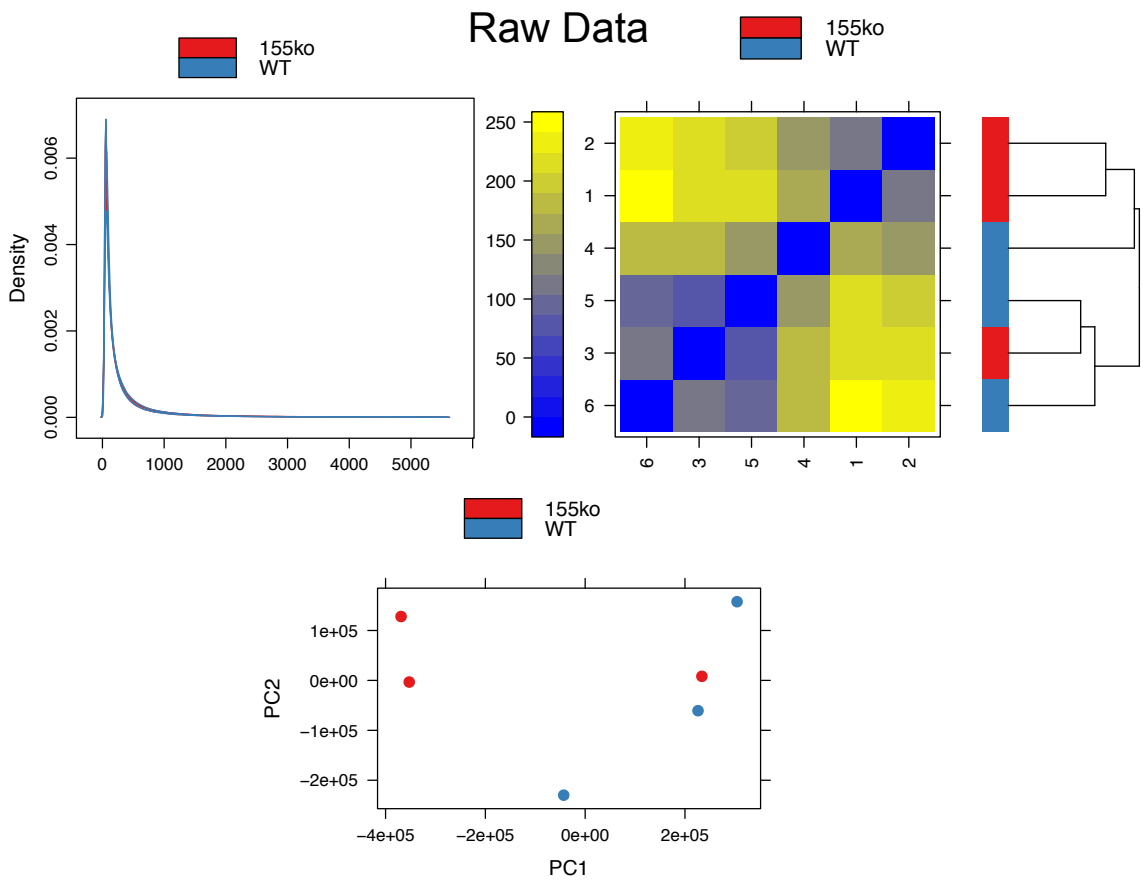


Figure 3-19 Raw quality control plots from wildtype versus microRNA 155 knockout mice show that data from one knockout array clusters with wildtype. Male C57BL/6 wild type and microRNA 155 knockout mice (n=3 each group) were fed on a high fat diet from 6 weeks old, livers harvested, total RNA prepared and a Mo Gene 1.0 ST array performed. Microarray files were read into R and processed using the oligo and arrayQualityMetrics packages. A density histogram shows no issues with signal distribution of raw data but both the correlation plot and PCA demonstrate one knockout array clustering with the wildtype

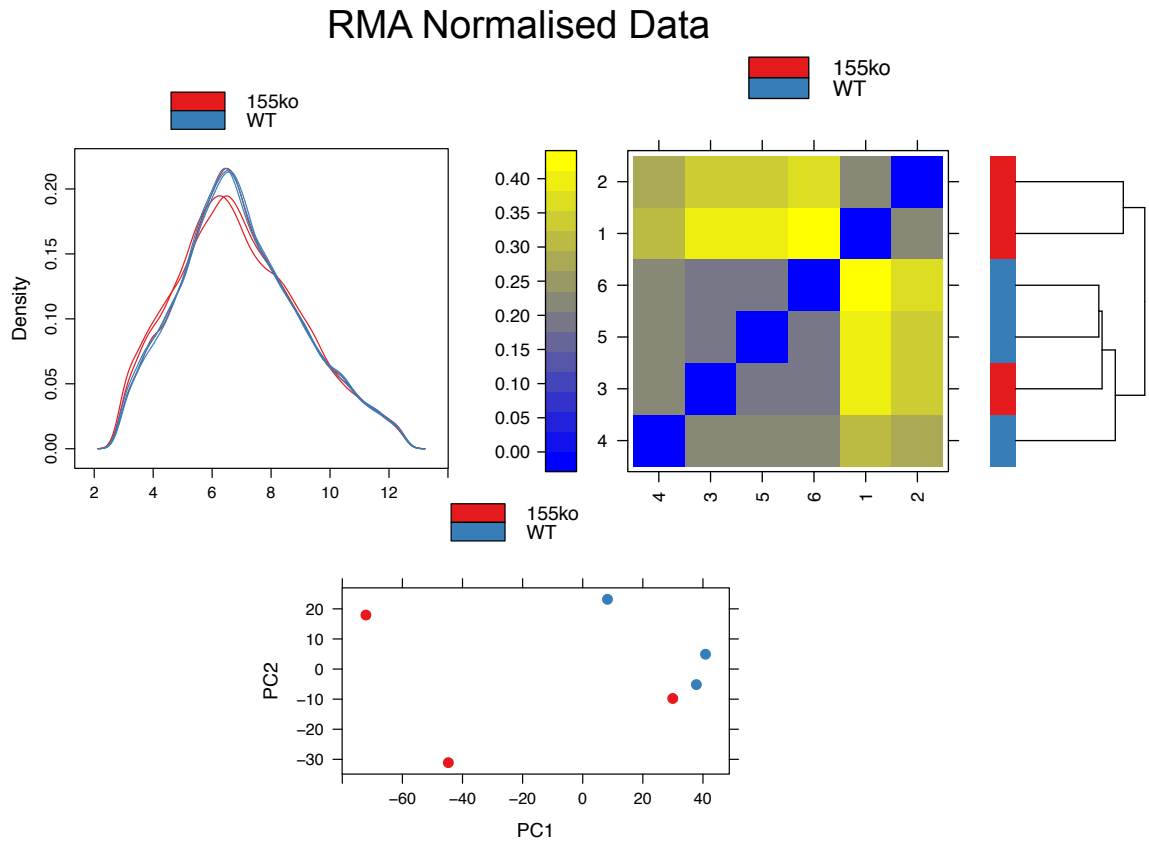


Figure 3-20 Normalised quality control plots from wildtype versus microRNA 155 knockout mice show that data from one knockout array clusters with wildtype. Male C57BL/6 wild type and microRNA 155 knockout mice (n=3 each group) were fed on a high fat diet from 6 weeks old, livers harvested, total RNA prepared and a Mo Gene 1.0 ST array performed. Microarray files were read into R and processed using the oligo and arrayQualityMetrics packages and normalised using RMA. A density histogram shows a normal signal distribution for both wildtype and knockout data. Both the correlation plot and PCA clearly demonstrate one knockout array clustering with the wildtype.

Figure 3-18 shows a density plot in the expected distribution for raw data. However both the correlation plot and PCA show that one of the knockout arrays is clustering with wild type. This is also apparent when the data is normalised (figure 3-19).

In this experiment there are three biological replicates in two groups and therefore discarding one array is going to have significant consequences for the experimental validity. Furthermore, the limma pipeline works best for samples where the replicates in each group is over 5 however I performed this in the first instance but no genes were differentially expressed with an adjusted p value of <0.05(223).

Therefore I decided to explore the use of non-parametric methods of differential gene expression to discover potentially differentially expressed genes in this difficult and noisy dataset. I decided to use the Rank Products(224) based method for differential gene expression because it is particularly useful for small and variable data sets(225,226). In this method the fold changes of genes is ranked in each replicate and then combined to give the rank product for each gene. Therefore if a gene is upregulated in two out of three samples in the same group, the rank product will not be affected as much as a t test as the relative rank is taken into account as opposed to taking the mean. This experimental design is typical for exploratory murine experiments where the assumption is that the inbred nature of the mouse will keep non specific biological variation to a minimum and therefore a wild type verses a knockout experiment should reveal changes in the transcriptome which are statistically significant even when using small numbers.

Therefore, I carried out a Rank Products based analysis on the complete data set along with a limma analysis removing the array that was clustering with the wild type samples. It should be noted that this was a purely exploratory analysis and that removing a data point because it is inconvenient could be considered manipulation of the data. However in this circumstance all of the aberrant array information is clustering with the wild type arrays and this was carried out in an unsupervised manner.

Employing a non-parametric statistic i.e. the Rank Product I was able to use the complete data set but still determine potentially differentially expressed genes. Figure 3-20 shows the overlap of probes that are differentially expressed when using three approaches: the original limma testing involving all of the data, limma with the aberrant

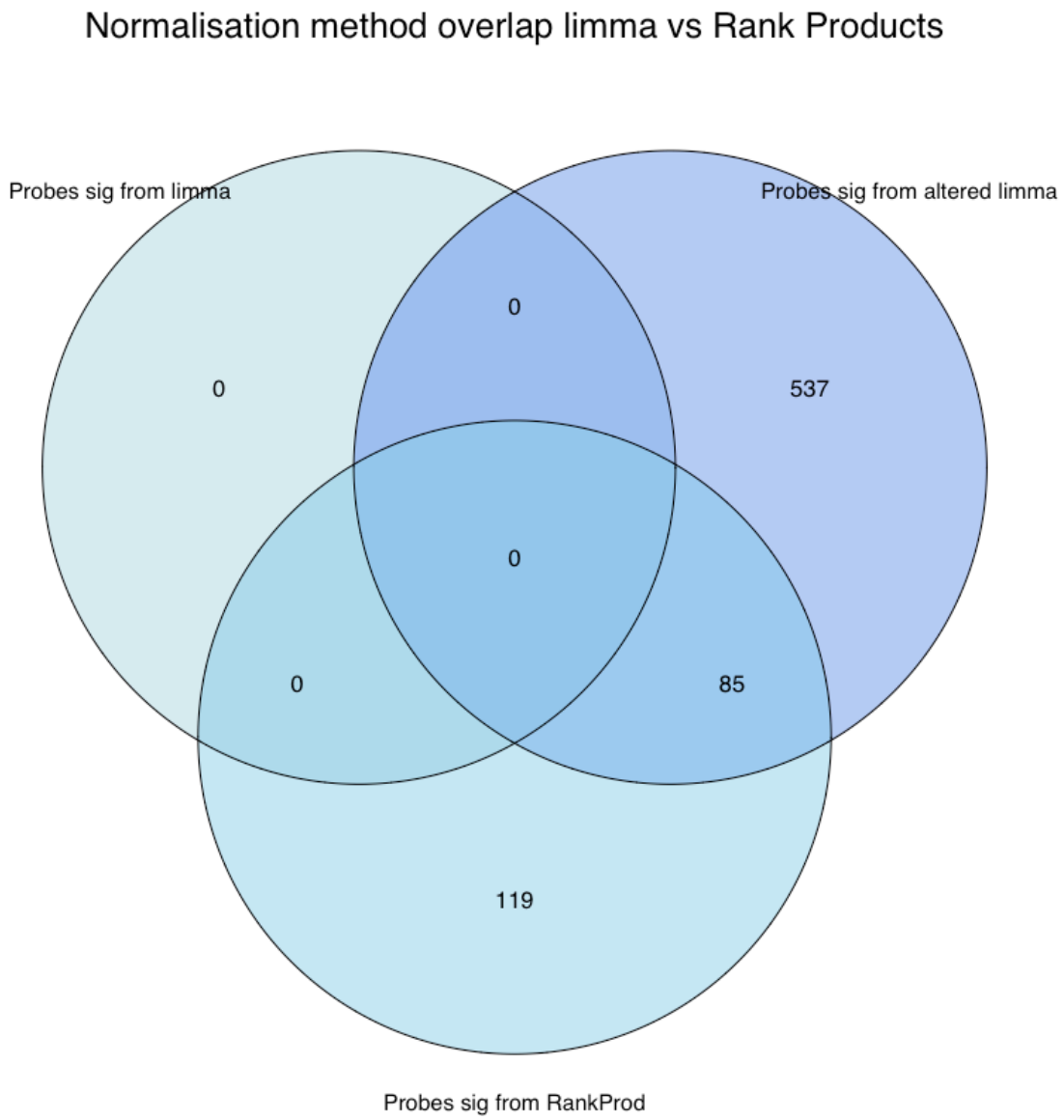


Figure 3-21 Non-parametric testing reveals differentially expressed genes in a highly variable murine dataset. Male C57BL/6 wild type and microRNA 155 knockout mice (n=3 each group) were fed on a high fat diet from 6 weeks old, livers harvested, total RNA prepared and a Mo Gene 1.0 ST array performed. Microarray files were read into R and processed using the oligo and arrayQualityMetrics packages and normalised using RMA. Differential gene expression was carried out using three methods: a standard limma pipeline using all three replicates in each group, an altered limma pipeline removing the outlier array and a rank products analysis using all three replicates in each group. The complete limma dataset reveals no differentially expressed genes after multiple testing corrections and the altered limma reveals more. Using rank products on the complete data set gives over 200 differentially expressed genes to validate, 85 of these are shared with the altered limma dataset.

array 3 removed i.e. 2 knockout arrays versus 3 wild type and also rank products using the complete data set.

The altered limma pipeline and the genes from the rank product analysis overlap to a degree and giving confidence that at least 85 genes are differentially expressed. I took both datasets forward for pathway analysis using Ingenuity. This employs a graph based method to determine how differentially expressed genes are related to each other. This unschooled clustering is combined with whether the genes are over represented in canonical curated pathways and so over time the same differentially expressed genes may be associated with newer pathways as these continue to be discovered.

I took both the rank products differentially expressed genes and the ones from the altered limma set and then carried out a Core analysis. The top canonical pathways are shown in figure 3-21 and although these have a higher proportion of genes present in the limma set, they are also present in the rank products set.

Furthermore although rank products allows us to use the complete data set and therefore not be accused of data manipulation, the test itself will cause false negatives to occur because of the inherent noise in the data set. Colleagues went on to further investigate the LXR pathway and members of the lipid and cholesterol metabolism pathways(227).

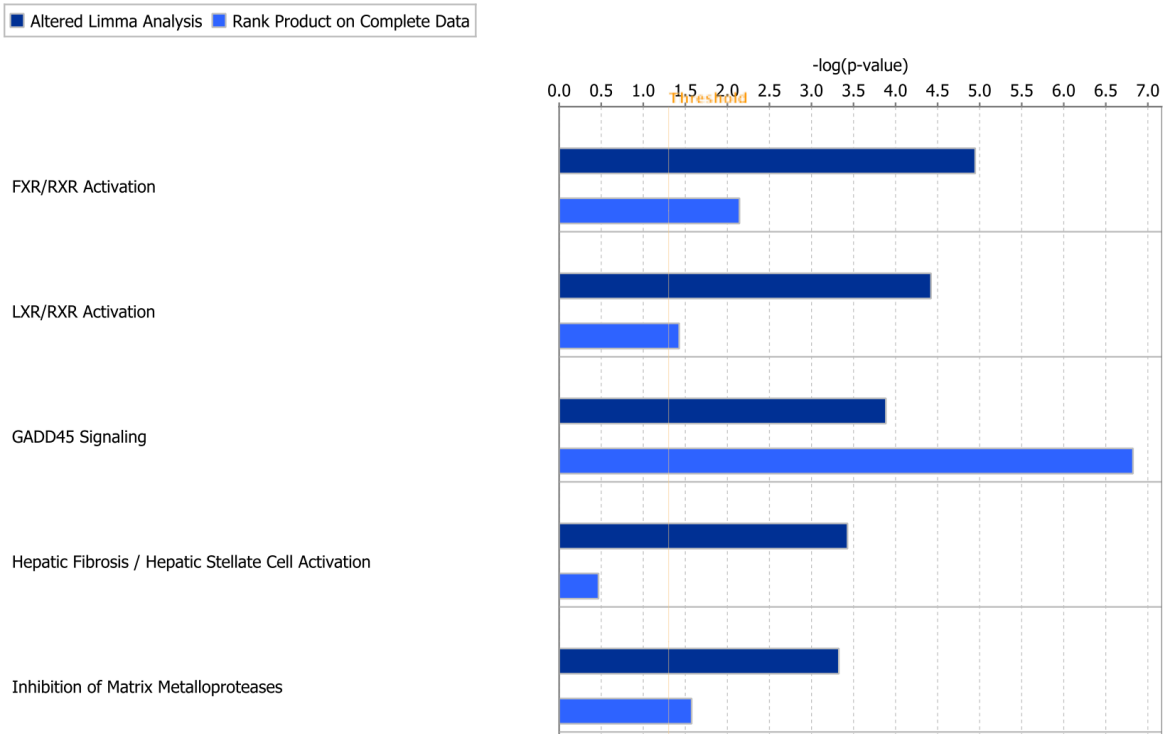


Figure 3-22 Comparison of canonical pathways that are altered between microRNA 155 knockout mice and wildtype show agreement between rank product and limma differential expression methods. C57BL/6 wild type and microRNA 155 knockout mice (n=3 each group) were fed on a high fat diet from 6 weeks old, livers harvested, total RNA prepared and a Mo Gene 1.0 ST array performed. Microarray files were read into R and processed using the oligo and arrayQualityMetrics packages and normalised using RMA. Differential gene expression was carried out using two methods: an altered limma pipeline removing the outlier array and a rank products analysis using all three replicates in each group. Pathway analysis was then carried out of genes that had a p value <0.05 in limma differential expression or a pfp value <0.05 in rank products. The length of the bar denotes the number of members of that pathway that are found in each data set and although in most cases fewer members are found in the rank products data set, the top four differentially expressed pathways are the same.

In conclusion, non-parametric tests such as rank product are useful if a dataset is small or particularly variable. However there have previously been studies where various methods of differential gene expression have been compared and limma tends to outperform other methods of differential expression once the sample size is over five(223,226,228,229).

3.3.4 Investigating the effect of various normalisation methods during microarray processing

I concluded that in the microarray data set from RA, PsA and healthy patients, a limma pipeline using the complete primary cell data set would be appropriate with the multiple testing corrected adjusted p value set at 0.05. I decided to investigate the effect of normalisation on this data using this method.

In this section, I went on to investigate the effect of using various methods of normalisation on the number of differentially expressed genes The three normalisation methods I employed are as follows: RMA, GC-RMA and FRMA. I processed the microarray data in R using the limma package and following normalisation with each method, I calculated the number of differentially expressed genes using Bayes moderated t-tests with a false discovery rate adjusted p value of < 0.05.

To investigate this I made the comparisons show in table 2. These comparisons serve two main purposes in that they investigate biology which of crucial interest to me, namely what are the differences in macrophages from two different disease states and also between macrophages and monocytes in a disease state. From the PCA plot (figure 3-15) I expect there to be a large difference in the transcriptome between monocytes and macrophages and very few consistent difference between synovial macrophages in RA and PsA. Therefore using both cell types gives me a probable yield experiment as well as a more exploratory analysis.

Comparisons
RA Synovial Macrophage vs PsA Synovial Macrophage
PsA Synovial Macrophage vs PsA Peripheral Blood Monocyte
RA Synovial Macrophage vs RA Peripheral Blood Monocyte

Table 3-2 Microarray comparisons to be made to investigate the effect of normalisation on differential gene calculation

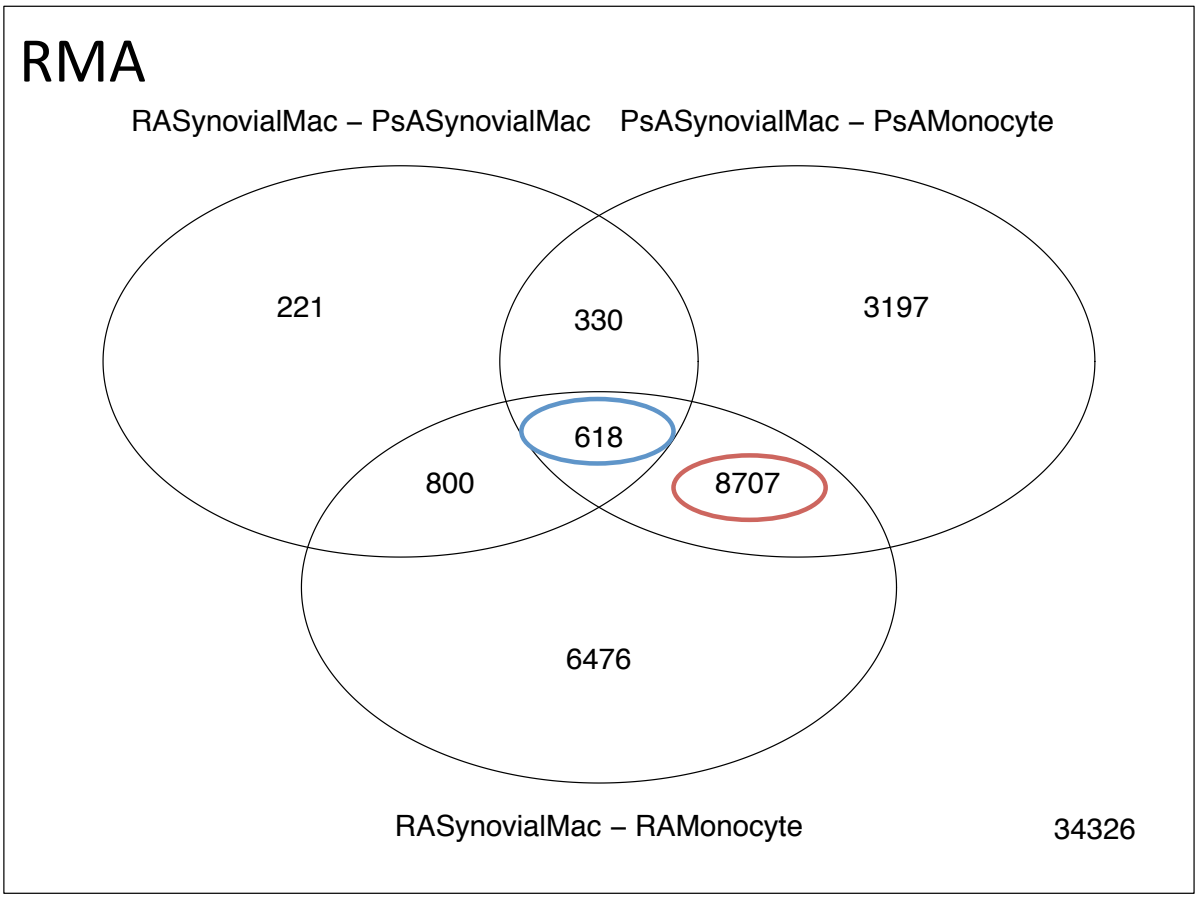


Figure 3-23 Differential gene expressions using RMA normalisation reveals more genes differences in comparisons between monocytes and macrophages. Microarray files were read into R and processed using RMA normalisation and differential genes calculated using Bayes moderated t-tests from the limma package with an fdr corrected p value of <0.05. Venn diagrams were generated using the decideTests function within limma. The number in the bottom right shows the total number of genes from this experiment analysis. A small number of genes are found in common between the three analyses although the most differentially expressed genes are found in the macrophage and monocyte comparisons. Although this Venn diagram shows overlap of genes, the direction of change may be different.

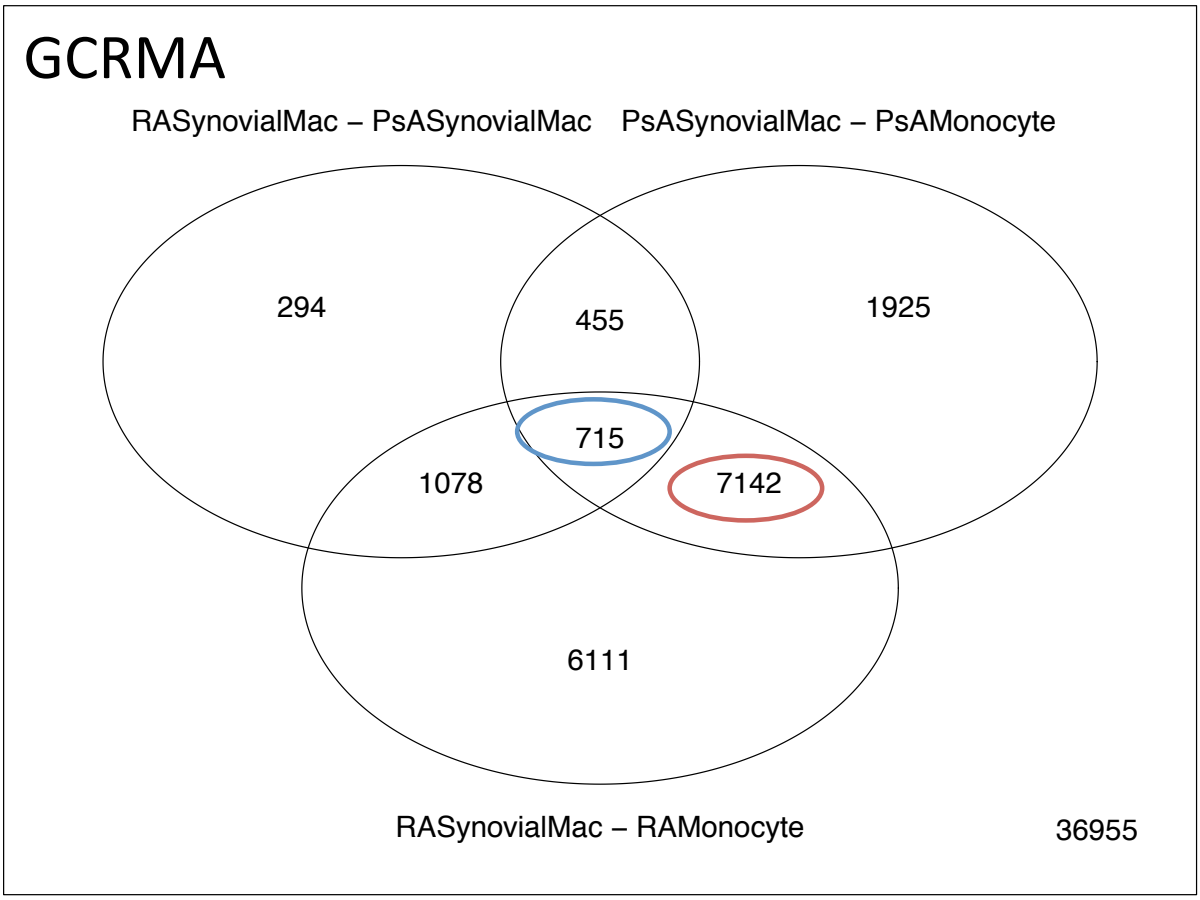


Figure 3-24 Differential gene expressions using GC-RMA normalisation reveals more genes differences in comparisons between monocytes and macrophages. Microarray files were read into R and processed using GC-RMA normalisation and differential genes calculated using Bayes moderated t-tests from the limma package with an fdr corrected p value of <0.05. Venn diagrams were generated using the decideTests function within limma. The number in the bottom right shows the total number of genes from this experiment analysis. A small number of genes are found in common between the three analyses although the most differentially expressed genes are found in the macrophage and monocyte comparisons. Although this Venn diagram shows overlap of genes, the direction of change may be different.

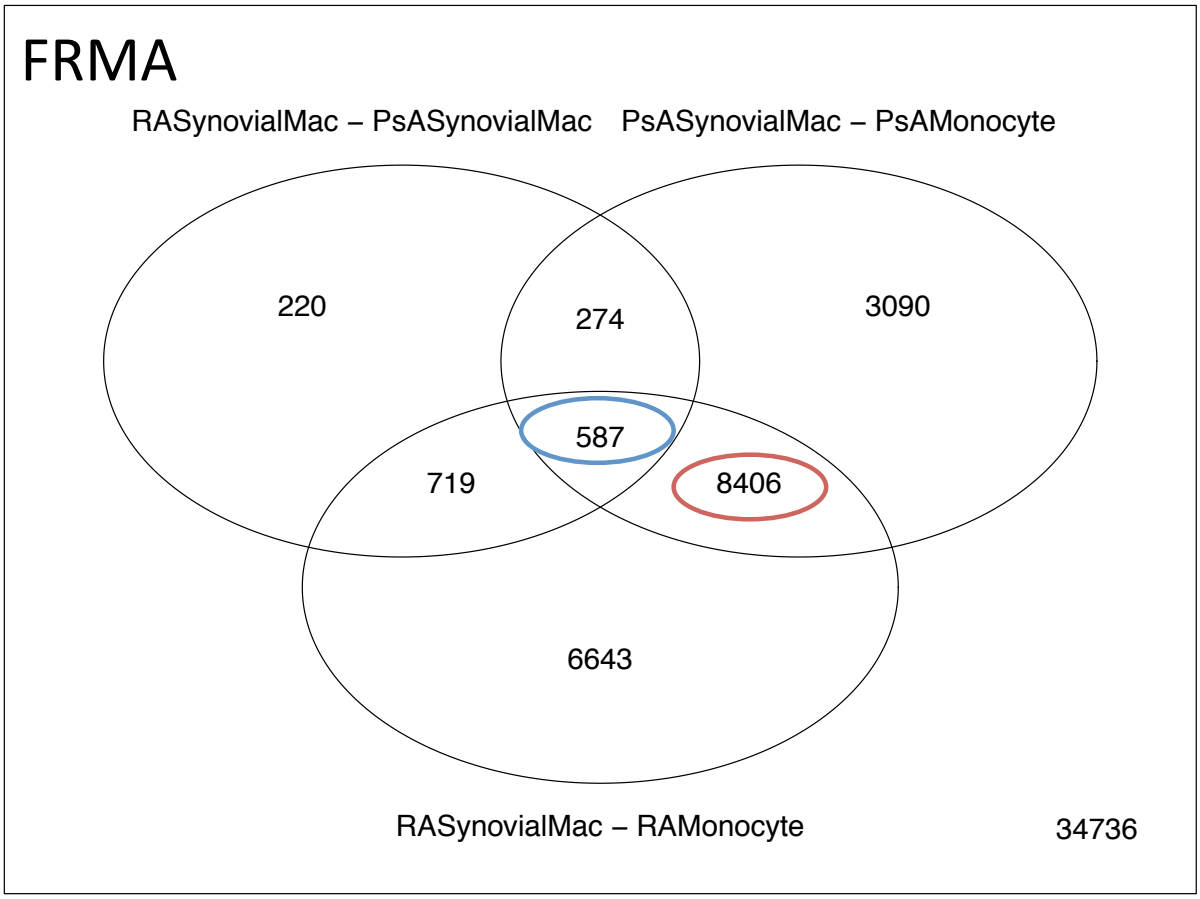


Figure 3-25 Differential gene expressions using FRMA normalisation reveals more genes differences in comparisons between monocytes and macrophages. Microarray files were read into R and processed using FRMA normalisation and differential genes calculated using Bayes moderated t-tests from the limma package with an fdr corrected p value of <0.05. Venn diagrams were generated using the decideTests function within limma. The number in the bottom right shows the total number of genes from this experiment analysis. A small number of genes are found in common between the three analyses although the most differentially expressed genes are found in the macrophage and monocyte comparisons. Although this Venn diagram shows overlap of genes, the direction of change may be different.

It is evident that each method of normalisation results in a different number of differentially expressed genes (figure 3-22 to 3-24). However the numbers of differentially expressed genes tend to be similar in each of the overlaps although there are fewer found with GCRMA in the Psoriatic Arthritis comparisons.

Furthermore it is clear that comparing RA and PsA synovial macrophages has resulted in a large number of differentially expressed genes which was not in keeping with my expectation due to the overlap of these samples on PCA (figure 3-15). Briefly, these cells were isolated from the synovial fluid of patients undergoing knee aspiration by using magnetic bead positive selection for CD14. Therefore we might expect these cells to be similar in terms of their transcriptomic profile as they both came from an inflamed joint. However, approximately 2000 genes are differentially expressed when we compare RA and PsA synovial macrophages and this may represent true biology and is a reflection of the different pathogenesis in each disease.

Also when macrophages are compared to monocytes within either RA or PsA, a large number of differentially expressed genes are found in common. This is in keeping with the hypothesis that the main effect would be due to the difference between the monocyte and macrophage transcriptome. Certainly out of the approximately 13000-14000 differentially expressed genes, 8-9000 of these are shared between RA and PsA suggesting that these are commonly changed. The Venn diagram does not show directionality of change and therefore it is possible that these genes could be changed in opposite directions. This could be further interrogated using heatmaps or pathway analysis.

Prior to drawing conclusions, I went on to determine the effect of different methods of normalisation within each comparison. To do this I had to process the microarray data three times using a different method of normalisation and then draw a Venn diagram of the overlap within R (figure 3-25 and 3-26).

3.3.5 RMA, GCRMA and FRMA give a significant overlap of probesets when comparing the RA synovial macrophage to the PsA synovial macrophage and respective disease monocytes

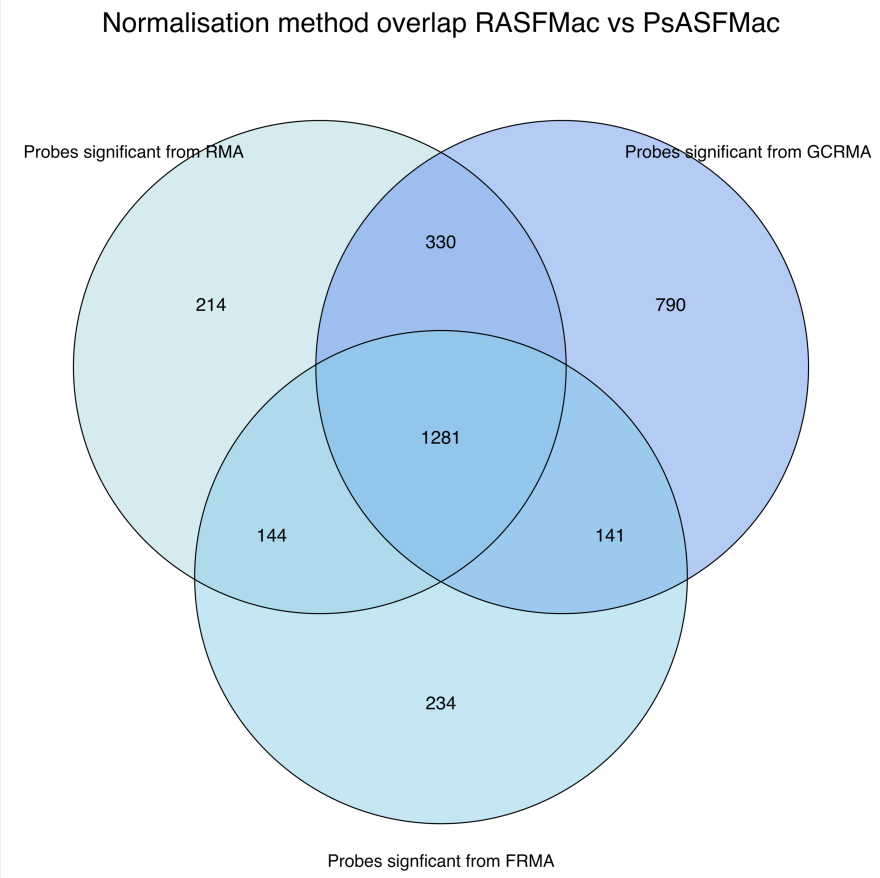


Figure 3-26: Methods of normalisation show significant overlap in the number of differentially expressed genes from RA versus PsA synovial macrophages. Microarray files were read into R and processed using three different normalisation methods to generate separate data sets. Differential gene expression was calculated using limma and a Bayes moderated t-test with fdr multiple testing corrected adjusted p value <0.05 for each data set and the overlap determined using the VennDiagram R package. Over 1200 of the approximately 2000 differentially expressed genes are discovered with any of the described methods of normalisation. However each method does also generate a number of genes that are unique to that method.

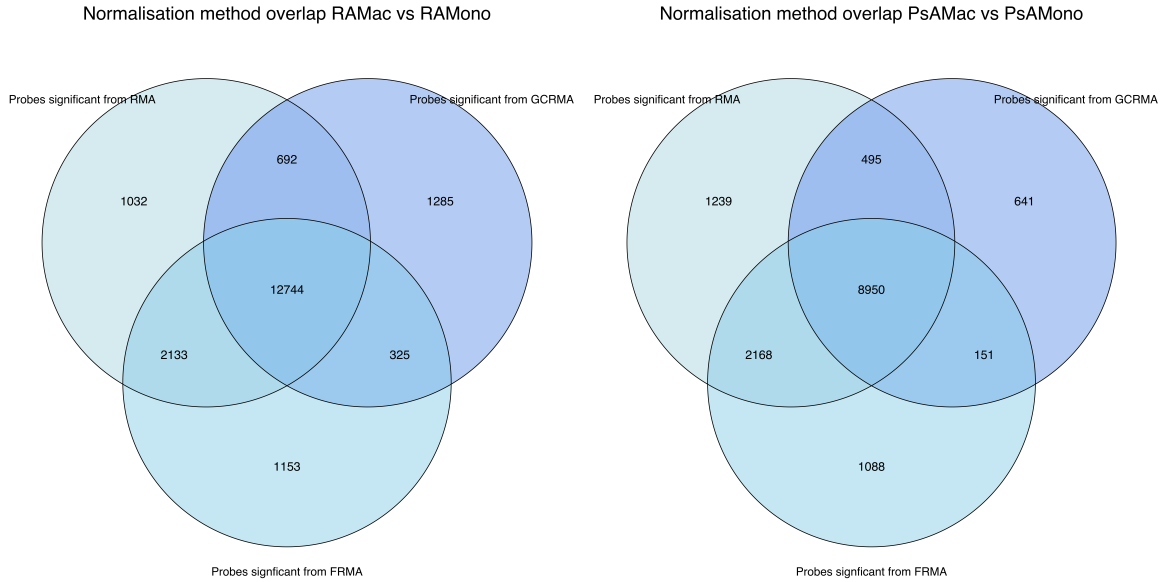


Figure 3-27: Methods of normalisation show significant overlap in the number of differentially expressed genes between monocytes and macrophages within diseases. Microarray files were read into R and processed using three different normalisation methods to generate separate data sets. Differential gene expression was calculated using limma and a Bayes moderated t-test with fdr multiple testing corrected adjusted p value <0.05 for each data set and the overlap determined using the VennDiagram R package. There are more differentially expressed genes found in RA macrophages versus monocytes compared to the corresponding analysis in PsA. The methods of normalisation generally agree with each other although there is more concordance between RMA and FRMA.

Two conclusions can be made from these figures: firstly there are far more differentially expressed probes when disease macrophages are compared to monocytes versus comparing the macrophages to each other and secondly that in all three cases, there is a large overlap between the methods of normalisation.

This could be expected given the techniques are built on RMA but is certainly encouraging that it is not grossly different. In fact it appears that RMA offers a good compromise and certainly there has been evidence that GCRMA can overestimate the number of differentially expressed probes and therefore lead to more false positive results(230).

FRMA would offer another option but in this case the QC plots in terms of the signal distribution, correlation plot or PCA do not provide evidence of a batch effect. Furthermore the experimental design was such that the same researcher collected samples and processed in one batch for RNA extraction and chip wet lab work although cell isolation was carried out as samples were received (Lucy Ballentine personal communication). Therefore FRMA in this instance may provide too stringent a framework for differential expression analysis.

Therefore I proceeded to analyse two of the datasets further: The RA synovial macrophage versus the PsA synovial macrophage to determine which pathways were changed in these diseases and also whether a JAK/STAT signature is evident between disease states and the RA synovial macrophage versus RA peripheral blood monocyte. From the investigations so far, I decided to use RMA normalisation with no batch correction and I employed a Bayes moderated t-test from the limma package to calculate differentially expressed genes.

3.4 Over 1300 genes are differentially expressed between RA and PsA synovial CD14 cells revealing evidence of differential JAK/STAT utilization between the diseases

Figure 3-28: Canonical pathways from RA and PsA synovial macrophages show IL-6 and acute phase response signaling upregulated in RA. Microarray files were read into R and processed using RMA normalisation and differentially expressed genes calculated using Bayes moderated t-tests in limma with an fdr adjusted p value <0.05. The gene list was then analysed using an Ingenuity Core Analysis to give these differentially expressed canonical pathways. Orange demonstrates relative activation in RA synovial macrophages compared with PsA and blue demonstrates relative inhibition in RA synovial macrophages compared to PsA

I therefore performed differential gene analysis, using limma and Bayes moderated t-tests, of RA versus PsA synovial macrophage samples. This resulted in over 1300 genes determined as differentially expressed and therefore to provide insight into this data I carried out a Core Analysis in Ingenuity Pathway Analysis. Figure 3-27 shows that certain canonical pathways are over represented in RA (orange) and PsA (blue).

Furthermore, IL-6 signaling features high as a pathway up regulated in RA synovial macrophages versus PsA synovial macrophages. Figure 3-28 demonstrates this where red molecules are genes that are upregulated in RA, green are upregulated in PsA and grey means that there is no differential expression. The colour intensity gives an indication of magnitude of fold change.

Clinically this is relevant because blocking IL-6R using tocilizumab is an effective treatment for RA but is not effective for PsA suggesting that although both diseases result in synovitis, the mechanisms responsible for disease pathogenesis must be different because of the differential response to therapy and also differential activation of inflammatory pathways in this array(12,54,231).

IL-6 signals through JAK2, STAT3 and also PI3K and Akt and we see that JAK2 expression levels are lower in RA synovial macrophages compared to PsA and also IL6R is highly upregulated in RA synovial macrophages compared to PsA. Furthermore overlaying the differentially expressed gene list over the canonical STAT3 pathway in Ingenuity highlights significant differences in expression levels of growth factor receptors FGFR1 and FLT1 or VEGFR-1 are highly expressed in RA compared to PsA (figure 3-29).

My analysis showed that PI3K/Akt kinases are upregulated while their inhibitor INPP5D is down regulated in RA synovial macrophages. This validates the key role of this pathways and it's epigenetic regulators in the pro-inflammatory activation of RA synovial macrophages that has previously been shown in our lab(232).

This gives credence to the hypothesis that there is differential JAK/STAT utilisation in RA and PsA. To investigate this further, I used the grow and connect functions in Ingenuity to construct my own JAK/STAT network starting with all JAK and STAT members and creating a network which ranged from growth factor receptors to nuclear transcription factors. I already saw that JAK2 levels were different between the conditions and

therefore this approach allowed me to explore whether this was evident in other JAK and STAT members and associated pathways (figure 3-30).

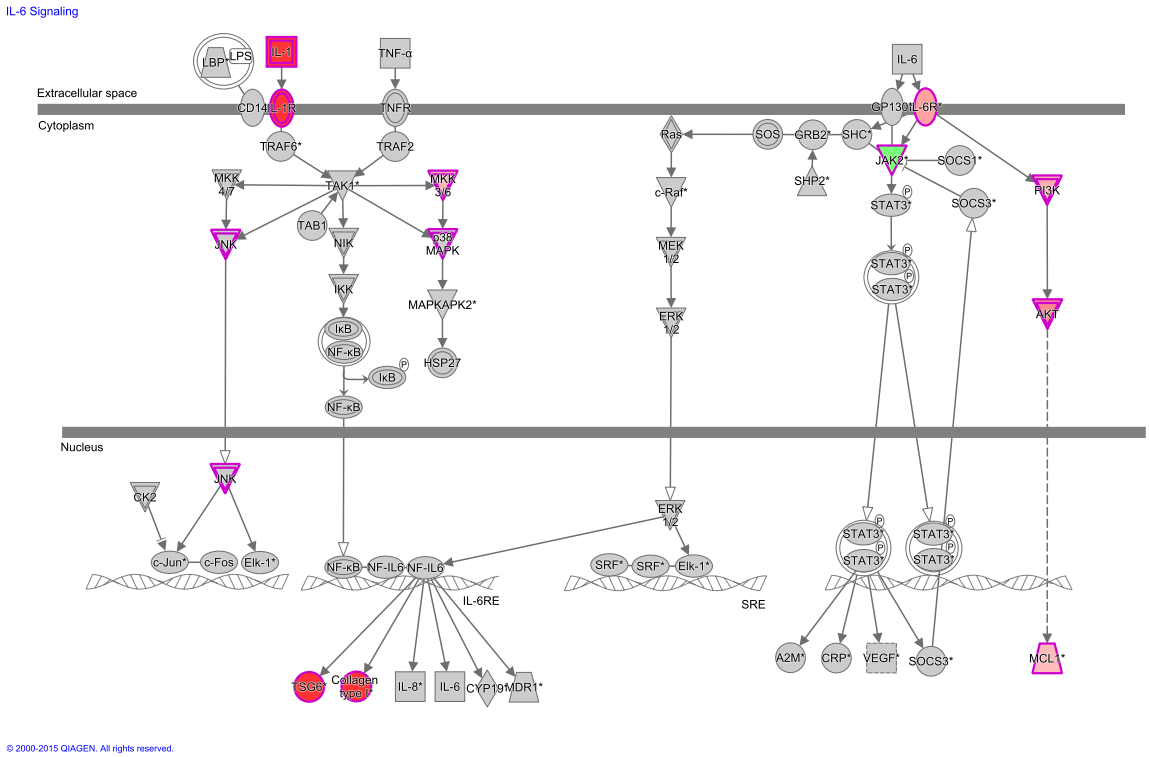


Figure 3-29: IL-6 signaling pathway showing that JAK2 is lower in RA synovial macrophages whereas IL-6R is upregulated. Microarray files were read into R and processed using RMA normalisation and differentially expressed genes calculated using Bayes moderated t-tests in limma with an fdr adjusted p value <0.05. The gene list was then analysed using an Ingenuity Core Analysis to give these differentially expressed canonical pathways. The IL-6 signaling pathways demonstrates genes in red as relatively upregulated in RA synovial macrophages compared to PsA and those in green as relatively upregulated in PsA synovial macrophages.

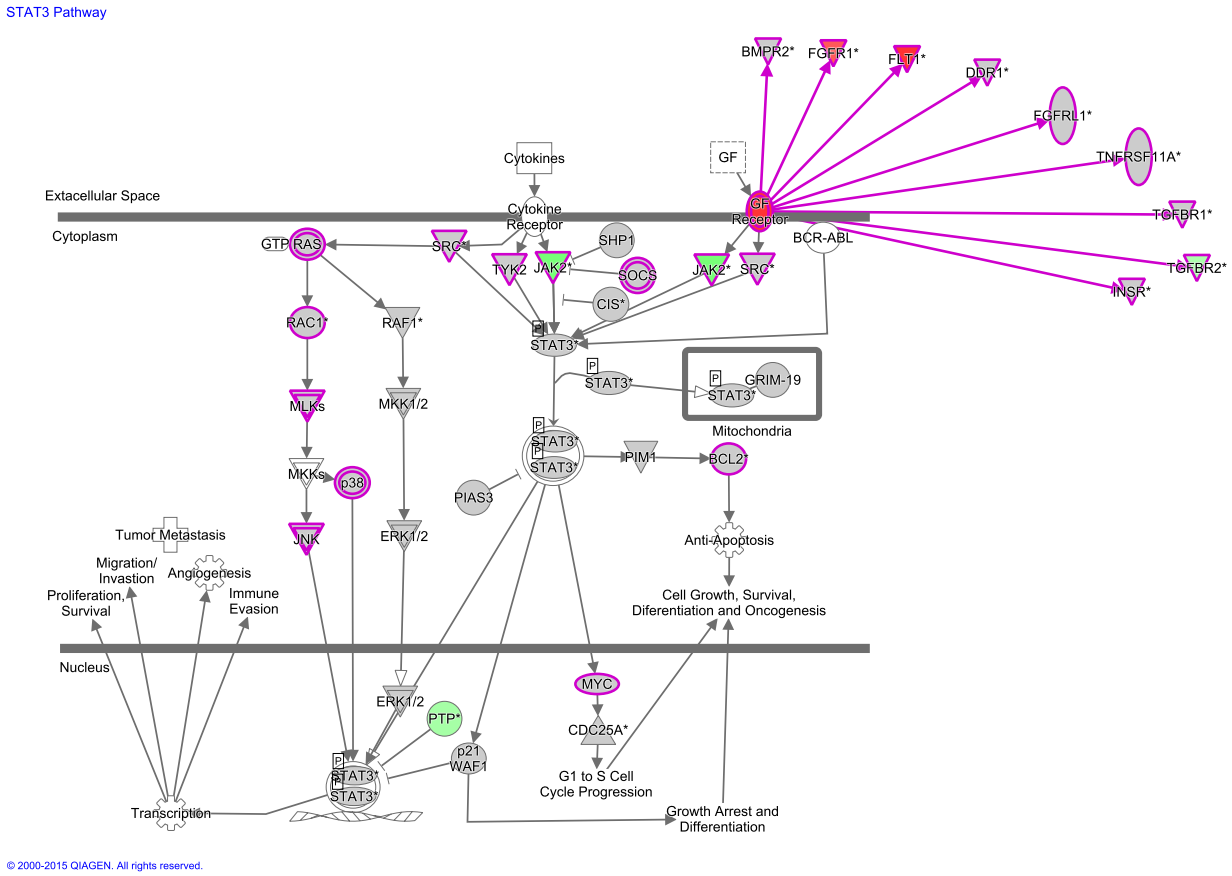


Figure 3-30: STAT3 signaling pathway shows that growth factor receptors that sit upstream of JAK2 and STAT3 are upregulated in RA synovial macrophages. Microarray files were read into R and processed using RMA normalisation and differentially expressed genes calculated using Bayes moderated t-tests in limma with an fdr adjusted p value <0.05. The gene list was then analysed using an Ingenuity Core Analysis to give these differentially expressed canonical pathways. The STAT3 signaling pathways demonstrates genes in red as relatively upregulated in RA synovial macrophages compared to PsA and those in green as relatively upregulated in PsA synovial macrophages.

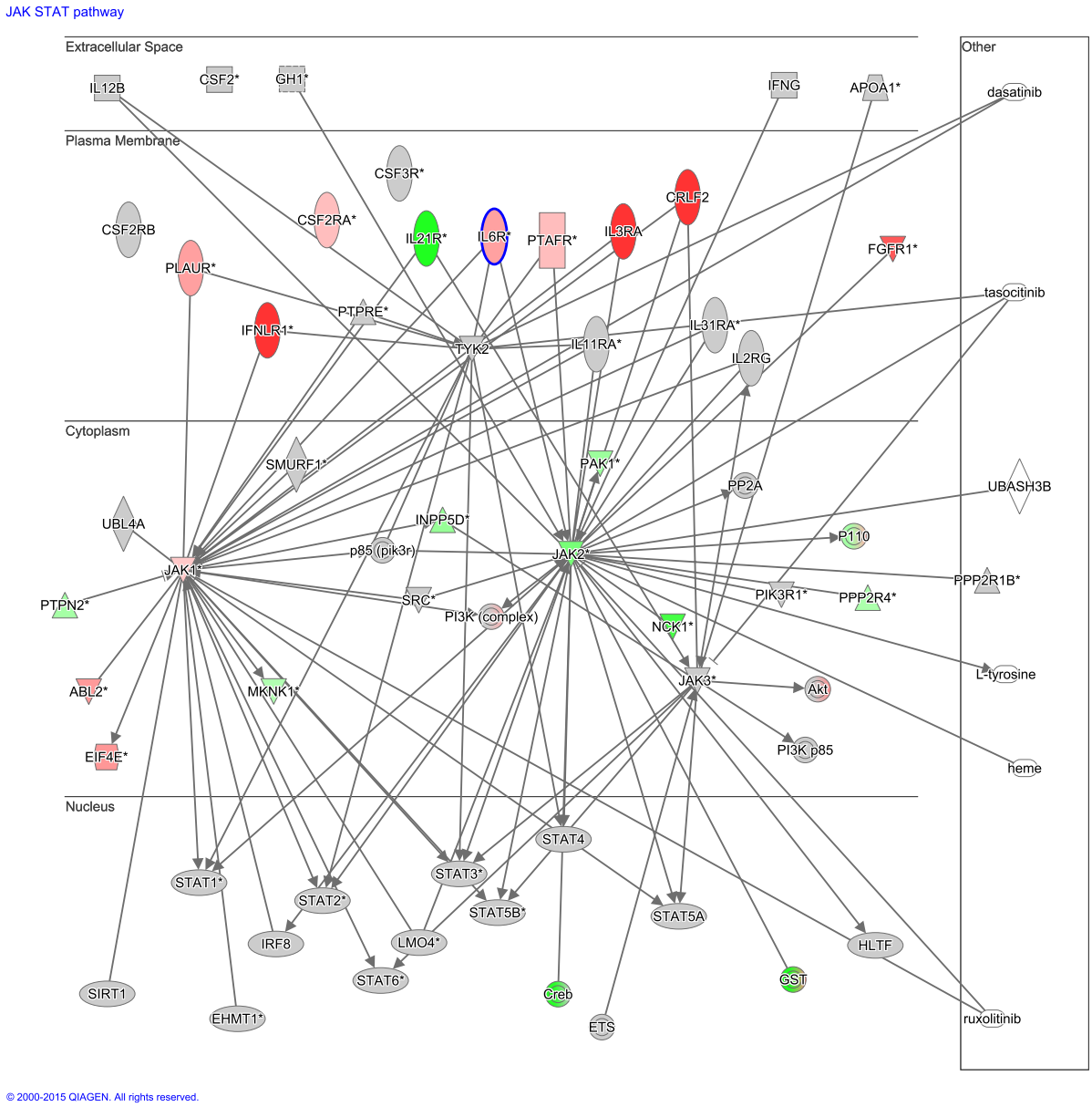


Figure 3-31 JAK1, JAK2 and associated molecules are differentially expressed between RA synovial macrophages and PsA synovial macrophages. Microarray files were read into R and processed using RMA normalisation and differentially expressed genes calculated using Bayes moderated t-tests in limma with an fdr adjusted p value <0.05. A network was then constructed using all JAK and STAT members and grown both up and downstream. Genes that were differentially expressed between RA synovial macrophages and PsA synovial macrophages were overlaid on this network. Genes in red are relatively upregulated in RA synovial macrophages compared to PsA and those in green as relatively upregulated in PsA synovial macrophages. JAK1, IL6R and FGFR1 levels are higher in RA with JAK2 and IL21R higher in PsA synovial macrophages.

Although JAK3 is not changed, JAK1 is upregulated in RA synovial macrophages and JAK2 is higher in PsA synovial macrophages. Furthermore other changes which we would expect to see include CSF2RA (233)and IL6R upregulated in RA and IL21R upregulated in PsA(234). Upregulation of IL21R and TGFBR2 in PsA macrophages compared to RA might represent the presences of a remodeling process in PsA but not RA joints. Both IL-21 and TGFβ have been described as drivers an M2 macrophage repair phenotype.

We can further explore the differences between the two inflammatory arthropathies by examining two further data sets: the RA synovial macrophage compared to blood monocyte and also the PsA synovial macrophage compare to blood monocyte. The reason for this would be to confirm finding that we have observed when comparing the synovial macrophages between disease states and dissect if the system of local environment drives the disease phenotype.

Figure 3-31 shows that JAK2 is altered in RA synovial macrophages and blood monocytes and there is upregulation of IFNLR1 and FGFR1. There is no change in IL6R transcript suggesting that IL6RT is upregulated in the periphery of RA patients.

When the PsA data set is overlaid (figure3-32), a different pattern is seen in that although IFNLR1, FGFR1 and TYK2 feature, IL21R and NCK1 are differentially expressed in keeping with the findings from the RA and PsA synovial macrophage analysis. Therefore I conclude that there is evidence of a JAK/STAT signature in RA (IL6R) is different from PsA (IL21R). Therefore exploring inhibition of the JAK/STAT pathway using small molecule inhibitors in a macrophage and cytokine activated T cell cell contact assay would make a logical next step in my doctoral studies.

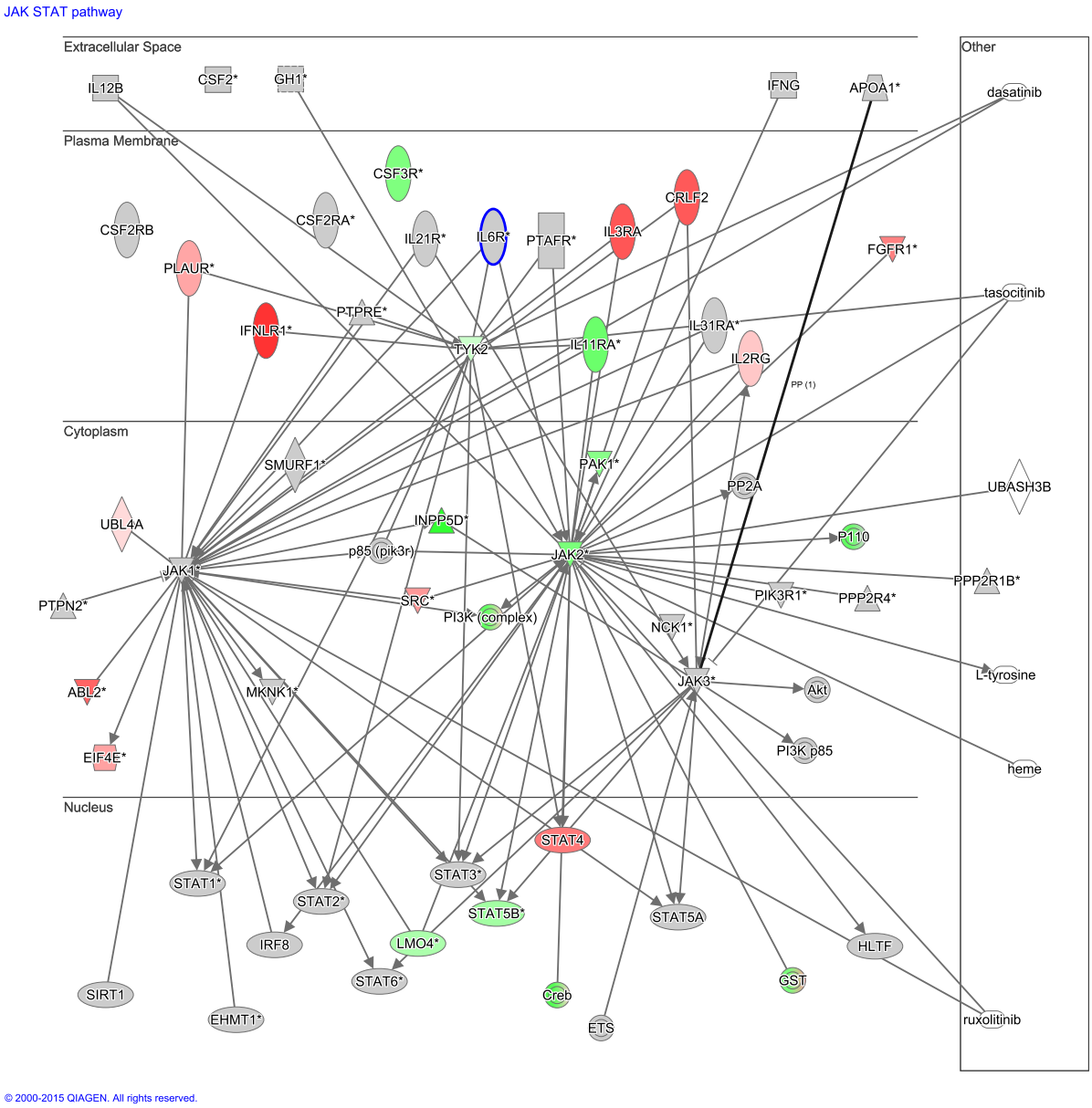


Figure 3-32 JAK2, IFNLR1 and FGFR1 are upregulated in RA synovial macrophages compared to blood monocytes. Microarray files were read into R and processed using RMA normalisation and differentially expressed genes calculated using Bayes moderated t-tests in limma with an fdr adjusted p value <0.05. A network was then constructed using all JAK and STAT members and grown both up and downstream. Genes that were differentially expressed between RA synovial macrophages and RA blood monocytes were overlaid on this network. Genes in red are relatively upregulated in RA synovial macrophages compared to RA monocytes and those in green as relatively upregulated in RA blood monocytes. JAK2, IFNLR1 and FGFR1 are upregulated in the RA synovial macrophage compared to blood monocyte. IL6R is not altered between the two suggesting that levels are similar although a strong difference was seen comparing the RA and PsA synovial macrophage.

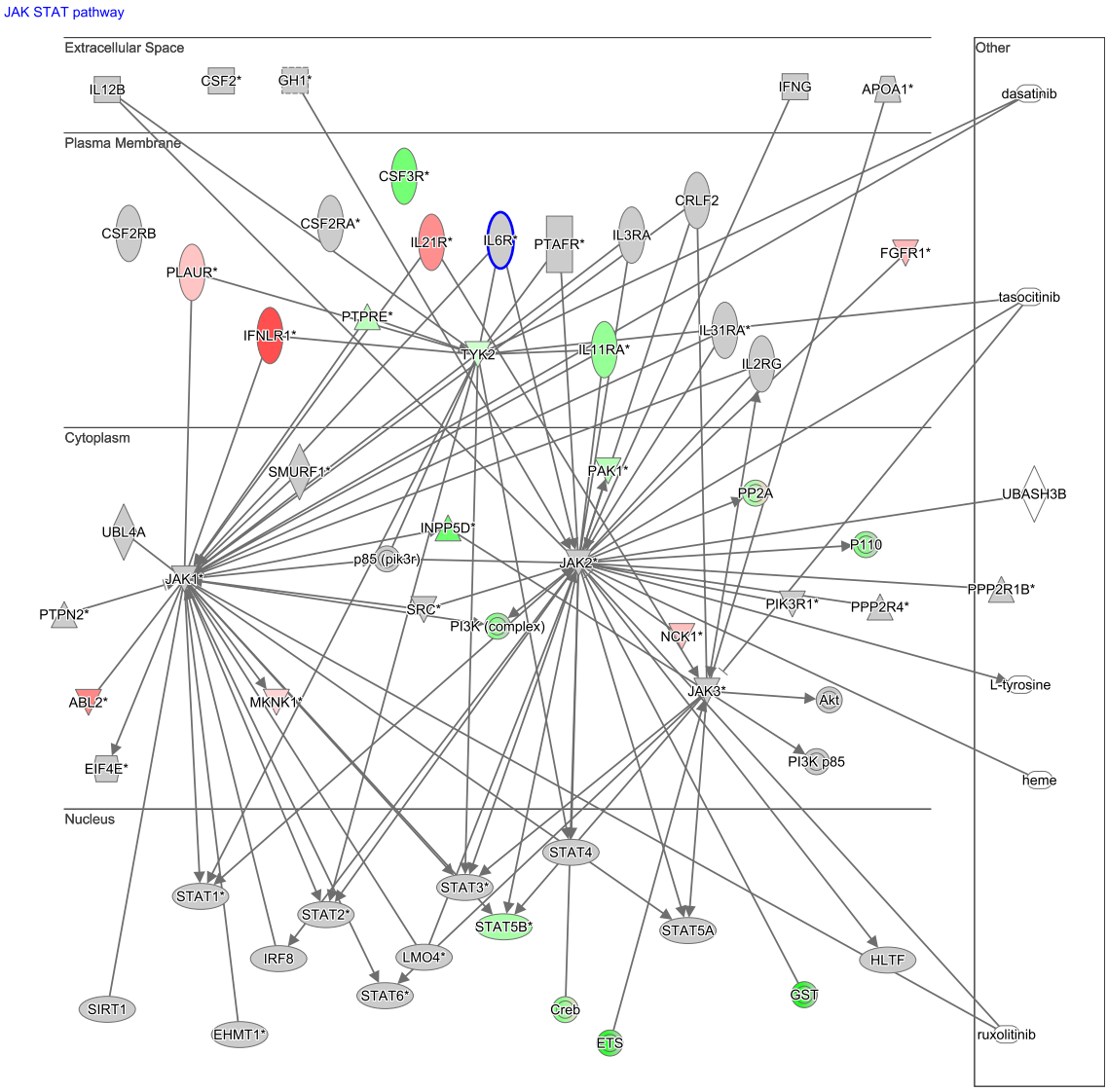


Figure 3-33 TYK2, IFNLR1 and IL21R are differentially expressed in PsA synovial macrophages compared to blood monocytes. Microarray files were read into R and processed using RMA normalisation and differentially expressed genes calculated using Bayes moderated t-tests in limma with an fdr adjusted p value <0.05. A network was then constructed using all JAK and STAT members and grown both up and downstream. Genes that were differentially expressed between PsA synovial macrophages and PsA blood monocytes were overlaid on this network. Genes in red are relatively upregulated in PsA synovial macrophages compared to PsA monocytes and those in green as relatively upregulated in PsA blood monocytes. IL21R, IFNLR1 and FGFR1 are upregulated in the PsA synovial macrophage compared to blood monocyte. IFNLR1 and FGFR1 are also altered in the corresponding RA comparison suggesting that these genes may be regulating monocyte to macrophage development. Furthermore, TYK2 is altered in PsA that is in keeping with GWAS studies.

3.5 Discussion

In this chapter I set out to:

1. Investigate how microarray dataset quality could be determined and demonstrate the influence of this on downstream analysis

2. Develop an analysis pipeline for microarrays and investigate whether batch effects, methods of normalisation and differential expression statistics affect analysis outcome

3. Using the outcome of the above objectives, determine whether a JAK/STAT signature is present in myeloid lineage cells from patients with RA and if this differs from PsA

In summary I used R and associated packages to perform quality control measures on a monocyte and macrophage dataset that had previously been generated in our lab. I determined that the data quality of MCSF and cell contact activated macrophage samples was insufficient for further analysis. I compared methods of normalisation and differential gene expression and concluded that in this case, RMA normalisation and Bayes moderated t-tests would be appropriate to process this data set.

Having decided the best way to process this dataset, I went on to interrogate the transcriptome of synovial macrophages from RA and PsA. Following pathway analysis and network creation I conclude that there is a JAK/STAT signature present in synovial macrophages with IL6R upregulated in RA and IL21R upregulated in PsA

The use of DNA microarrays in translational research in rheumatology has grown from 2004 and this technology is still being used today despite the availability and also relative ease of RNA sequencing. Researchers have become familiar with particular manufacturers such as Affymetrix and Illumina, their wet lab and dry lab preparation and processing protocol and there is inherent trust in the results from these platforms. Furthermore, the use of blackbox data analysis packages such as Partek Genome Studio and Genespring mean that researchers can access information that is useful to them in a quick and easy fashion. However, misinterpretation or not analysing quality control plots can mean that there are no differentially expressed genes from an experiment or there is a high proportion

of false results(235). Analysis of data from public repositories has revealed unreported quality control issues that affect the subsequent interpretation of the data. Awareness of such issues enables removal of suspect data sets whose inclusion could bias the downstream interpretation by building poor hierarchical models.

Unfortunately there is no way to take account of a random technical issue and therefore in that instance such as overhybridisation in cell contact 22 it is better to remove the data and not include it in a limma analysis. The reasoning for this is that limma builds a hierarchical model that borrows data from all other genes and arrays in order to perform differential expression analysis. If the data going into the model is poor, it will underperform.

Removing the data is important because subsequent normalisation for the entire array data set will take account of some information in the poor array. Furthermore the cell contact activated macrophage and MCSF macrophage data did not help in achieving my experimental objectives and therefore were removed from subsequent analysis.

Batch effects occur in array and other omics datasets and therefore it is crucial that as much data as possible is collected about the wet lab process because any one of these steps could contribute to the batch effect. If the batch effect is systematic ie a chip effect or in the case of the RA and SLE dataset that was presented, a time effect, this can be corrected by modeling the batch effect. Methods to correct this involve computationally removing the "cause" of the batch effect either by including that factor as part of the differential expression model or using a package to remove it. One useful method is the ComBat method that is implemented in the Bioconductor sva package. This package models the batch effect and then aims to remove it by using an empirical Bayes framework similar to that used in the linear models for microarray data package, limma. Therefore although batch effects were not an issue in my microarray dataset, I have the knowledge now to assess them and implement strategies to take account and correct for them if present.

There are many methods of differential gene expression in microarray experiments ranging from calculating fold change, simple t-tests, significance analysis of microarray experiments, Bayes moderated t-tests as part of the limma package and also non parametric methods such as Rank Products. I decided to investigate which statistical test was appropriate prior to assessing whether methods of normalisation had an effect on differential gene expression as a pragmatic approach. Therefore, I controlled the method

of normalisation while varying the method of differential gene expression. The limma method of normalisation is widely used for various array types and also RNA sequencing data. It has been shown to outperform other methods of differential expression and is generally robust when there are more than 5 replicates in a microarray experiment and so in our data set with a minimum of 5 biological replicates of healthy blood monocytes and 8 replicates of paired patient synovial and blood samples, it is likely that limma will be the best method(216,223).

However, to address this question I employed another data set from our lab that had a high degree of variation possibly due to an element of technical error. The quality control plots of wild type versus microRNA 155 knockout mice clearly showed that a knockout sample was clustering with the wild type samples on both raw and normalised data analysis. Furthermore, a standard limma analysis pipeline revealed no differentially expressed genes when an fdr adjusted p value of <0.05 was employed. Therefore, in this situation one can discard the array that is failing quality control or employ a non-parametric statistic that is suited to smaller and variable data sets. I decided to compare both of these approaches in this data set.

In the case of the microRNA 155 knockout versus wild type arrays we see that the Rank product methods outperforms limma when used on the whole dataset and therefore would give the researcher an indication of which pathways or genes to assess further. The caveat in this situation is that with a noisy dataset and poor quality data, the chance of a false positive result are much higher and therefore validation with other technologies such as quantitative real-time PCR or quantifying protein levels of a particular gene product would be crucial. In this particular project, the researchers validated the differentially expressed genes that my analysis generated with qRT-PCR of additional samples.

Therefore in the case of small numbers, non-parametric methods and even fold change can be helpful to analyse high throughput data. This can be useful if an experiment has to be carried out on a degraded sample because it is particularly rare. If larger numbers of replicates are available, methods such as limma outperform because they allow the modeling and removal of systematic batch effect and also enable the researcher to implement sophisticated experimental designs such as paired or time course sample analysis.

Microarray experiments need to be normalised to take account of technical differences that occur during the experimental process in order to allow researchers the best chance to discover a biological signal. The nature of microarray and sequencing technology means that the distribution of data tends to be positively skewed and therefore a log transformation partly normalises the data. This tends to be the case for many of the 'omics technologies and therefore assessing the distribution of proteins or metabolites from high throughput experiments for positive skew and then log transforming means that many of the parametric methods which are used for differential expression from microarrays can be deployed onto other technologies.

Robust multiarray normalisation has been shown to be a reliable method of normalisation for microarray data especially when batch effects are not an issue. When I compared the different methods of RMA normalisation in the RA synovial macrophages versus the PsA synovial macrophages, there was a good overlap of differentially expressed genes between the methods. However, GC-RMA normalisation produced almost 50% more differentially expressed genes and there is literature suggesting that this method over estimates the number of differentially expressed genes(230). There was good overlap between the RMA and frozen RMA methods although each method generated a small number of genes unique to that particular process.

I decided to use the RMA method of normalisation as part of my pipeline because I had not discovered any batch effects that would make me want to use the frozen RMA method and it is widely accepted in the literature. In reality, confirming that a particular method of normalisation is better than another is difficult and although it can be assessed with density histograms to assess the distribution, there is still a degree of subjectivity involved. The only robust way to determine a superior method of normalisation would be to confirm using RT-qPCR whether the genes that are differentially expressed are validated.

Therefore to determine whether a JAK/STAT signature is evident in myeloid lineage cells from RA patients I decided to use R to analyse the data using RMA to normalise and implement Bayes moderated t-tests as a method of differential expression. An example of the programme script that I generated is shown in the appendix.

Initial microarray studies in rheumatoid arthritis tended to use proprietary platforms and also focused on synovial tissue obtained from biopsy(236). Although this approach is being refined, the conclusion of these studies was often that RA was a heterogeneous

disease and did not take account of the fact that the "transcriptome" in these samples consisted of a mixture of cells: fibroblasts, innate immune cells and also elements of lymphoid aggregates which vary from patient to patient.

Microarray studies then focused on those where a drug intervention such as an anti-TNFα medication was employed as part of a clinical trial and therefore a patient sample became an internal control but was still dependent of the cellular makeup of the particular biopsy(237). At the same time, researchers started carrying out studies of primary cells or cultured cells in particular conditions and work focused on fibroblasts isolated from synovial membranes of patients undergoing arthroplasty. This approach resulted in homogenous populations of cells conditions such as osteoarthritis and RA could be compared(238). In addition, cultured cells were subjected to cytokine stimulation or blockade thereby improving the experimental design and also the reliability of the experimental results(239).

As well as profiling cultured cells, interest began to grow in the ability to profile the transcriptome of human whole blood(240). This approach led to further issues of an excess of globin transcript levels from red blood cells but paved the way for a whole host of studies involving the whole blood transcriptome many of which have been and are still being carried out in our lab (James Dale, personal communication).

To remove the effect of the red blood cell, many studies were also carried out on human PBMC fractions. However, alteration in the relative makeup of the PBMC fraction by improvement in systemic inflammation could lead to a change in transcripts that was only as a result of a change in proportion of a particular cell type(241).

One publicly available study comparing human synovial macrophages from patients with RA involved 5 donors with RA and 3 healthy donors from which CD14+ monocytes were selected by positive selection from PBMCs(242). These were then cultured with MCSF and compared to the RA synovial CD14+ cells. These samples led to the hypothesis that TNFα treatment of primary macrophages from patients led to an increased expression of interferon response genes, a finding that others have replicated in conditions such as SLE, JIA and also tuberculosis infection.

An elegant study from You et al (243) showed how a systems biology approach can be used to compare two cell types: namely the synovial fibroblast and synovial macrophage.

They showed that factors involved with fibroblast were upregulated in RA and accentuated by RA synovial macrophages that were typified by a pro-inflammatory phenotype.

Therefore the microarray study carried out in our lab added to this information because of the paired monocyte and macrophage samples from patients and also the ability to compare synovial macrophages between RA and PsA. I used this data to devise a quality control and data analysis pipeline using open source bioinformatics software which would provide me with transferable skills across the various 'omics fields and also in the field of data science.

Analysis of the synovial macrophage transcriptome showed that the IL-6 pathway was upregulated in RA compared to PsA. This is insightful in two ways: firstly, blockade of the IL6R by the monoclonal antibody tocilizumab is used clinically when DMARDs fail to control disease activity in patients and secondly, IL-6 signals via the JAK/STAT pathway(244). Therefore this discovery confirms that JAK/STAT signaling is crucial in myeloid lineage cells in RA and that it is different from other arthropathies such as PsA.

As well as being released from synovial macrophages, IL-6 is secreted by fibroblast like synoviocytes and anecdotally there is evidence that in patients with resistant RA, inhibition of IL-6 is a better therapeutic option. IL-6 receptor expression is upregulated in RA synovial macrophages compared to PsA synovial macrophages. Furthermore other components of the IL-6 pathway such as PI3 Kinase and AKT are also upregulated. IL-6 binds to IL-6R and GP130 and then via JAK1 and JAK2 phosphorylates STAT3. IL-6 can also signal via the Ras/Raf pathway and therefore act with IL-1β and TLR signaling to activate NF-κB and further pro-inflammatory cytokine production.

On further inspection, JAK2 is relatively downregulated in RA synovial macrophages compared to PsA and this is likely to reflect that when a comparison is made, the changes are relative and not absolute. Furthermore, JAK1 is also involved with IL6 signaling and this is relatively increased in RA synovial macrophages. A JAK 1 inhibitor, filgotinib, has shown excellent response rates in the treatment of RA(245).

Furthermore, in PsA, IL21R is upregulated and IL21 signals via the common gamma chain and JAK3. This is not seen in RA synovial macrophages and also demonstrates the differential utilisation of the JAK/STAT pathway in these arthropathies.

Kinase inhibition is particularly attractive in the treatment of inflammatory arthritis because small molecules and therefore oral preparations can be used to treat patients as opposed to infusions or subcutaneous injections of cytokines. Furthermore, whereas inhibiting a cytokine at the soluble or receptor level will inhibit a particular pathway, inhibition of signal transduction machinery such as Janus kinases means that many more pathogenic pathways can be modulated. Tofacitinib showed excellent response rates in patients with RA when compared to methotrexate (69)however concerns over safety and also cost may have prevented a license in Europe being granted.

Other kinases such as PI3K sit further downstream of JAK and may offer further solutions although inhibition of larger parts of immune transduction pathways may be problematic. Tofacitinib was thought to be useful in this respect in that JAK2 is known to transduce the signal from growth factors such as erythropoietin and therefore relative inhibition of this kinase may lead to blood dyscrasias. Ruxolitinib, an inhibitor of JAK1/2 is licensed for the treatment of myelofibrosis but another JAK1/2 inhibitor baricitinib is under clinical trial for RA.

Therefore pathway analysis techniques have facilitated my exploration of this microarray dataset and I have determined that the JAK/STAT pathway is a potentially tractable target in myeloid lineage cells as well as what is known from the literature regarding targeting lymphocytes.

In conclusion, the analysis of array experiments should be predetermined at the outset in the same manner as for a clinical trial data analysis plan. The number of biological replicates, cell type or disease and whether the samples are being collected at one or multiple centers will allow a researcher and bioinformatician to make an informed decision *a priori* as to what is the best analysis plan i.e. which method of normalisation will be used and which statistical tests for differential expression. Furthermore on analysis of raw and normalised QC data, one can then decide whether a deviation from the plan is required and whether this is justified. In the end this should mean that more results move forward to validation and further investigation in hypothesis driven experiments.
Therefore, to further explore the role of JAKs in the synovial environment, I went on to employ a macrophage: T cell co-culture assay, which has been shown to simulate the RA synovial microenvironment. Therefore I will assess the role of JAK inhibition with various compounds to further explore this pathway in RA with a view to determining the

effect on two of the critical cell types, the synovial macrophage and also the synovial T-cell.

Chapter 4 The effect of Tofacitinib, a pan JAK inhibitor on a Macrophage: Tck cell contact activation assay

Some TNFα ELISAs were performed by Miss Ashley Gilmour. Luminex was performed by Mr Jim Reilly and Mrs Shauna Kerr.

4.1 Introduction

My prior bioinformatic analysis of the RA and PsA macrophage transcriptome revealed differences in the JAK/STAT pathway and therefore I decided to further explore this as inhibitors of this pathway exist although their exact mechanism of action is unclear. The perpetuation of synovial inflammation by macrophages stimulated by cytokine activated T-cells (Tck) may be modeled *in vitro* by the macrophage: T-cell co-culture cell contact assay(121,246-249). As well as antigen specific activation of T-cells by APCs, many groups, including our own, proposed that cytokines alone would provide the stimulus for memory T-cells to activate macrophages. Initial work showed that this effect was cell contact dependent and could also be recapitulated using T-cell membrane fragments and also externally fixed Tcks. The initial stimulus cocktail comprised of IL-2, IL-6 and TNFα but Tcks could also be generated by substituting IL-2 for IL-15 as well as by IL-15 alone. These cytokines are found within the inflamed rheumatoid joint and their role in pathogenesis is well documented in preclinical models with anti IL-15 antibodies being used in a proof of concept clinical trial in RA patients(61). In addition, blockade of TNFα and IL-6 have been employed clinically for the treatment of RA further demonstrating the crucial role for these cytokines.

Macrophages may be generated using an MCSF maturation protocol and specific functional subtypes can be generated using GMCSF, IFNγ, LPS or IL-4 in the final stage of maturation to give rise to inflammatory (GMCSF, IFNγ, LPS) or repair (IL-4/13) subtypes.. Tck express levels of CD18, CD69 and CD49d that are similar to RA synovial fluid derived T-cells(116). Furthermore, blockade of these markers leads to a partial reduction in TNFα production.

CD18 or integrin beta-2 is the beta chain that can be combined with various members of CD11 to form a complete integrin. It is responsible for the ability of leucocytes to extravasate from the blood to the tissue compartment. Furthermore a deletion in the CD18 gene in humans results in the inability of leucocytes to migrate from the blood compartment to tissues. This suggests that T-cells that can eventually form Tck are derived from circulating T-cells that are drawn into the tissue(250).

CD69 is a C-type lectin that is an early marker of activation on T-cell, platelets and NK cells. CD49d is an integrin alpha unit and therefore is involved with cell to cell signaling

and has been shown to interact with paxillin a molecule that may have a role in the pathogenesis of RA(251). Paxillin itself anchors cells to the extracellular matrix and is responsible for focal adhesion of cells and signals via tyrosine and src kinases.

CD49d also makes up part of the adhesion molecule VLA-4 that can be activated by fibronectin fragments and VCAM-1(252). VCAM-1 is upregulated on the endothelium of blood vessels in the RA synovium and soluble VCAM-1 is secreted by RA FLS and can act as a chemoattractant. Furthermore blockade of VCAM-1 by antibodies results in reduced Tck chemotaxis and migration across endothelium.

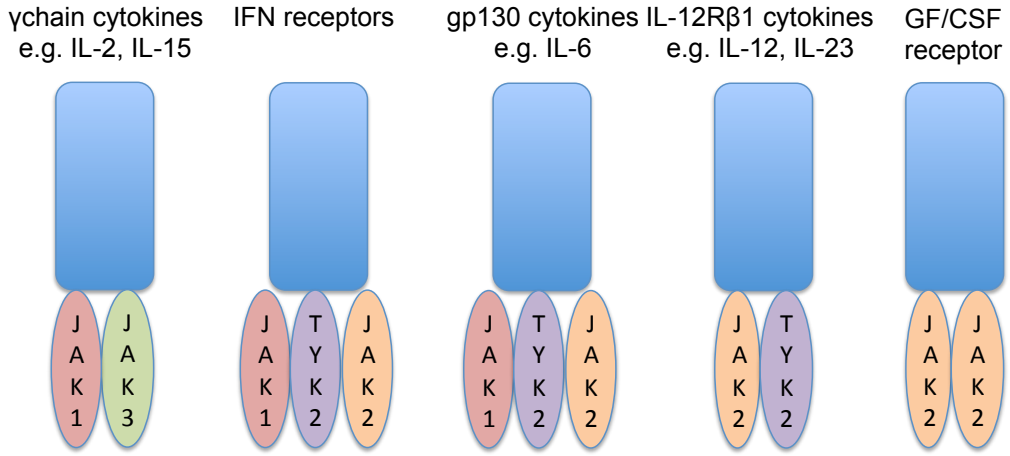
This therefore provides an ideal assays system in which to explore the effect of JAK inhibition on both the macrophage and T-cell, two cells that are crucial in the pathogenesis and the perpetuation of inflammation in RA. Tofacitinib is licensed for the treatment of RA and is approved by the FDA in the USA and is also available in Japan, Australia as well as other countries. However, it was not licensed by the EMA because of concerns regarding side effects and the lack of information regarding the exact mechanism of action.

Kinase inhibitors have been employed for many years in hematology with one of the first great triumphs of translational medicine being imatinib mesylate (Glivec) which targeted the ATP binding site of the constitutively activated tyrosine kinase bcr-abl implicated in chronic myeloid leukaemia (CML). This tyrosine kinase inhibitor was also found to be useful in other conditions such as gastrointestinal stromal tumours (GIST).

The initial mechanism of action of tofacitinib was thought to be via T-cells given that it is a JAK1/3 inhibitor and will therefore inhibit signals downstream of interferon receptors and cytokines which use the common gamma chain as a receptor which include but are not limited to IL-2, IL-6 and IL-15.

In addition to tofacitinib, other JAK inhibitors have been developed and ruxolitinib, a JAK1/2 inhibitor has been used to treat high-grade myelofibrosis(253). Clinical trials are currently underway to determine efficacy in psoriasis and alopecia areata. Baricitinib, a newer JAK1/2 inhibitor, has shown significant efficacy in rheumatoid arthritis phase II trials(78). For this reason I used ruxolitinib as a further interventional drug in my *in vitro* studies. A final kinase inhibitor, AG-490 tyrphostin, is not employed clinically but is a potent inhibitor of JAK2 but at a higher IC_{50} than for either tofacitinib or ruxolitinib and

was used as a further experimental condition. A summary of the various cytokine families targeted by tofacitinib and ruxolitinib is presented in figure 4-1 below.



	γchain cytokines e.g. IL-2, IL-15	IFN receptors	gp130 cytokines e.g. IL-6	IL-12Rβ1 cytokines e.g. IL-12, IL-23	GF/CSF receptor
Tofacitinib JAK1/3	✔	✔	✔	✗	✗
Ruxolitinib JAK1/2	✔	✔	✔	✔	✔

Figure 4-1 Cytokine and growth factor receptor families that are targeted by tofacitinb and ruxolitinib. JAK1/3 inhbition with tofacitinib would target signaling from gamma chain cytokines as well as interferon and gp130 associated cytokines such as IL-6. Although originally thought to target only JAK3, there is ample evidence of JAK1 inhibition. Ruxolitinib targets both JAK1/2 with similar efficacy and would also target signals from IL-12 receptor and growth factor and colony stimulating factor receptors. The later could contribute to side effects but may deliver improved efficacy in RA if GM-CSF is inhibited to a degree.

I hypothesise that inhibiting the JAK/STAT pathway in both T-cells and Macrophages will impact their activation and that tofactinib will be able to hit multiple pathogenic cell targets as well as multiple cytokine pathways.

The key questions addressed in this chapter are as follows:

1. Does JAK inhibition prevent the release of TNFα when macrophages are cell contact activated by Tck?

2. Does JAK inhibition prevent the formation of Tck?

3. Does JAK inhibition prevent LPS driven TNFα release from macrophages?

4. Does JAK inhibition prevent the release of inflammatory cytokines and chemokines that are important in the pathogenesis of RA?

4.2 MCSF macrophages produce TNFα in a concentration dependent manner when cell-contact activated by live cytokine activated CD4 T-cells

Before investigating the effect of JAK inhibition on the macrophage: Tck cell contact activation assay, I wished to demonstrate that I could perform the assay in my hands consistently, that cell contact was a required event and that DMSO, used as a drug vehicle, would not affect the assay.

Briefly, monocytes were cultured at density of $5x10^5$ cells per ml in a 96 well plate with MCSF supplementation. CD4 T cells ($1x10^6$ cells/ml) were cultured in flasks for six days and were supplemented with IL-2, IL-6 and TNFα. Prior to co-culture, T-cells were thoroughly washed, counted and assessed for viability using trypan blue. T cells were added to macrophages at various concentrations and allowed to remain in culture for 24 hours. Supernatants were harvested and TNFα concentration measured using an ELISA and results from two representative donors are shown in figure 4-2.

TNFα is produced when macrophages are cell contact activated by Tck in keeping with what has previously been demonstrated in the literature(254,255). TNFα production is dependent on the concentration of Tck added to macrophages and macrophages alone produce little TNFα without stimulation. Finally, there is variability in the amount of cytokine produced by representative donors although the trends seen above are consistent.

Figure 4-3 demonstrates that Tck do not produce significant amounts of TNFα when cultured alone and therefore the large amount of TNFα produced in the assay is likely due to cell contact. Although there is a significant literature supporting the use of this assay, I decided to use transwell membranes to inhibit cell contact and therefore show that this event was crucial in this assay.

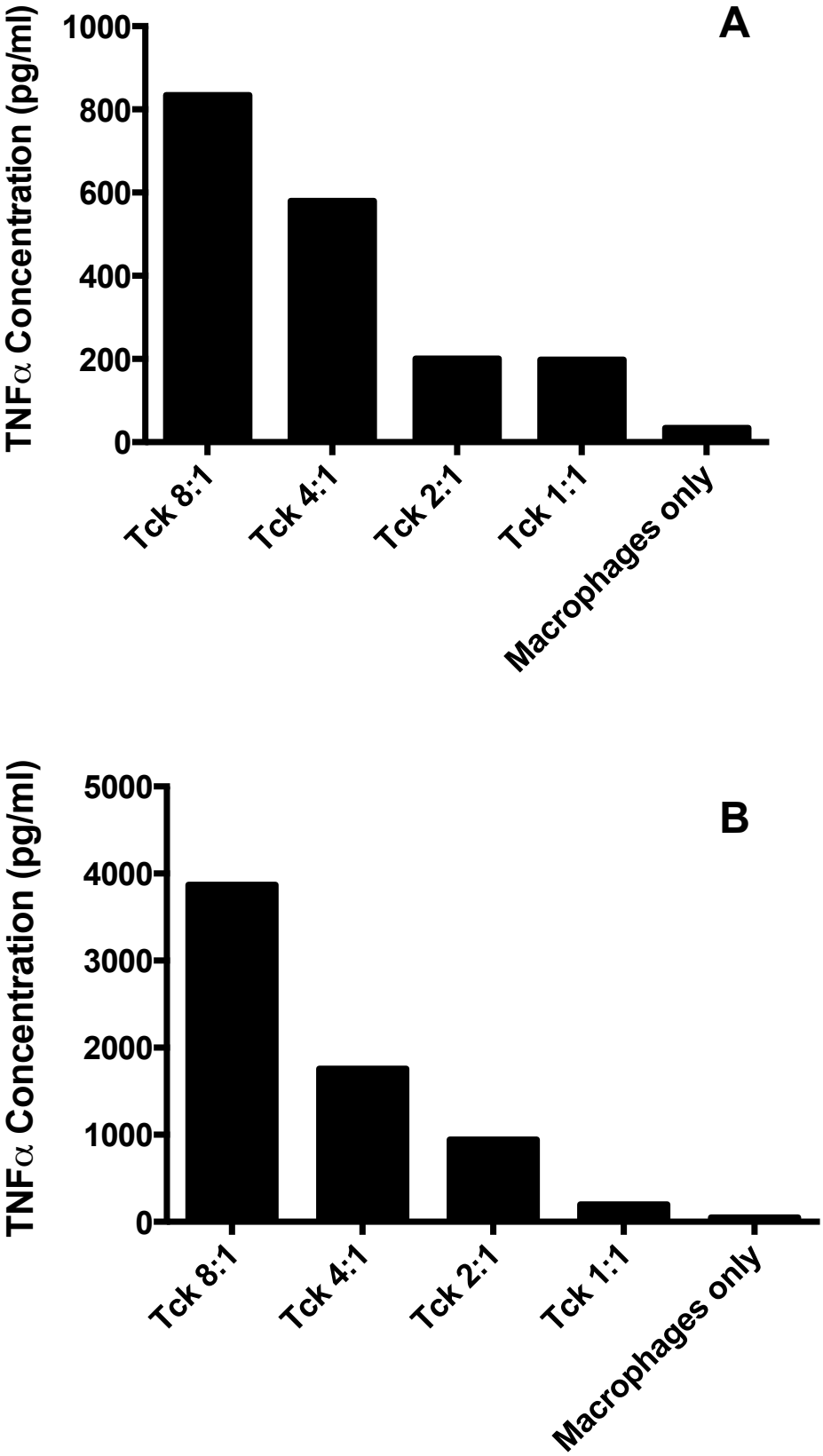


Figure 4-2 TNFα is produced in a concentration dependent manner when macrophages are cell contact activated by Tck. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. They were cultured at a density of $5x10^5$ cells/ml for 6 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-2 (25ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at the concentrations shown and co-cultured for 24 hours. Supernatants were harvested and an ELISA for TNFα performed. (**A**) – representative donor 1 (**B**) – representative donor 2. TNFα is not produced by macrophages alone. There is variability between the two donors.

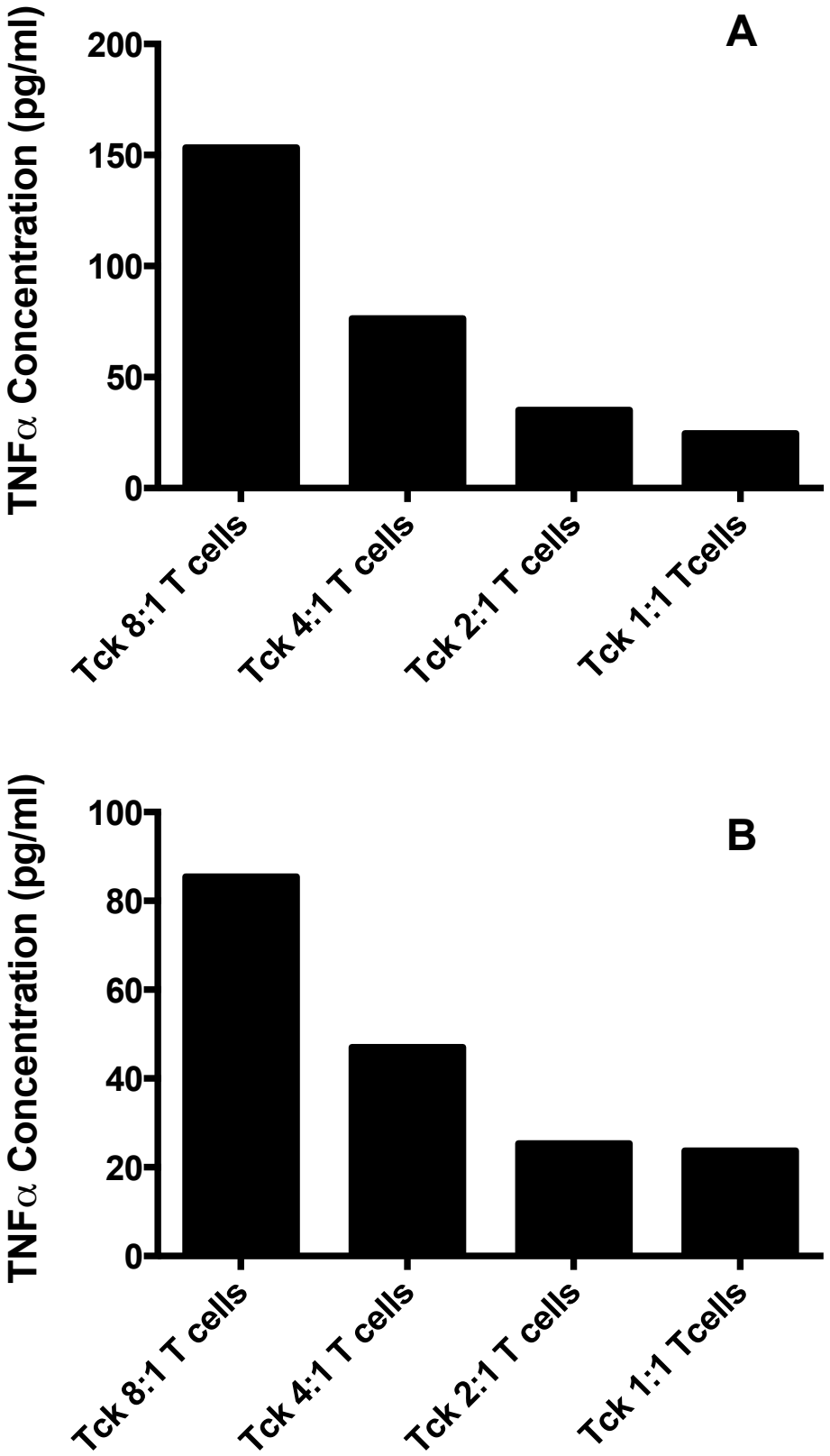


Figure 4-3 TNFα is not produced in significant amounts by Tck. CD4 T cells were obtained by density centrifugation of healthy donor buffy coats and positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-2 (25ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and cultured in a 96 well plate for 24 hours. Supernatants were harvested and an ELISA for TNFα performed. (**A**) – representative donor 1 (**B**) – representative donor 2. TNFα is not produced at significant levels by Tck alone.

4.3 Preventing cell contact by using transwell membranes prevents the production of TNFα in the macrophage: Tck cell contact activation assay

Figure 4-4 demonstrates that cell contact is a crucial factor in the macrophage: Tck cell-contact activation assay. When Tck were placed into a cell culture insert containing a membrane to prevent cell contact but still allow soluble factor transition, TNFα release was prevented to a concentration in keeping with that produced by macrophages in a resting state.

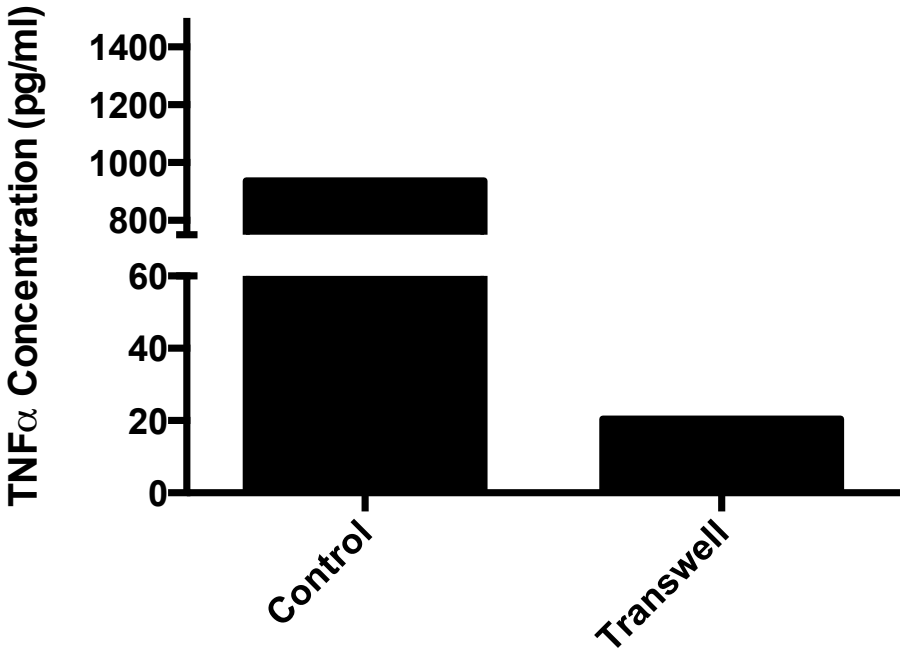


Figure 4-4 Preventing cell contact by using a transwell membrane to separate macrophages and Tck prevents production of TNFα. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. They were cultured at a density of $5x10^5$ cells/ml for 6 days in complete medium in the presence of MCSF (50ng/ml) in a 12 well plate. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-2 (25ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at the concentrations of 1 macrophage to 4 Tck and co-cultured for 24 hours. In the transwell condition, Tck were placed into an insert containing a transwell membrane with pore size of 0.4µm therefore inhibiting cell contact. Supernatants were harvested and an ELISA for TNFα performed. Data shown is representative of three biological replicates.

4.4 Using DMSO as a drug vehicle does not affect the production of TNFα in the macrophage: Tck cell contact activation assay

I had to dissolve tofacitinib in DMSO for use in my *in vitro* experiments. However high concentrations of DMSO are toxic and therefore I went on to demonstrate that DMSO concentration which I would use in subsequent experiments would not adversely affect the production of TNFα in the cell contact activation assay.

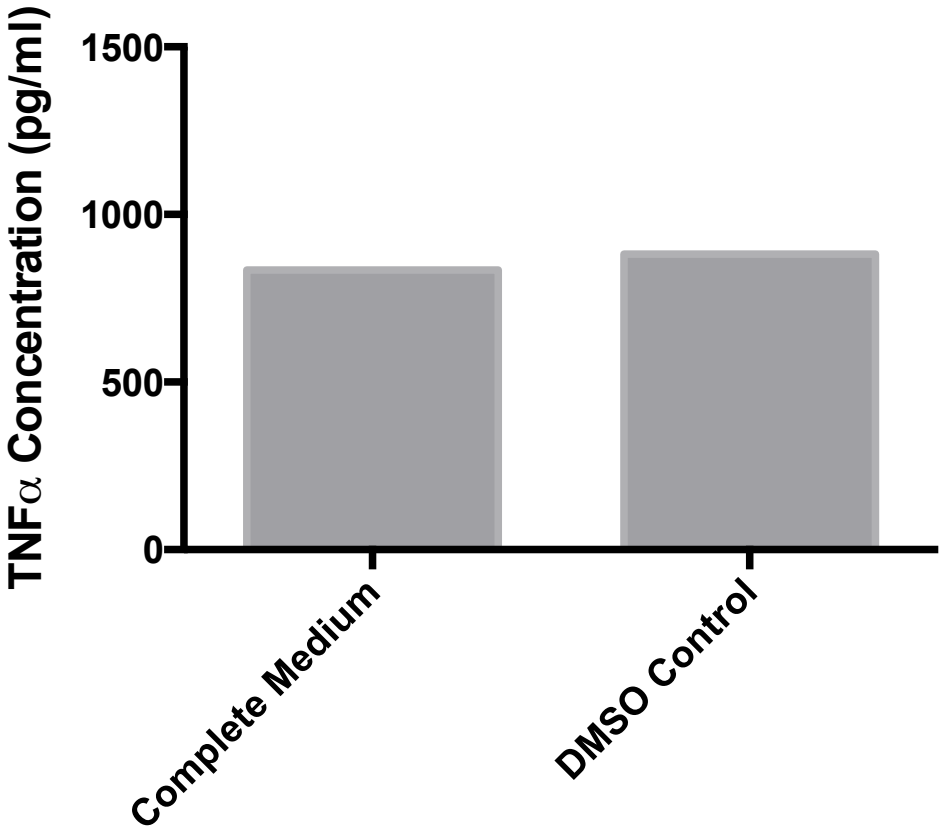


Figure 4-5 TNFα is produced at a similar level in the presence of 0.001% DMSO. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. They were cultured at a density of $5x10^5$ cells/ml for 6 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-2 (25ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 8 Tck and co-cultured for 24 hours. DMSO was added at a concentration of 0.001% in keeping with what would be required in subsequent experiments. Supernatants were harvested and an ELISA for TNFα performed. Data is representative of one donor from three biological replicates.

4.5 Tofacitinib may lead to a reduction in TNFα production in the macrophage: Tck cell contact activation assay

I went on to investigate the effect of JAK inhibition in the context of varying concentrations of Tck in the cell contact activation assay. My aim was to determine a Tck concentration that I could use in subsequent experiments with various JAK inhibitors.

Therefore, having prepared macrophages and Tck as described previously, I performed co-culture experiments using tofacitinib at two concentrations: 100nM and 1000nM. These concentrations correspond to approximate blood levels that would be reflected by peak and trough dosage of tofacitinib by mouth(256). Both macrophages and Tck were pre-incubated with tofacitinib for one hour prior to co-culture and supernatants were collected after overnight culture.

Tck at a concentration of 1 macrophage to 4 Tck give a strong cell contact activation signal to macrophages (figure 3-5). Furthermore, 1000nM tofacitinib inhibits the production of TNFα in this assay to approximately 25% of control conditions. Therefore fixing Tck concentration at 1:4 gives a balance between a strong activation signal, the ability to detect inhibition of cytokine production by JAK inhibition and is also experimentally practical.

In conclusion, I have shown that in my hands, TNFα is produced when macrophages are cell contact activated by Tck. I have further demonstrated that this is in fact cell contact dependent by the use of transwell membranes. Also, TNFα production depends on the concentration of Tck in the assay and the production of TNFα is prevented by tofacitinib.

Therefore I have decided to further investigate the role of JAK inhibitors on TNFα production in the macrophage: Tck cell contact activation assay with cell concentration controlled at 1 macrophage to 4 Tck. Furthermore, I decided to replate macrophages at day 3 of culture to make cell numbers more consistent in further assays. Finally, I decided to use IL-15 in lieu of IL-2 to produce Tck because IL-2 levels in synovial fluid are low(257), IL-15 is detectable and also shown to be important in Tck development(254).

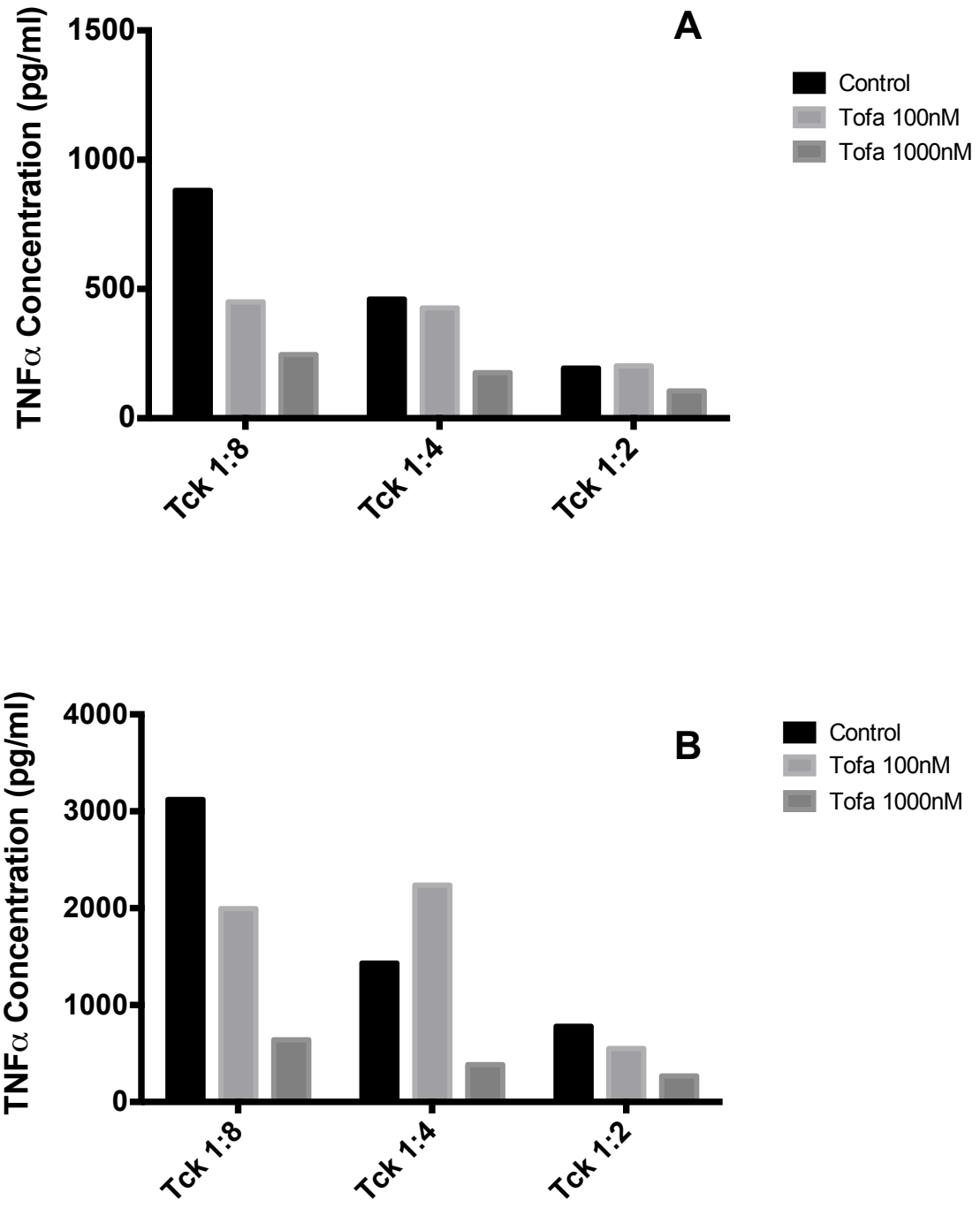


Figure 4-6 Tofacitinib reduces cell contact mediated TNFα production in a concentration dependent manner. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 6 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-2 (25ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at concentrations shown and co-cultured for 24 hours. Tofacitinib was dissolved in DMSO and added to both macrophages and Tck one hour prior to co-culture. Supernatants were harvested and an ELISA for TNFα performed. Data is representative of two donors from three biological replicates. Tck at a concentration of I macrophage to 4 Tck provide a strong cell contact activation signal and this is inhibited by tofacitinib at 1000nM.

4.6 Tofacitinib decreases TNFα production in the Macrophage: T Cell assay in a concentration dependent manner

Tofacitinib decreases the production of TNFα from macrophages following Tck cell contact activation in a concentration dependent manner (figure 4-7). Briefly, CD14+ monocytes were differentiated into macrophages from five healthy buffy coat donors by density centrifugation, CD14+ positive magnetic bead separation and culture with MCSF for 6 days. Macrophages were replated at day 3 to improve experimental consistency. CD4+ T-cells were isolated from the CD14 negative fraction and isolated using CD4+ positive magnetic bead selection. CD4+ T-cells were cultured in complete medium supplemented with IL-6, IL-15 and TNFα for 6 days to produce Tck.

Both macrophages and Tck were incubated with tofacitinib or vehicle control in complete medium for 1 hour prior to co-culture for 24 hours. Supernatants were harvested for TNFα ELISA and luminex analysis.

Tofacitinib inhibits TNFα production in a concentration dependent manner following cell contact activation of macrophages by Tck (figure 4-7). These results show that 1000nM tofacitinib reduces TNFα production in this assay to 10% of control. Therefore I have demonstrated that JAK 1/3 inhibition is able to reduce TNFα production in the macrophage: Tck cell contact activation assay.

Following this, I went further to investigate whether two other small molecule inhibitors, Ruxolitinib, a JAK1/2 inhibitor and AG-490 Tyrphostin, a JAK2 and EGFR inhibitor would have a similar effect or whether JAK3 inhibition was crucial to disrupting TNFα production in the cell contact activation assay.

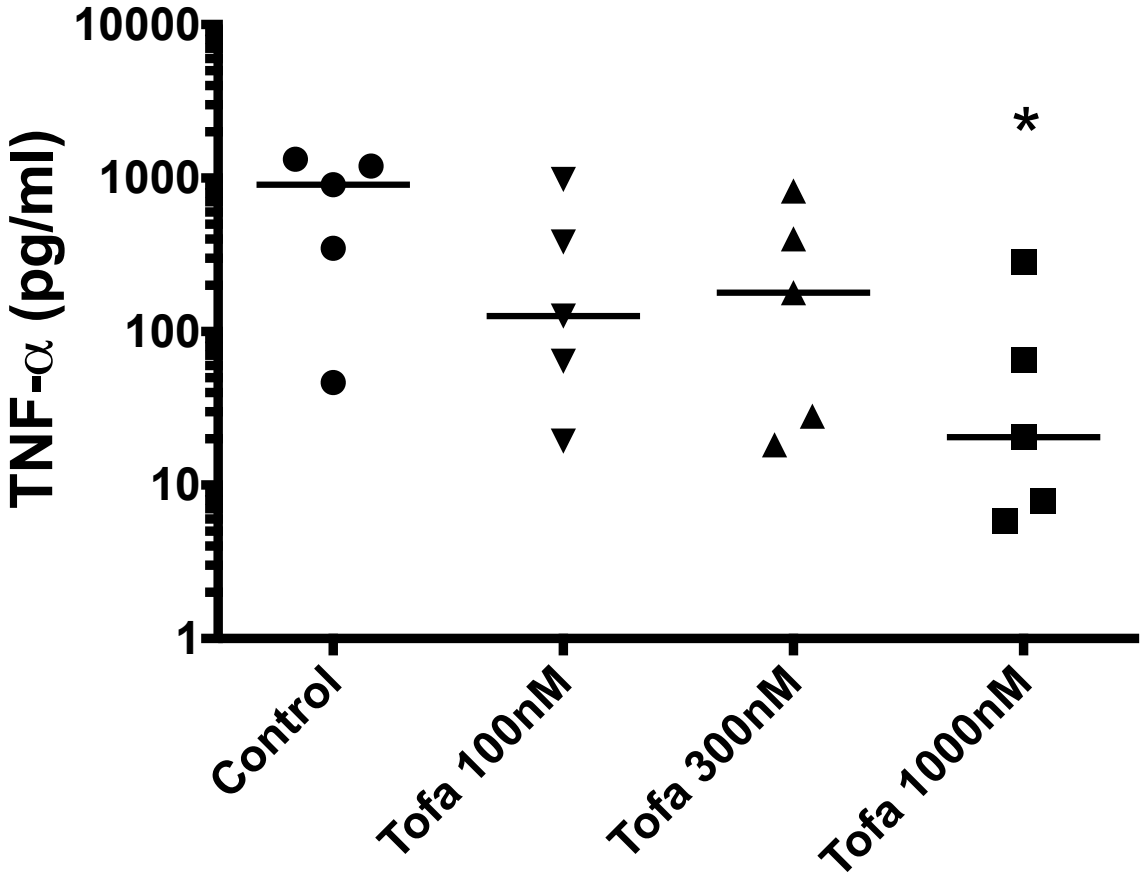


Figure 4-7 Tofacitinib consistently reduces cell contact mediated TNFα production in a concentration dependent manner. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with tofacitinib for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for TNFα performed. Data shows the results of five biological replicates with line representing median. Tofacitinib decreases the production of TNFα in the macrophage: Tck cell contact assay. This effect is concentration dependent and there is no significant production of TNFα by macrophages at rest. Friedman test with Dunn's multiple comparison test was used to calculate statistical difference of inhibitor treated conditions versus control. * = $p<0.05$. (n=5 biological replicates)

4.7 Ruxolitinib, a JAK1/2 inhibitor and AG-490 tyrphostin, a JAK2/EGFR inhibitor also decrease TNFα production from macrophages following Tck stimulation

Ruxolitinib is another kinase inhibitor that is used in the treatment of myelofibrosis but is also currently in clinical trial as a treatment for RA. Furthermore, another JAK1/2 inhibitor, baricitinib is under investigation and recently reported promising results at the American College of Rheumatology 2015 meeting with significantly better efficacy signals compared to TNFα inhibition.

I was therefore interested to explore the effect of relatively sparing JAK3 from inhibition, as it was the inhibition of JAK 3 that was theoretically most important for autoimmunity when JAKs were discovered. The reasoning behind this was that JAK3 was predominantly expressed on immune cells; in particular lymphocytes, and therefore inhibition of this target would result in fewer off target effects. In contrast, JAK2 is heavily involved with growth factor signaling molecules such as erythropoietin and therefore inhibition could lead to anaemia and indeed we do see this in clinical trials of the JAK1/2 inhibitors although it is concentration responsive(253,258).

Both ruxolitinib and AG 490 tyrphostin decreased TNFα production in the cell contact assay (figure 4-8). Ruxolitinib was particularly effective at decreasing cell contact mediated TNFα production at the same concentrations as tofacitinib. Despite this, there was no statistically significant difference in the production of TNFα when tofacitinib or ruxolitinib were used. Ag 490 tyrphostin also decreased TNFα production but the dose of this drug was 50 times higher than either tofacitinib or ruxolitinib.

Whereas both tofacitinib and ruxolitinib are licensed medications, AG 490 tyrphostin failed because of lack of potency. I decided to investigate it because it was able to weakly inhibit JAK2 but also had actions on other kinases out with the JAK family and therefore allowed me to compare JAK inhibition with inhibition of other kinases in this assay.

Having demonstrated that both tofacitinib, a JAK1/3 inhibitor, and ruxolitinib, a JAK1/2 inhibitor, decreased cell contact mediated TNFα production, I then went on to investigate whether these drugs could also prevent the formation of Tck *in vitro*.

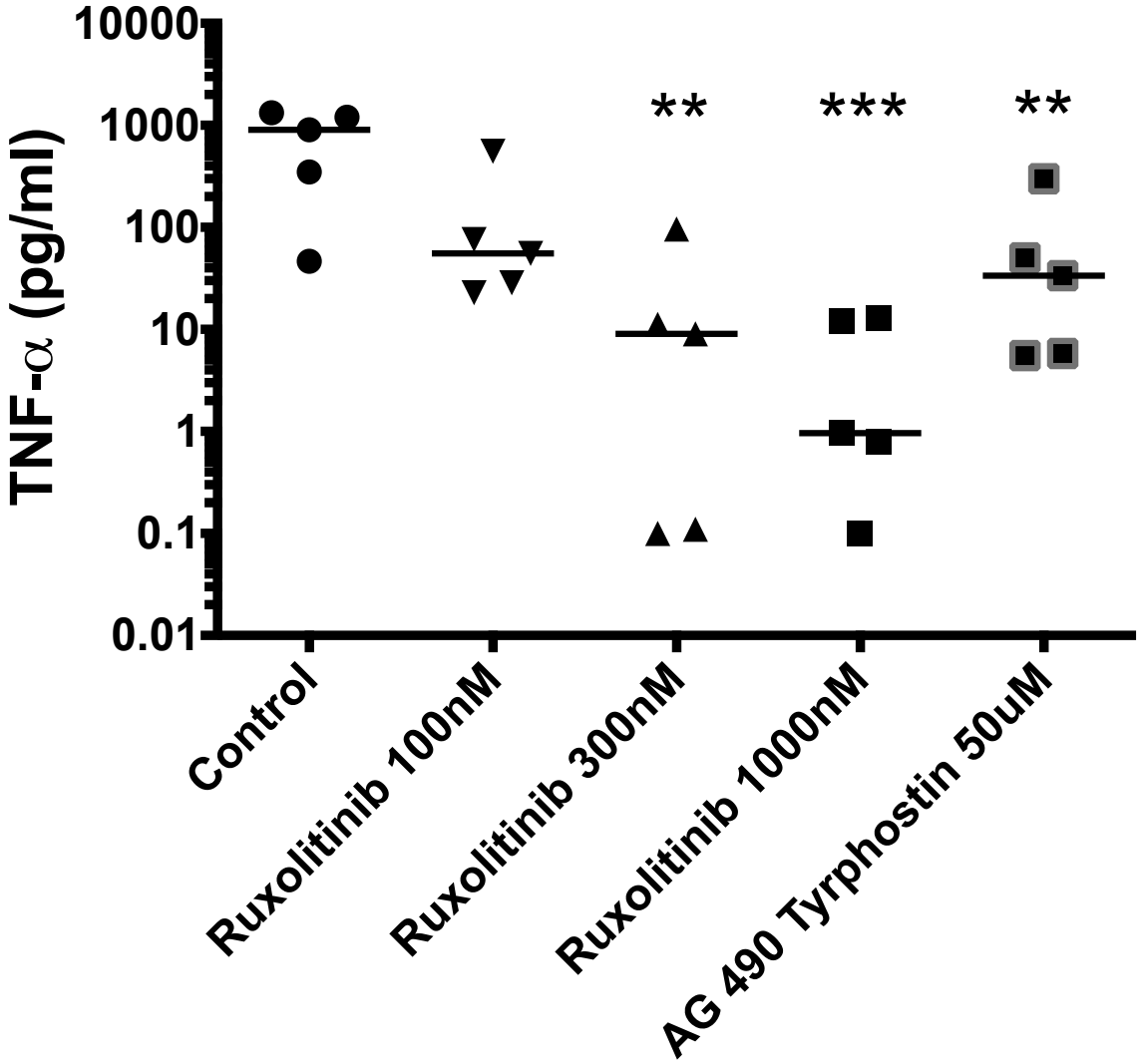


Figure 4-8 Ruxolitinib and AG 490 tyrphostin consistently reduce cell contact mediated TNFα production in a concentration dependent manner. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with JAK inhibitors for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for TNFα performed. Data shows the results of five biological replicates with line representing median. JAK inhibitors decrease the production of TNFα in the macrophage: Tck cell contact assay. This effect is concentration dependent and there is no significant production of TNFα by macrophages at rest. Friedman test with Dunn's multiple comparison test was used to calculate statistical difference of inhibitor treated conditions versus control. **= $p<0.01$, *** = $p<0.001$. (n=5 biological replicates)

4.8 Tofacitinib and ruxolitinib prevent the cytokine induced maturation of Tck *in vitro*

Briefly, Tck are matured by obtaining CD4+ T-cells from blood and culturing in complete media supplemented with TNFα, IL-6 and IL-15. Theoretically, JAK inhibition should disrupt signaling via IL-6 and IL-15. Tck arise from the effector memory compartment as described by Brennan et al (116)and are similar to synovial CD4 T cells in that they express high levels of CD69, CD18 and CD49d making these cells an accepted surrogate for a CD4 T-cell in an inflamed rheumatoid joint.

Therefore, I went on to explore whether tofacitinib and ruxolitinib would prevent the formation of Tck by adding JAK inhibitors to the culture medium during the 6 day maturation period. The T-cells were then washed thoroughly and cultured with mature macrophages as previously described at a concentration of 1 macrophage to 4 T-cells.

When tofacitinib and ruxolitinib were added to the Tck maturation cocktail, the resulting T-cells were unable to induce TNFα production from macrophages in co-culture (figure 4-9). Furthermore, when these cells were counted, I noticed that they were smaller than Tck with their morphology in keeping with freshly isolated CD4+ T-cells. Therefore I can conclude that JAK inhibition with tofacitinib and ruxolitinib does prevent the maturation of Tck.

It is likely that both molecules inhibit both IL-6 and IL-15 signaling and others have demonstrated that these cytokines are crucial in the formation of mature Tck(117,254,255). Furthermore, there is also evidence from McInnes et al (254)that IL-15 alone is sufficient to form Tck and this cytokine signals via JAK1 and JAK3 and has been shown to alter the balance between effector and memory T-cells.

Finally in this section I went on to investigate whether JAK inhibitors could prevent LPS mediated TNFα production from macrophages. TLR4 is key pattern recognition receptor on myeloid lineage cells and there is evidence that ACPA can act as TLR4 ligands and therefore perpetuate inflammation(35,135,259). JAK inhibitors should not be able to prevent LPS mediated TNFα production, as the JAK/STAT pathway is not used for primary signal transduction.

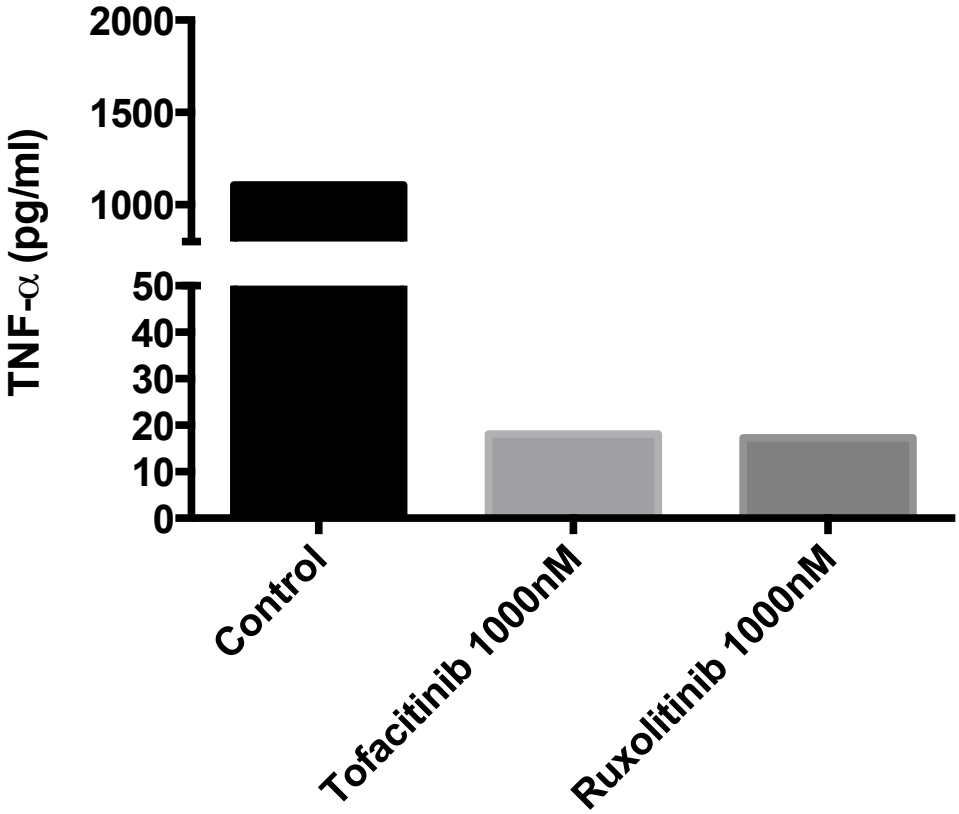


Figure 4-9 Addition of tofacitinib or ruxolitinib during Tck maturation results in T-cells that are unable to drive cell-contact mediated TNFα production. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Culture medium was supplemented with inhibitor for the duration of Tck culture. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Supernatants were harvested and luminex for TNFα performed. Data shows an example of one representative experiment. JAK inhibition results in T-cells which are unable to drive cell-contact mediated TNFα production from macrophages.

4.9 LPS induced production of TNFα, by macrophages is not inhibited by tofacitinib

TNFα production in response to TLR4 stimulation by LPS was used as a positive control in the co-culture experiments since JAK/STAT inhibition is not expected to contribute to the primary effect of this pathway. However TLR4 stimulation of myeloid lineage cells is a key inflammatory pathway that is implicated in the pathogenesis of RA. I therefore explored the effect of inhibiting JAKs in conjunction with LPS stimulation and found no significant evidence of an effect on TNFα production by macrophages in response to LPS (figure 4-10) although this is based on two replicates. This result is similar to findings from other groups, although they showed that JAK inhibition increased the production of TNFα in response to LPS by inhibiting IL-10 release. I could not replicate this finding but with more biological replicates this could be confirmed because there was a trend to higher TNFα production with LPS and tofacitinib. However, the significance of this result in a system where cytokine concentration is already high is unclear(260).

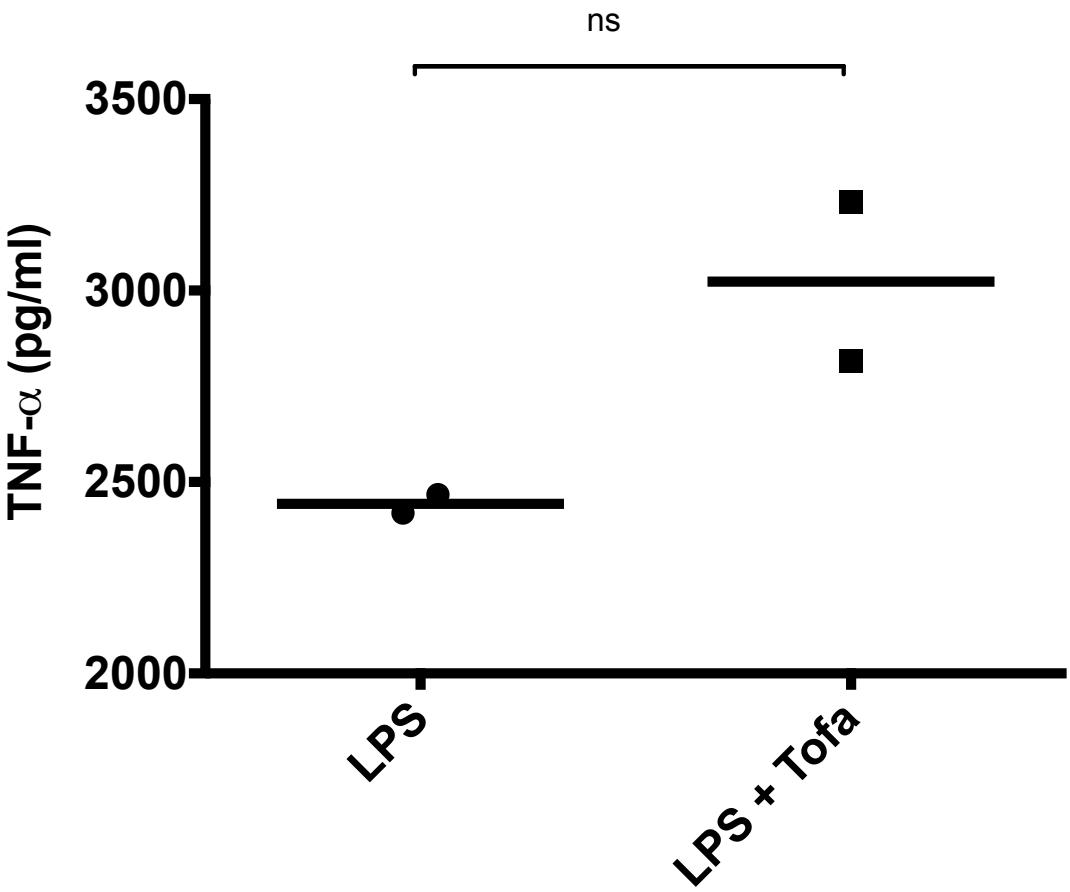


Figure 4-10 Tofacitinib is unable to prevent LPS mediated TNFα production from macrophages. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. Macrophages were incubated with 1000nM tofacitinib or DMSO vehicle control for one hour and then stimulated with LPS (1ng/ml) for 24 hours. Supernatants were harvested and luminex for TNFα performed. Data shows two biological replicates. Tofacitinib does not prevent LPS mediated production of TNFα from macrophages. A paired t-test with a significance level of 0.05 was performed to detect statistical difference.

In summary:

1. JAK inhibition using small molecules decreases TNFα production in a concentration dependent manner when macrophages are cell contact activated by cytokine activated T-cells

2. JAK inhibition with tofacitinib and ruxolitinib prevents the formation of cytokine activated T-cells *in vitro* from freshly isolated blood CD4+ T-cells

3. Tofacitinib does not inhibit LPS mediated TNFα production from macrophages

Therefore in this assay, that is an accepted model of the interactions that occur in the synovial microenvironment, we see that the production of TNFα, a crucial cytokine in the pathogenesis of RA is decreased by tofacitinib. This demonstrates that JAK1/3 inhibition interrupts TNFα production although it I cannot conclude whether this is a direct or indirect mechanism.

Furthermore this effect is not specific to JAK1/3 inhibition and both ruxolitinib and AG 490 tyrphostin are able to decrease TNFα production in this assay suggesting that JAK1 and JAK2 inhibition mediates this effect. This finding is in keeping with encouraging clinical trial results of other JAK inhibitors in RA such as baricitinib, which targets JAK 1/2 (78)and filgotinib(245), which specifically targets JAK1. Both of these agents showed efficacy in improving joint swelling and inflammation and had favourable side effect profiles compared to tofacitinib. This is compared to decernotinib, a JAK3 specific inhibitor that has been dropped from further development following publication of a recent trial(79). Although the drug was effective at controlling inflammation in RA, abnormal liver function tests, changes to plasma lipids and also herpes zoster infection occurred more commonly in the decernotinib treated group. These side effects were seen in clinical trials of tofacitinib suggesting that JAK3 inhibition may be playing a role in these side effects as they were not as marked in the baricitinib or filgotinib trials.

If I had more time, I would explore the role of specific JAK inhibition using tool compounds such as those available from the MRC Technology drug library. This would allow me to specifically inhibit each JAK and determine which combination of JAK inhibition is most effective at reducing TNFα production.

Furthermore, tofacitinib and ruxolitinib prevent the formation of cytokine activated T-cells that is likely due to inhibition of signaling downstream of IL-6 and IL-15 receptors. This finding in combination with the previous result in the co-culture assay suggests that JAK inhibition will be effective at not only disrupting cell contact mediated inflammation at a synovial site but will also prevent the development of further Tck. This has clinical implications: firstly JAK inhibition would be effective at treating inflammation where the process is established, such as in the case of RA, by disrupting macrophage and T-cell cross talk and preventing T-cells from developing an inflammatory phenotype and secondly they are unlikely to be useful in preventing the transition of pre-RA typified by autoantibody positivity without overt inflammation. Agents that target the B-cell or antigen presenting cells specifically may be more suited to this although there has been encouraging work with Bruton's tyrosine kinase in the treatment of RA(80).

Tck express CD18, CD49d and CD69 in keeping with synovial CD4 T-cells. Therefore, logical further experiments would be to investigate whether JAK inhibition prevents the upregulation of these surface molecules during Tck maturation and also if their expression is downregulated when mature Tck are exposed to JAK inhibitors. Blockade of these markers was shown by Brennan et al (116) to reduce TNFα production and therefore if JAK inhibitors down-regulate these this may suggest a mechanism for the reduction in TNFα produced in this assay.

Finally all experiments have been performed using live Tck and both macrophages and Tck are able to produce TNFα. Therefore using fixed Tck would help to determine the effect on the macrophage specifically thereby further elucidating the mechanism for this finding.

Although TNFα is a key cytokine in the pathogenesis of RA, IL-6, IL-15 and IFNγ also have a role(261). Furthermore anti-inflammatory cytokines may be affected and therefore I went on to investigate the production of various cytokines and chemokines in the cell contact activation assay using a human cytokine luminex. Unfortunately the IFNγ assay failed in this system and this would be a crucial further experiment if I had more time and to further explore this a time course cell contact activation experiment would reveal which cytokine IFNγ or TNFα increased first.

4.10 IL-6 production following macrophage cell contact activation by Tck is reduced by JAK inhibitors

IL-6 receptor blockade has been employed clinically in RA (53)as well as other inflammatory conditions such as juvenile idiopathic arthritis (262,263)and giant cell arteritis(264). IL-6 is produced following cell contact activation of macrophages by Tck and is shown to be contact dependent (figure 4-11). IL-6 is also found in significant concentrations in both RA synovial fluid and is used to produce Tck and therefore decreasing IL-6 by JAK inhibition may result in an alteration of synovial T-cell to an anti-inflammatory phenotype.

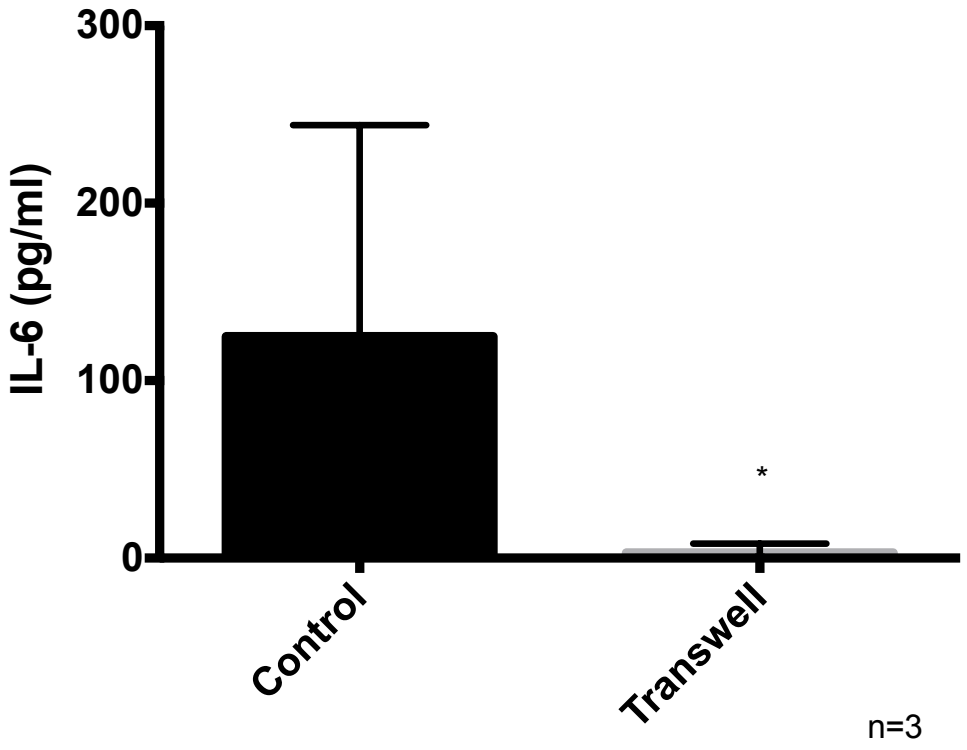


Figure 4-11 Preventing cell contact by using a transwell membrane to separate macrophages and Tck prevents production of IL-6. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. They were cultured at a density of $5x10^5$ cells/ml for 6 days in complete medium in the presence of MCSF (50ng/ml) in a 12 well plate. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at the concentrations of 1 macrophage to 4 four Tck and co-cultured for 24 hours. In the transwell condition, Tck were placed into an insert containing a transwell membrane with pore size of 0.4μm therefore inhibiting cell contact. Supernatants were harvested and a luminex for IL-6 performed. Student's t-test was used to determine statistical difference. * = $p < 0.05$. n=3 biological replicates. Error bars show standard deviation.

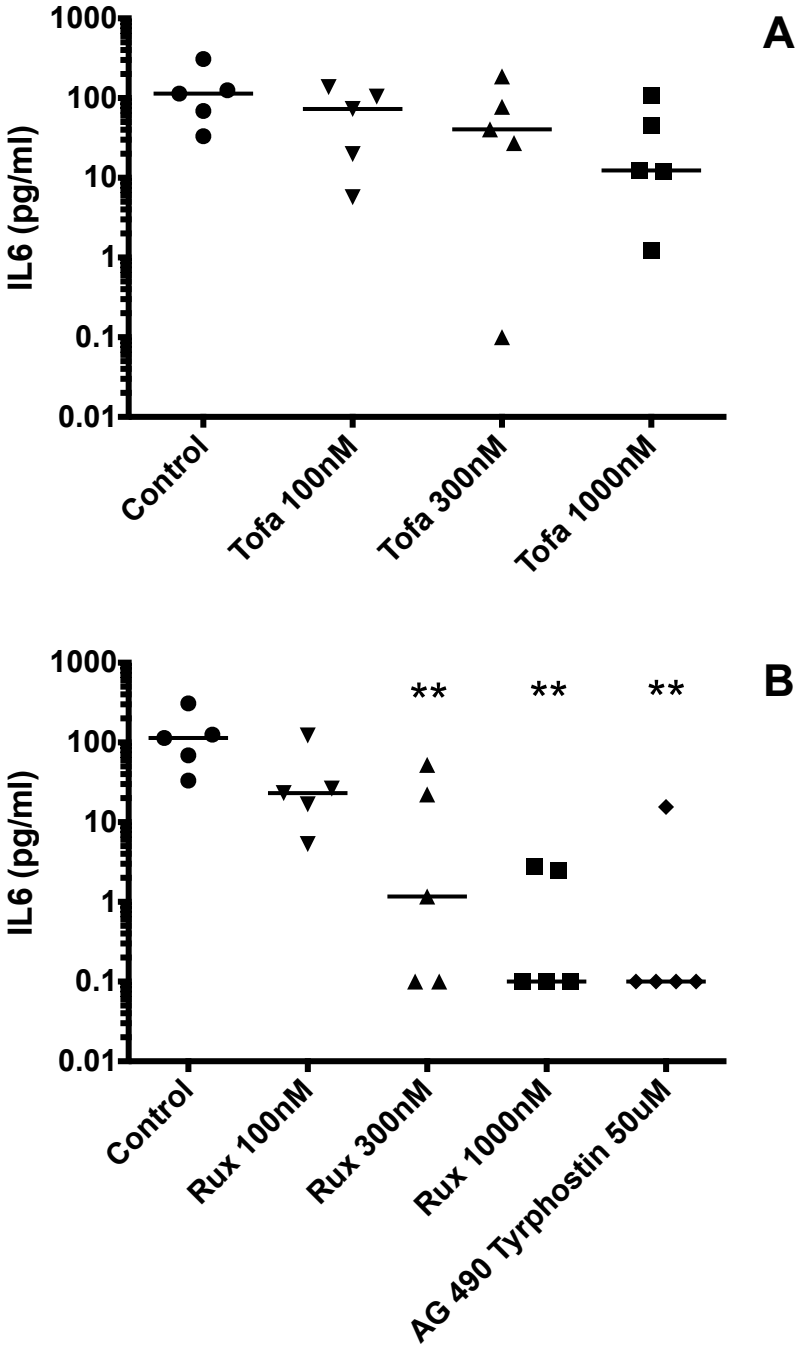


Figure 4-12 JAK inhibitors reduce cell contact mediated IL-6 production in a concentration dependent manner. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with inhibitors for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for IL-6 performed. Tofacitinib may decrease the production of IL-6 in the macrophage: Tck cell contact assay but it is not statistically significant. There is no significant production of IL-6 by macrophages at rest. Ruxolitinib and Ag 490 Tyrphostin decrease IL-6 production. Friedman test with Dunn's multiple comparison test was used to calculate statistical difference of inhibitor treated conditions versus control. **= $p<0.01$. (n=5 biological replicates)

There is no IL-6 production by macrophages (figure 4-12) or T-cells at rest (data not shown). IL-6 is produced by cell contact activation and this is decreased by tofacitinib, ruxolitinib and AG 490 tyrphostin (figure 4-12). Although ruxolitinib statistically decreases IL-6 production at a lower concentration of inhibitor, when tofacitinib is compared to ruxolitinib there is no statistically significant difference (data not shown). Finally tofacitinib does not alter LPS mediated IL-6 release from macrophages (data not shown).

4.11 IL-15 is not produced by macrophages or Tck at rest but is produced following co-culture and this is reduced by JAK inhibition

IL-15 is produced when macrophages are cell-contact activated by Tck (figure 4-13). IL-15 production is decreased by ruxolitinib more than the equivalent concentration of tofacitinib. The production of IL-15 is completely cell contact dependent and is prevented by transwell membranes (data not shown). Furthermore, in this case, ruxolitinib 1000nM is statistically better than tofacitinib 1000nM at decreasing IL-15 production following cell contact activation (data not shown). Additionally, compared to TNFα and IL-6, AG 490 tyrphostin had no statistically significant effect on cell contact activation mediated IL-15 production.

In contrast with TNFα and IL-6, tofacitinib inhibits LPS mediated IL-15 production by macrophages (figure 4-14). This combined with results of the cell contact assay suggests that the mechanism leading to IL-15 production by both cell contact activation and LPS is different from that leading to the production of TNFα and IL-6. In the case of both TNFα and IL-6, JAK inhibition had no effect on LPS mediated production of cytokine from macrophages but this is clearly decreased in the case of IL-15. Further experiments using other TLR agonists and TLR agonism as a further stimulus in the macrophage: Tck cell contact assay would lead to further insight into whether this effect occurs across TLR agonists and and also suggest which pathway is over riding in the case of IL-15: cell contact activation or TLR mediated production from macrophages.

Furthermore, AG 490 tyrphostin is not effective at significantly decreasing IL-15 production in the macrophage: Tck cell contact activation assay.

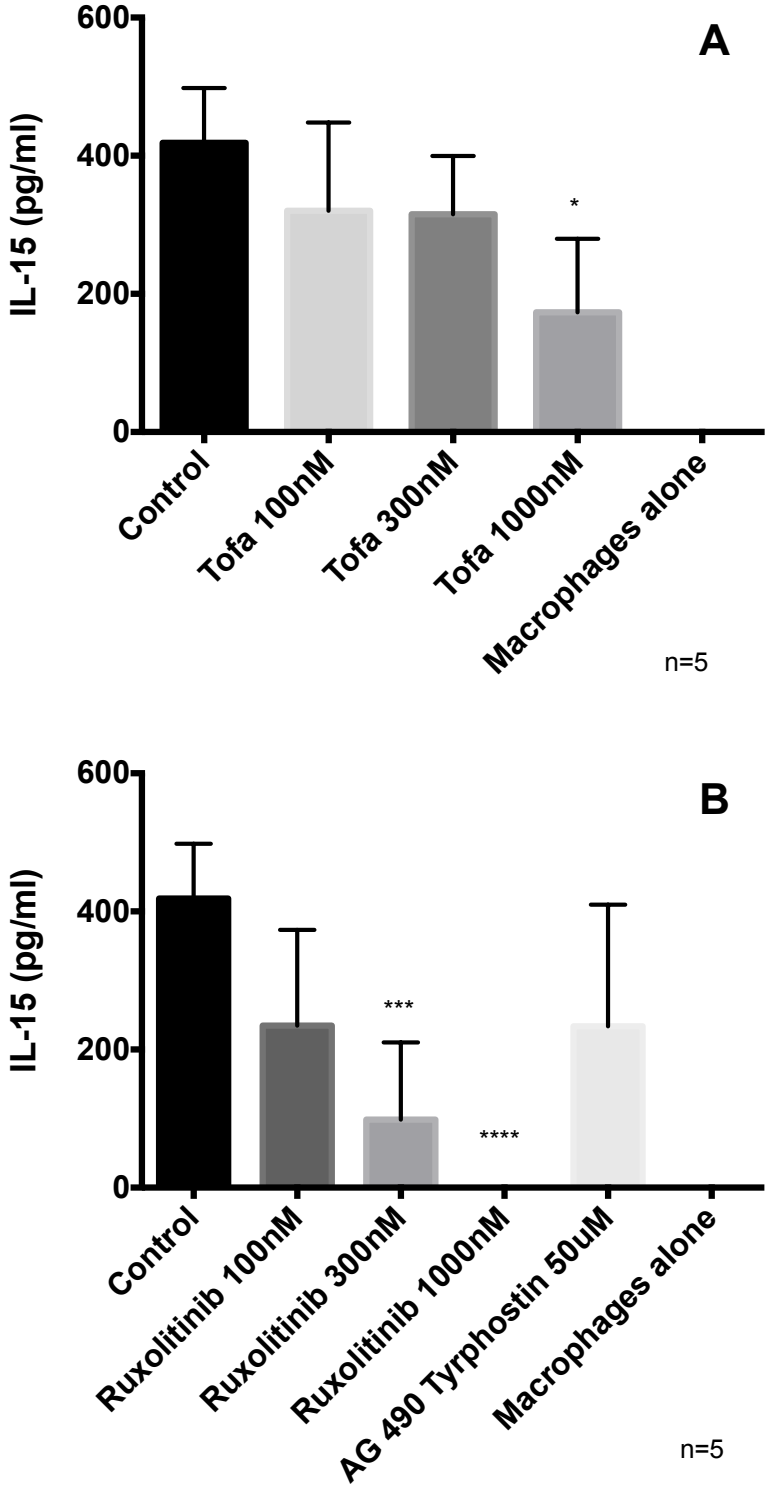


Figure 4-13 Ruxolitinib decreases cell contact mediated IL-15 production in a concentration dependent manner. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with inhibitors for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for IL-15 performed. Both tofacitinib and ruxolitinib decrease the production of IL-15 in the macrophage: Tck cell contact assay. This effect is concentration dependent and there is no significant production of IL-6 by macrophages at rest. AG 490 tyrphostin does not significantly reduce IL-15 production in this assay. Repeated measures one-way ANOVA with Bonferroni's multiple testing was used to calculate statistical difference of inhibitor treated conditions versus control. * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$, **** = $p<0.0001$. (n=5 biological replicates). Error bars show standard deviation.

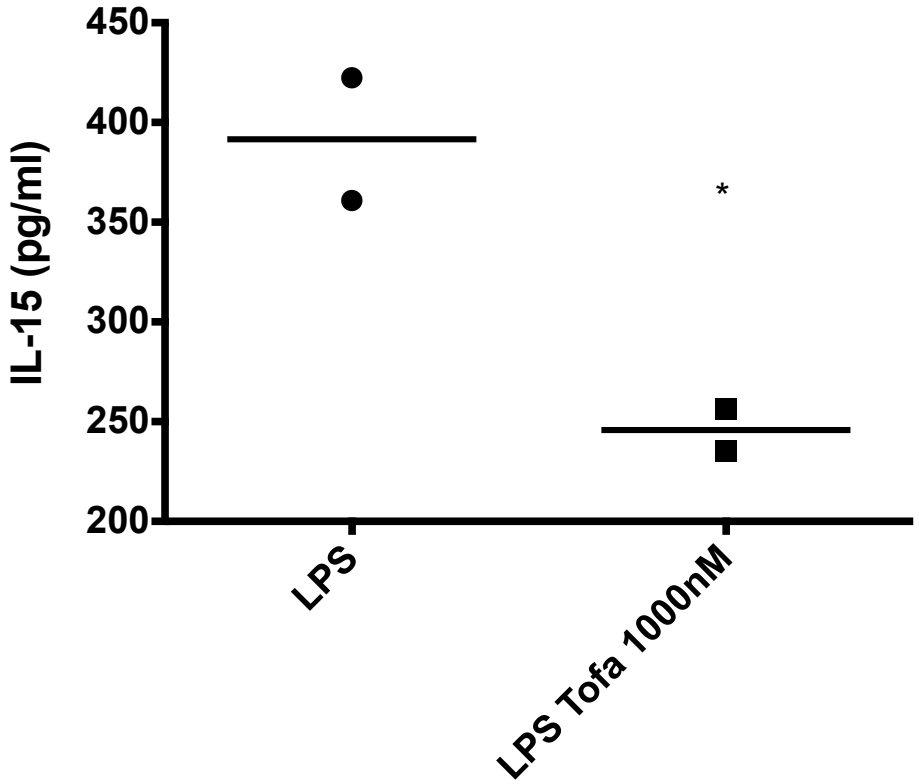


Figure 4-14 Tofacitinib statistically decreases LPS mediated IL-15 production from macrophages. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. Macrophages were incubated with 1000nM tofacitinib or DMSO vehicle control for one hour and then stimulated with LPS (1ng/ml) for 24 hours. Supernatants were harvested and luminex for IL-15 performed. Data shows two biological replicates. Tofacitinib decreases LPS mediated production of IL-15 from macrophages. A paired t-test was performed to detect statistical difference. * = $p<0.05$

4.12 IL1RA is decreased by JAK inhibitors when macrophages are cell contact activated by Tck

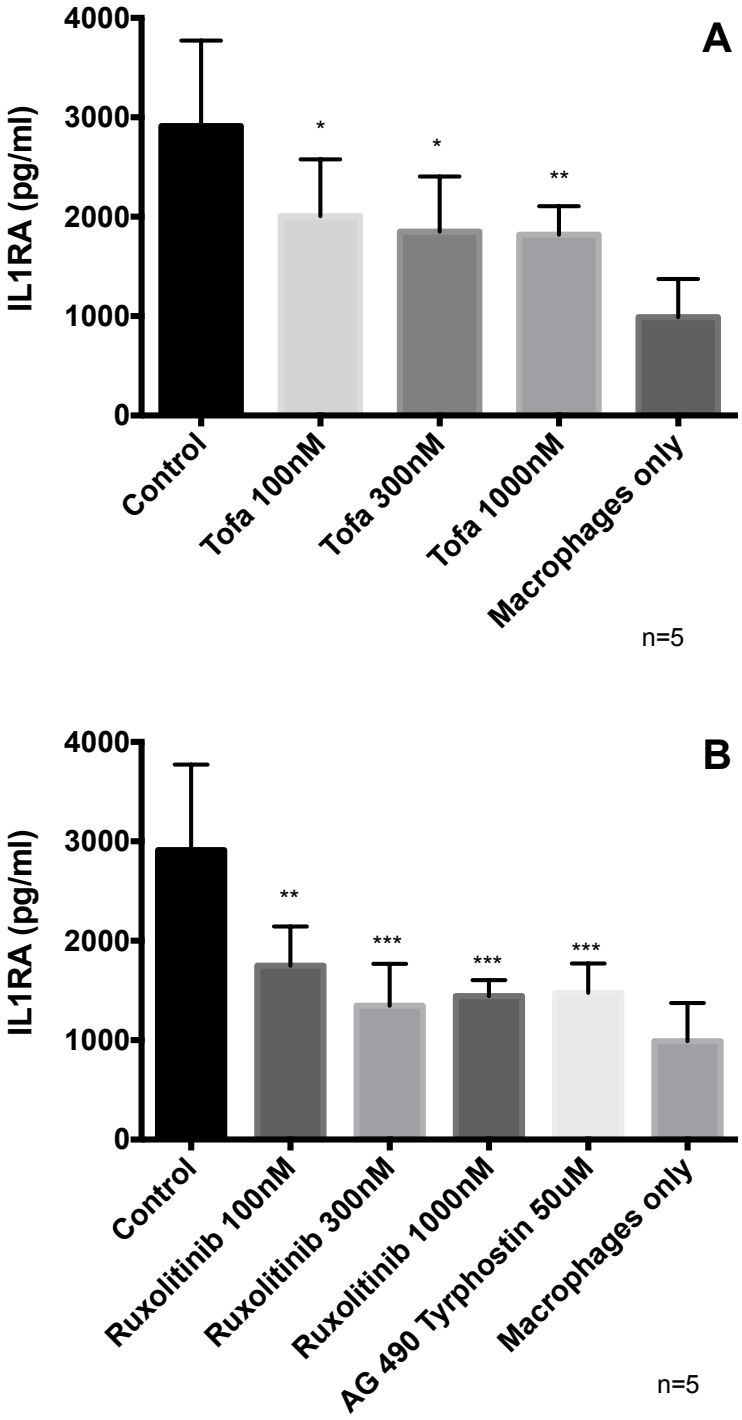


Figure 4-15 Macrophages produce IL1RA at rest, this is increased by cell contact activation and decreased by JAK inhibitors. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with inhibitors for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for IL1RA performed. Both tofacitinib and ruxolitinib decrease the production of IL1RA in the macrophage: Tck cell contact assay. This effect is concentration dependent there is production at rest by macrophages. Repeated measures one-way ANOVA with Bonferroni's multiple testing was used to calculate statistical difference of inhibitor treated conditions versus control. * = $p<0.05$, **= $p<0.01$, *** = $p<0.001$. (n=5 biological replicates). Error bars show standard deviation.

Macrophages produce IL1RA at rest (figure 4-15) and this is increased by cell contact activation. Tck produce small amounts of IL1RA (data not shown) but the increased production is far in excess of an additive effect. All of the small molecule inhibitors reduce IL1RA production with a suggestion that ruxolitinib may be more effective than tofacitinib although this is not statistically significant. Furthermore, levels of IL1RA do not reduce below basal production suggesting that a JAK dependent mechanism causes upregulation of this anti-inflammatory cytokine. Tofacitinib does not affect LPS mediated production of IL1RA from macrophages (data not shown).

4.13 LPS mediated production of IL-10 from macrophages is decreased by tofacitinib

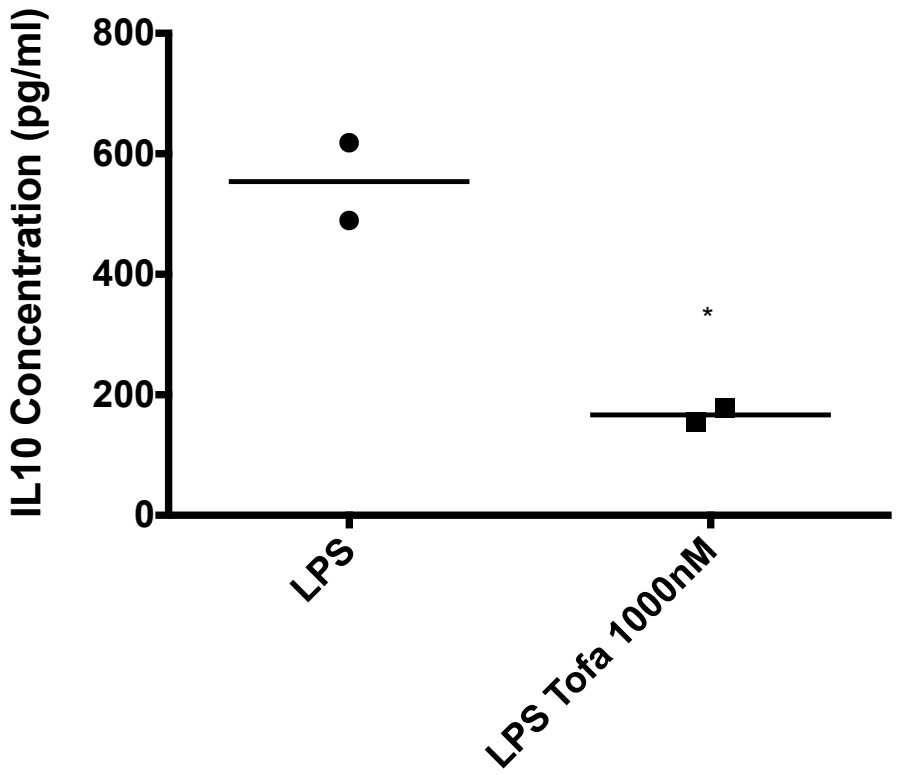


Figure 4-16 Tofacitinib statistically decreases LPS mediated IL-10 production from macrophages. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. Macrophages were incubated with 1000nM tofacitinib or DMSO vehicle control for one hour and then stimulated with LPS (1ng/ml) for 24 hours. Supernatants were harvested and luminex for IL-10 performed. Data shows two biological replicates. Tofacitinib decreases LPS mediated production of IL-10 from macrophages. A paired t-test was performed to detect statistical difference. * = $p<0.05$

IL-10 is not produced by macrophages or Tck at rest and is not upregulated after cell contact activation by Tck (data not shown). It is produced following LPS stimulation of macrophages, as others have shown, and is decreased by tofacitinib.

4.14 MIP1α and MIP1β are both released by macrophages following cell contact activation by Tck and are decreased by JAK inhibition

Neither MIP1α nor MIP1β are produced at rest by macrophages (figure 4-17 and 4-18) or Tck (data not shown). They are both upregulated on cell contact activation and this is inhibited by the use of transwell membranes (data not shown). Tofacitinib, ruxolitinib and AG 490 tyrphostin decrease MIP1α and MIP1β although in the case of MIP1α this is statistically significant for ruxolitinib 1000nM. The error bars denote standard deviation and they are wide showing that there is considerable variation between experiments. Despite this, the trend within an experiment shows that JAK inhibition decreased MIP1α and MIP1β in a concentration dependent manner.

Both MIP1α and MIP1β are produced following LPS stimulation of macrophages and tofacitinib is unable to alter this (data not shown).

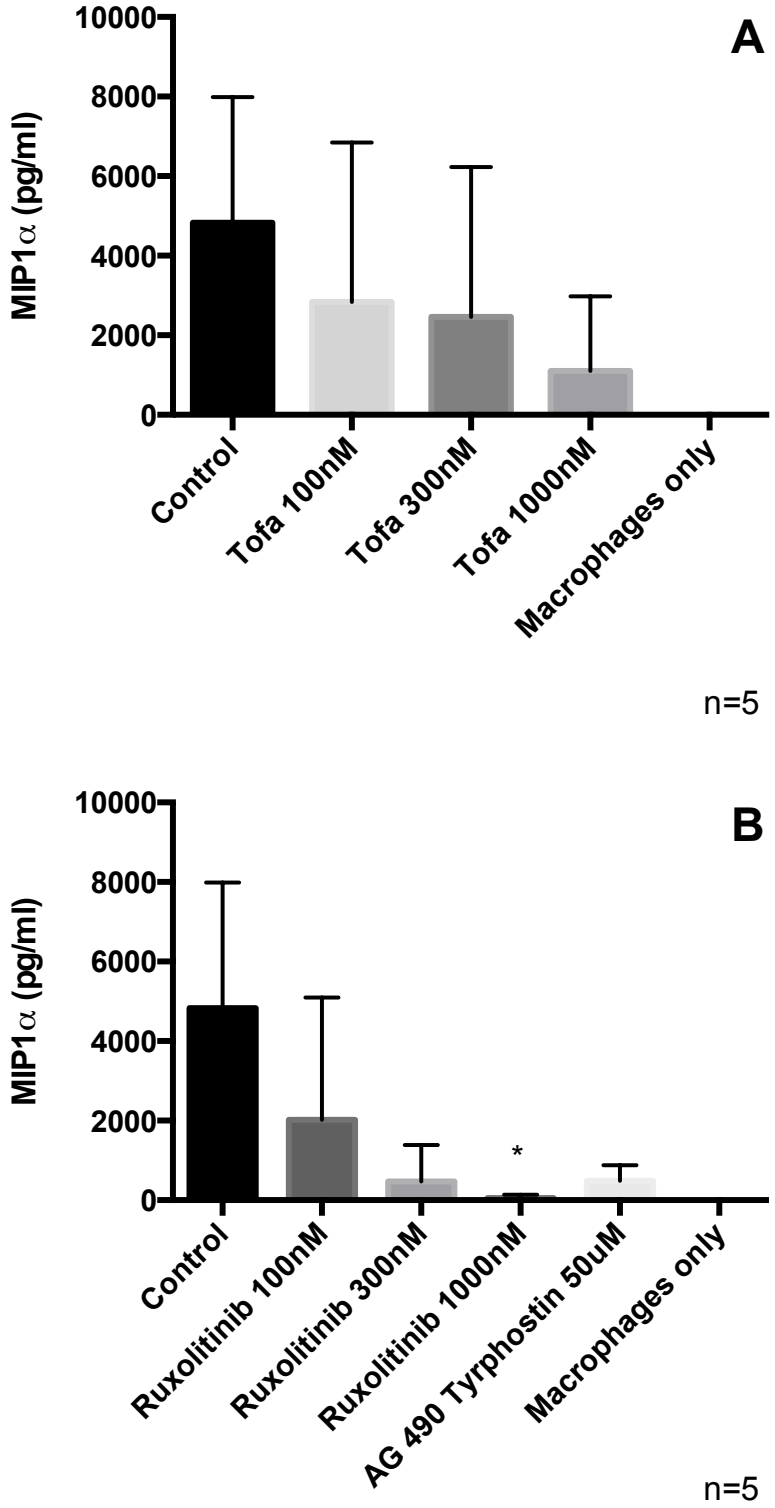


Figure 4-17 On cell contact activation by Tck, macrophages produce MIP1α and this is decreased by ruxolitinib. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with inhibitors for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for MIP1α performed. Both tofacitinib and ruxolitinib decrease the production of MIP1α in the macrophage: Tck cell contact assay and this is statistically significant with ruxolitinib 1000nM. Repeated measures one-way ANOVA with Bonferroni's multiple testing was used to calculate statistical difference of inhibitor treated conditions versus control. * = $p<0.05$. (n=5 biological replicates). Error bars show standard deviation.

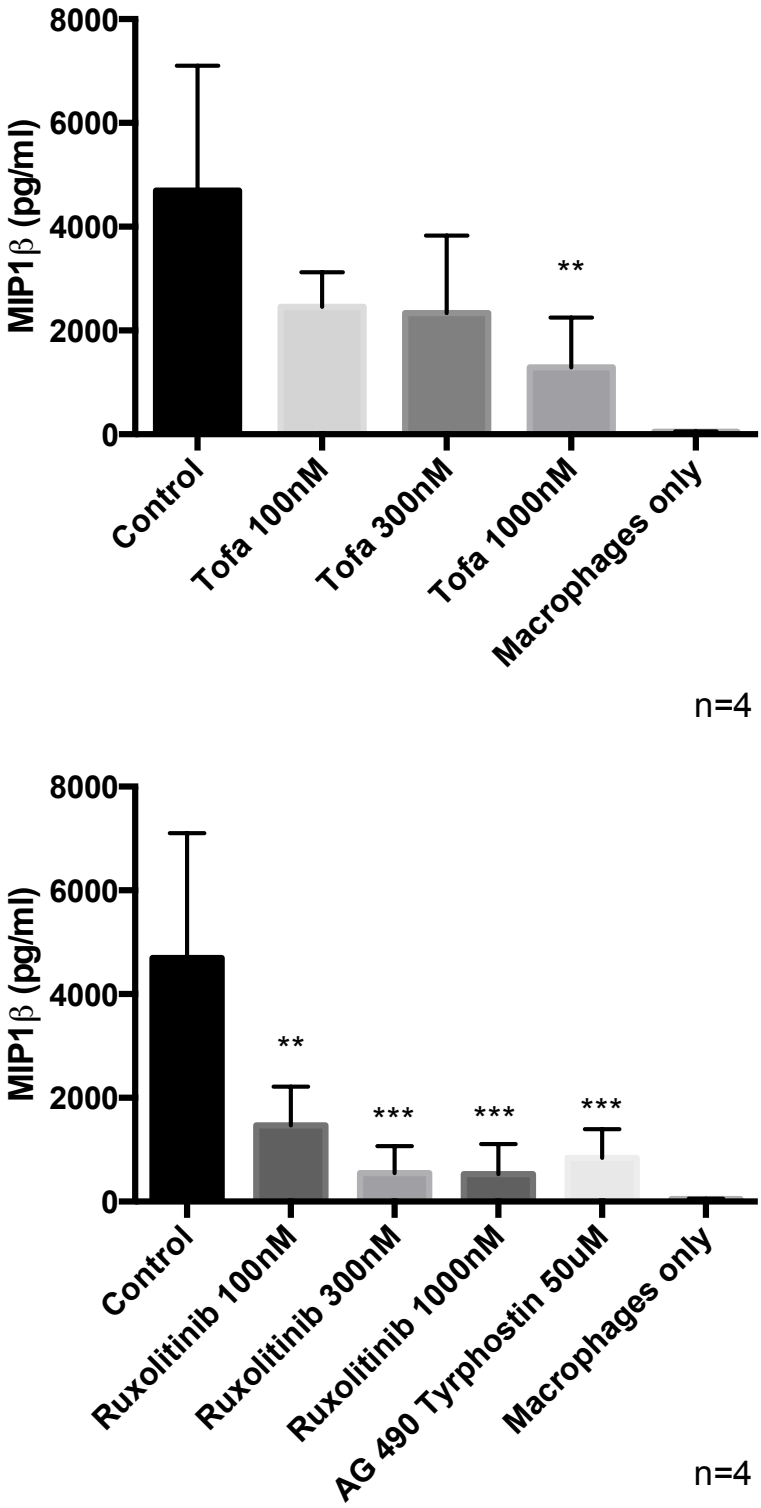


Figure 4-18 On cell contact activation by Tck, macrophages produce MIP1β and this is decreased by JAK inhibition. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with inhibitors for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for MIP1β performed. Tofacitinib, ruxolitinib and AG 490 tyrphostin decrease the production of MIP1β in the macrophage: Tck cell contact assay and this is statistically significant at multiple inhibitor doses. Repeated measures one-way ANOVA with Bonferroni's multiple testing was used to calculate statistical difference of inhibitor treated conditions versus control. * = $p<0.05$, **= $p<0.01$, *** = $p<0.001$. (n=4 biological replicates – the luminex assay failed for one experiment). Error bars show standard deviation.

4.15 IP10 is produced following LPS stimulation of macrophages and this is decreased by tofacitinib.

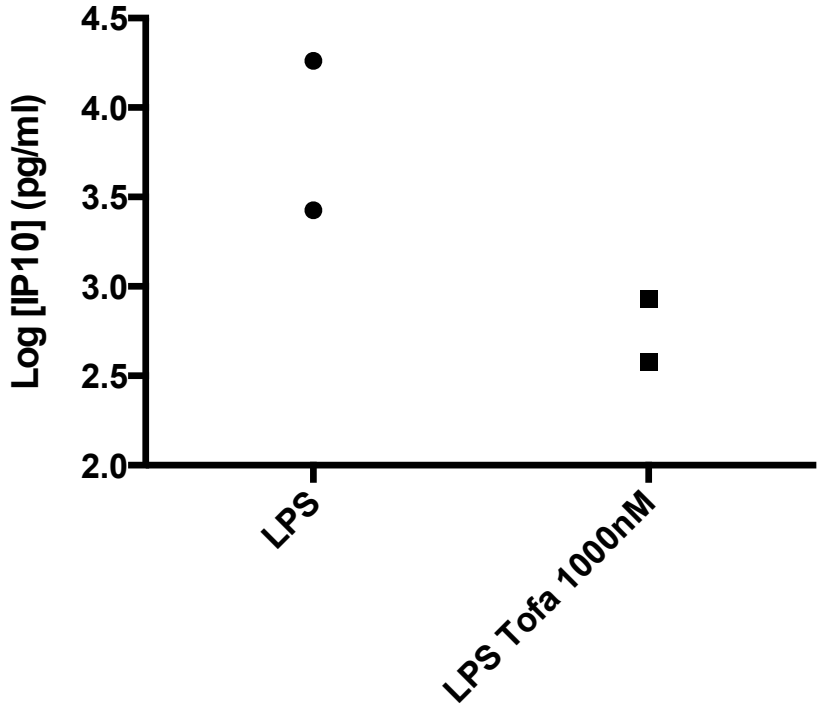


Figure 4-19 Tofacitinib may decrease LPS mediated IP10 production from macrophages. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. Macrophages were incubated with 1000nM tofacitinib or DMSO vehicle control for one hour and then stimulated with LPS (1ng/ml) for 24 hours. Supernatants were harvested and luminex for IP10 performed. Data shows two biological replicates. Tofacitinib may decrease LPS mediated production of IP10 from macrophages, however the result is non significant. A paired t-test was performed to detect statistical difference.

IP10 is produced by macrophages following LPS stimulation and this is reduced by treatment with tofacitinib (figure 4-19). IP10 is secreted in response to IFNγ (265)and in this case it is likely that IFNγ release following LPS stimulation leads to IP10 production. As stated previously, the luminex for IFNγ failed, and this is a limitation of my current work. Therefore I suggest that tofacitinib via JAK1 inhibition is disrupting IFNγ signaling in this instance.

Both tofacitinib and ruxolitinib decrease IP10 production following cell contact activation of macrophages by Tck (figure 4-20). Use of transwell membranes to prevent cell contact prevented the release of IP10 (data not shown).

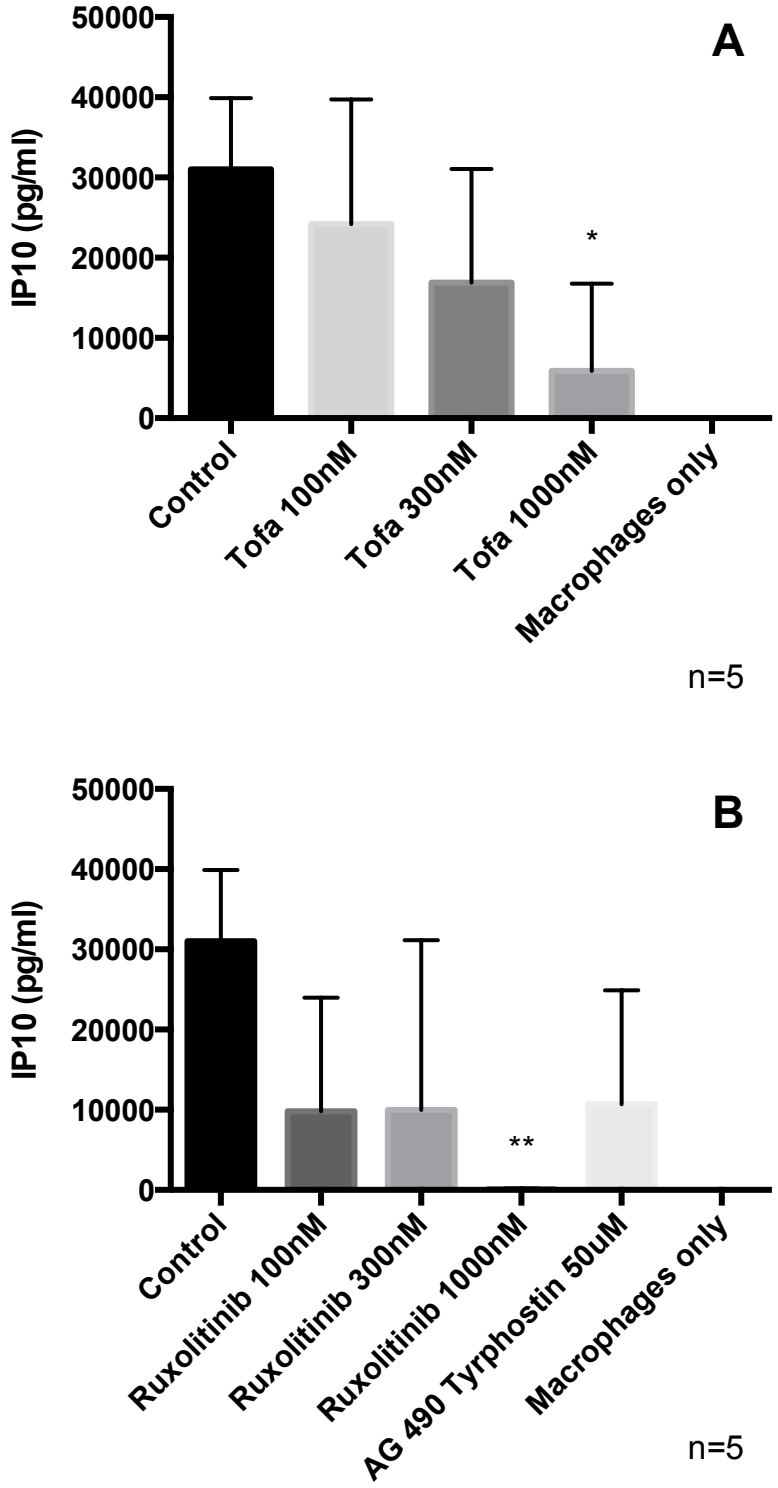
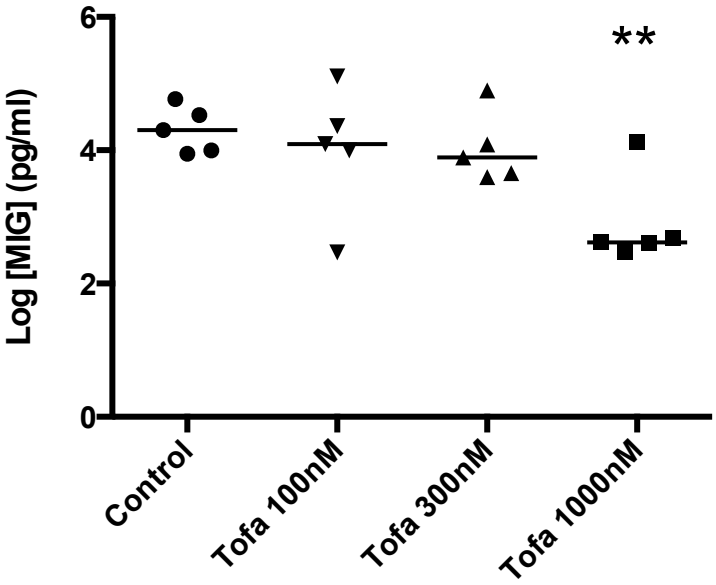


Figure 4-20 On cell contact activation by Tck, macrophages produce IP10 and this is decreased by JAK inhibition. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with inhibitors for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for IP10 performed. Tofacitinib and ruxolitinib decrease the production of IP10 in the macrophage: Tck cell contact assay and this is statistically significant.. Repeated measures one-way ANOVA with Bonferroni's multiple testing was used to calculate statistical difference of inhibitor treated conditions versus control. * = $p<0.05$, **= $p<0.01$. (n=5 biological replicates). Error bars show standard deviation.

4.16 MIG is released following cell contact activation of macrophages by Tck and is decreased by JAK inhibition



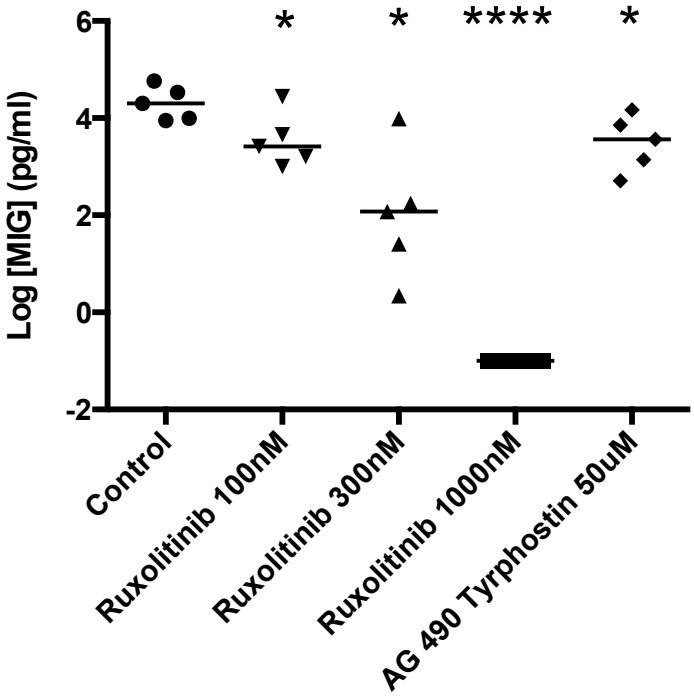


Figure 4-21 MIG is produced following cell contact activation of macrophages and Tck is decreased by JAK inhibition. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. CD4 T cells were positively selected using magnetic beads from the CD14 negative fraction and cultured in 25ml flasks at a density of $1x10^6$ cells/ml for 6 days in the presence of IL-15 (100ng/ml), IL-6 (100ng/ml) and TNFα (25ng/ml) to produce Tck. Tck were washed and added to macrophages at a concentration of 1 macrophage to 4 Tck. Both macrophages and Tck were pre-incubated with inhibitors for one hour prior to co-culture for 24 hours. Supernatants were harvested and luminex for MIG performed. Tofacitinib and ruxolitinib decrease the production of MIG in the macrophage: Tck cell contact assay. Repeated measures one-way ANOVA with Bonferroni's multiple testing was used to calculate statistical difference of inhibitor treated conditions versus control after log transforming data. * = $p<0.05$, ** = $p<0.01$, **** = $p<0.0001$. (n=5 biological replicates)

MIG (monokine induced by gamma interferon) is not released by macrophages or Tck at rest (data not shown). It is upregulated by cell contact activation (figure 4-21) and this is prevented by the use of transwell membranes demonstrating the necessity of cell contact (data not shown). There is biological variation in MIG levels following cell contact activation but the highest dose of ruxolitinib prevents production of MIG following cell contact activation.

Furthermore, although it is produced at much lower levels following LPS stimulation of macrophages than cell contact activation, it is still decreased by tofacitinib (figure 4-22).

MIG, IP10 and CXCL11 are located on chromosome 4 and form a cluster of interferon gamma inducible genes. The failure of the IFNγ assay limits the interpretation of this data. However the data suggests that JAK inhibition is disrupting IFNγ mediated signaling and this may be by two possible mechanisms: firstly, there may be decreased production of IFNγ and secondly JAK1/2 inhibition prevents signaling from the IFNγ receptor.

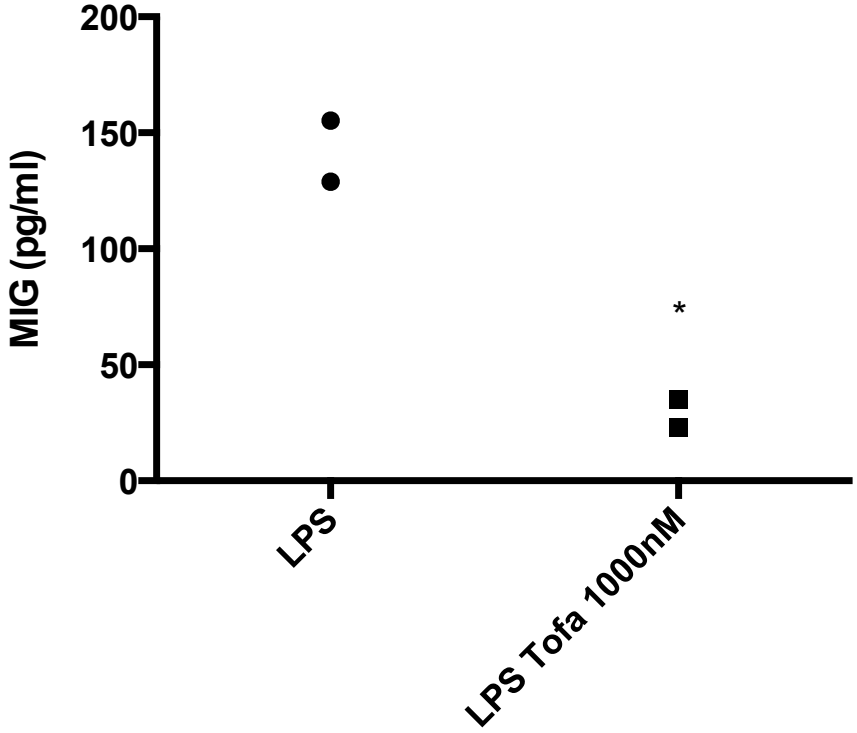


Figure 4-22 Tofacitinib decreases LPS mediated MIG production from macrophages. CD14+ monocytes were obtained following density centrifugation of healthy donor buffy coats and positive magnetic bead selection for CD14. Cells were cultured at a density of $5x10^5$ cells/ml for 3 days in complete medium in the presence of MCSF (50ng/ml) in a 96 well plate and then replated on day 3 at a density of $5x10^5$ cells/ml for a further 3 days. Macrophages were incubated with 1000nM tofacitinib or DMSO vehicle control for one hour and then stimulated with LPS (1ng/ml) for 24 hours. Supernatants were harvested and luminex for MIG performed. Data shows two biological replicates. Tofacitinib decreases LPS mediated production of MIG from macrophages. A paired t-test was performed to detect statistical difference. *= $p<0.05$.

4.17 Discussion

In this chapter I set out to answer the following questions:

1. Does JAK inhibition prevent the release of TNFα when macrophages are cell contact activated by Tck?

2. Does JAK inhibition prevent the formation of Tck?

3. Does JAK inhibition prevent LPS driven TNFα release from macrophages?

4. Does JAK inhibition prevent the release of inflammatory cytokines and chemokines that are important in the pathogenesis of RA?

The data presented here shows that JAK inhibition reduces the release of TNFα, IL-6 and IL-15, cytokines key in the pathogenesis of RA, in a macrophage and Tck cell contact activation assay. To my knowledge, others have not demonstrated this, although the effects of JAK inhibition on macrophage and T-cell activation have been studied separately(67,73,266,267). In addition, tofacitinib did not have an effect on LPS mediated TNFα release from macrophages but did alter release of interferon responsive chemokines such as MIG and IP-10 in this case.

There are limitations to the data presented herein that include the use of an in vitro culture system as opposed to a murine model of arthritis to assess the role of JAK inhibition. However murine models have been criticised as models of human disease (268)and therefore I decided to investigate this mechanism further in a human model system.

The role of cell contact activation of macrophages by activated T-cells has been investigated extensively in the past(246,269,270). Stimuli including activation of CD3 and CD28 by antibody and cytokine cocktails produce T-cells that are capable of driving pro-inflammatory cytokine production from myeloid lineage cells. Furthermore, based on immunohistochemical studies, we see that both macrophages and CD4 T-cells are present in the joint(99,100).

Furthermore, to further validate the co-culture system, T-cells derived from RA synovial fluid and synovial tissue stimulate inflammatory cytokine production from mononuclear cells derived from healthy volunteers or RA patients. Work by Brennan et al at the Kennedy Institute of Rheumatology showed that CD18, CD69 and CD49d were upregulated on cytokine activated T-cells and these markers were also highly expressed on RA synovial fluid derived T-cells(116,117). Further, antibody blockade of these surface markers reduced the amount of TNFα generated by approximately 30%.

In summary, although not perfect I believe that this assay system allowed me to explore how JAK inhibition would affect the crosstalk between macrophages and T-cells in the pathogenesis of RA.

Both macrophages and T-cells that were used in my experiments were derived from healthy buffy coats although this is in common with pervious studies. RA patient derived samples of blood and synovial fluid derived cells would be useful in the future to further

explore the mechanism. Donor to donor variability was an issue in the experiments but controlling macrophage cell numbers by re-plating partially addressed that.

In my experiments I used live macrophages and T-cells, which leads to difficulties in concluding which cell type is responsible for the secretion of a particular cytokine. Both macrophages and T-cells can secrete TNFα following stimulation and therefore limits my ability to comment on whether JAK inhibition is affecting one cell type more so than the other.

An alternative approach I would have explored had I had more time would have been to use fixed Tck and T-cell membrane preparations to cell contact activate macrophages. This approach would help tease apart the role of JAK inhibition on the macrophage in this circumstance. Furthermore, by using membrane preparations as opposed to fixed cells, I could also explore changes that occur to the transcriptome of the macrophage following contact activation and treatment with a JAK inhibitor.

I used CD4 T-cell positive selection to prepare Tck and did not subset these into Th1, Th2 or other T-cell subsets. Brennan et al (116)showed that CD4+, CD45RO+, CCR7-, CD49dhigh T-cells mediated TNFα production in the monocyte contact activation and these are derived from the effector memory T-cell pool in peripheral blood. It may be that these cells are being affected by JAK inhibition to reduce surface marker expression and this could be investigated using FACS analysis in further experiments.

Finally, although other groups have shown that IFNγ levels are increased in macrophage and Tck co-culture(271,272), I was unable to demonstrate this due to failure of the luminex bead assay for that cytokine. This cytokine may be mediating TNFα release from macrophages and therefore potential blockade of IFNγ signaling by JAK inhibitors may be responsible for my findings. Despite these limitations, I believe that my work demonstrates that JAK inhibition is able to reduce inflammatory cytokine production in the macrophage and Tck co-culture and therefore adds to current knowledge regarding the mechanism of action of this class of drug in both the innate and adaptive immune system.

Sebbag et al (255)used the macrophage and T-cell co-culture assay to demonstrate that CD3 T-cells could be activated by cytokines in an antigen independent fashion and stimulated macrophages and monocytes to produce TNFα but not IL-10 (255).

Furthermore, they also demonstrated that cell-contact was a necessary component by using transwell membranes and found that the T-cells activated macrophages in a concentration dependent fashion. My findings are in keeping with the discoveries in this paper. They also demonstrated that IL-2 alone was unable to stimulate Tck formation but IL-6 and TNFα were necessary. Also they showed that IL-15 alone or in combination with IL-6 and TNFα made functional Tck that could stimulate monocytes to produce TNFα. They also demonstrated that exogenous IFNγ and GM-CSF led to an increased production of TNFα following Tck stimulation of monocytes. Finally, this group used fixed Tck in their experiments and therefore it a reasonable assumption that the TNFα levels seen in my experiments is of macrophage origin.

McInnes et al demonstrated the role of IL-15 in T-cell migration and activation (273)and also the ability of IL-15, at concentrations that are found in RA synovial fluid, to produce Tck that were capable of cell contact activating U937 cells to produce TNFα (254). I found that macrophages do not produce IL-15 at rest but when Tck are added, significant concentrations of IL-15 are produced and this is dependent on cell contact.

Furthermore, tofacitinib reduces IL-15 production at the highest concentration but Ruxolitinib has a more significant effect. Although IL-15 itself signals through JAK1/3, from my experimental JAK1/2 inhibition may more effectively reduce IL-15. Furthermore, a proof of concept trial of IL-15 inhibition was encouraging in RA patients with 63% reaching an ACR20 response, 38% an ACR50 and 25% ACR70(61).

Wenink et al showed that abatacept (CTLA4-Ig) was able to prevent cell contact mediated TNFα release from GM-CSF macrophages (274). They did shows that IL-6 was not released in their system but was upregulated with concomitant LPS stimulation. Furthermore, IFNγ and IL-12p70 were also released in this situation. Finally they state that abatacept had no effect on TLR mediated cytokine release alone. These findings demonstrate that abatacept can inhibit inflammatory cytokine production in the co-culture assay where an antigen specific T-cell response is not required. GM-CSF macrophages may respond differently to MCSF macrophages and therefore this could be a further avenue of further investigation.

Taking account of the studies above, JAK inhibition may be able to break the vicious cycle of inflammation in the synovial joint by decreasing TNFα, IL-6 and IL-15 production when macrophages are cell contact activated by T-cells. The exact mechanism of this is

unclear but studies have been carried out of JAK inhibition in both cell types that may suggest a mechanism, which could be tested in further studies.

Ghoreschi et al carried out an extensive study of the effect of tofacitinib on mouse and human T-cells. They found that tofacitinib disrupted both gamma and non gamma chain cytokine signaling by measuring phosphostat activation by both Western blotting and phosphoFACS (73). They also showed that both Th1 and Th2 differentiation was affected by tofacitinib and suggested that IFNγ and IL-12 signaling may be affected in the former and IL-4 in the latter.

Furthermore in the murine system, in vivo, they found that LPS mediated TNFα and IL-6 production was blocked by tofacitinib. Furthermore, IL-10 was increased in this situation which is contrary to my findings. In vitro, I found that tofacitinib did not alter LPS mediated TNFα or IL-6 production from macrophages and furthermore, IL-10 was reduced with tofacitinib treatment. This could be accounted for by the oversimplified nature of single cell culture compared to an in vivo model, differences in mouse and human biology or it may be a drug concentration effect as my LPS experiments utilised 1000nM tofacitinib and in vivo drug levels are likely to be from 100-300nM. It should be noted that others have demonstrated in human macrophages that JAK inhibitions leads to a reduction in LPS mediated IL-10 production by possibly interrupting IFNβ signaling. (260)

Yarilina et al showed clearly that both tofacitinib and ruxolitinib inhibited IFNγ mediated phosphorylation of STAT1 in human macrophages (267). They also linked the TNF and JAK/STAT pathways by showing that JAK inhibition interfered with TNFα mediated stat activation. In this circumstance they showed that the production of IP-10 and CXCL11 was reduced following JAK inhibitor treatment of TNFα activated macrophages.

A previous study from the same group showed that TNFα treated macrophages set up an autocrine loop involving IFNβ and this went on to activate interferon associated genes such as IRF7(242). Therefore JAK inhibition may be able to break this loop and therefore intervene on the self-sustaining inflammatory macrophage phenotype. This mechanism may account for my findings where both tofacitinib and ruxolitinib decreased IP-10 following cell contact activation and LPS mediated stimulations of macrophages.

This group has previously shown that prolonged exposure of macrophages to TNFα activates c-Jun and NFATc and promotes osteoclastogenesis(275). JAK inhibition also

increased NFATc levels above TNFα stimulation and was associated with increased osteoclastogenesis. This was not borne out in the clinical trial of tofacitinib versus methotrexate (69) where the tofacitinib treated group had fewer erosions.

Maeshima et al studied RA patient derived CD4 T-cells and showed that tofacitinib inhibited IFNγ and IL-17 production from these cells but did not have an effect on IL-6 or IL-8 (276). The exact mechanism of this was unclear but was proposed to involve inhibition of IL-2 mediated STAT activation. This suggests that tofacitinib may be able to inhibit IFNγ production by Tck.

Finally, in a related study, Kubo et al demonstrated that there was no effect on apoptosis of immature monocyte derived dendritic cells, measured by AnnexinV and propidium iodide staining, when up to 1000nM tofacitinib was used in culture(266). They also found that immature monocyte derived dendritic cells when stimulated with LPS increase the proliferation and production of IFNγ by naïve T-cells. Tofacitinib decreased this but also increased IL-10 production in this circumstance.

Both IP-10 and MIG were increased in macrophage and T-cell co-culture and decreased by tofacitinib. Tofacitinib also prevents LPS mediated release of both chemokines from macrophages. This mechanism may be mediated by IFNβ but I did not measure this in my assays. Rosengren et al in a study of RA synovial fibroblasts demonstrated that TNFα mediated release of IP-10, MCP-1 and RANTES(74). Furthermore, IP-10 release was decreased by tofacitinib and they went on to show that a blocking antibody to IFNβ could reproduce this effect. Therefore in fibroblasts, TNFα induces IFNβ production by three hours and this is responsible for the increase in IP-10. A similar mechanism may be responsible for my results as IP-10 was also reduced following JAK inhibition of TNFα stimulated macrophages(267). It may also explain why tofacitinib decreases LPS mediated release of IP-10 and MIG if their production is dependent on type I interferon.

IP-10 was also reduced in the treatment arm of a double-blinded clinical trial of tofacitinib which included synovial biopsy at day -7 and also day 28 of treatment(277). The comparing the biopsies demonstrated that pSTAT1 and pSTAT3 were reduced following tofacitinib treatment and this correlated with disease activity. Furthermore, interferon response genes such as IP-10 were reduced with tofacitinib treatment. Also, IP-10 was detectable in patient blood samples and was statistically reduced when patients were on drug. Finally, there was no difference, following tofacitinb treatment, to the cellular

makeup or the cellularity of the synovial biopsy demonstrating that it does not cause a depletion of cell populations but mediates effects through changes to transcription.

IP-10 has been targeted therapeutically in RA in a Phase II clinical trial and showed an improvement in ACR20 response rates(278). However the response was modest and there was no effect on ACR50 or ACR70 responses. One possible reason for this is that the IP-10 receptor CXCR3 also binds MIG and I-TAC. My data shows that MIG is increased in co-culture and that there is a trend to reduction with tofacitinib treatment. This suggests that tofacitinb is acting upstream of both chemokines to reduce their production by likely disrupting interferon signaling.

This leads to a disparity in that on one hand tofacitinib acts on Janus kinases that act as signal transduction machinery for cell surface receptors that include interferons and cytokines that signal through the common gamma chain such as IL-15. However in a system which requires cell contact and in which pro inflammatory cytokine production is prevented by cell contact inhibition, tofacitinib still has an effect. Therefore JAK inhibition may be directly interfering with this.

Both tofacitinib and ruxolitinib were able to prevent the formation of functional Tck when added into T-cell culture with the cytokine cocktail of TNFα, IL-6 and IL-15. Furthermore, I have demonstrated that in this assay, levels of these cytokines are decreased suggesting that JAK inhibition may be able to interrupt the cross talk between macrophages and T-cells and therefore disrupt the inflammatory process (figure 4-22).

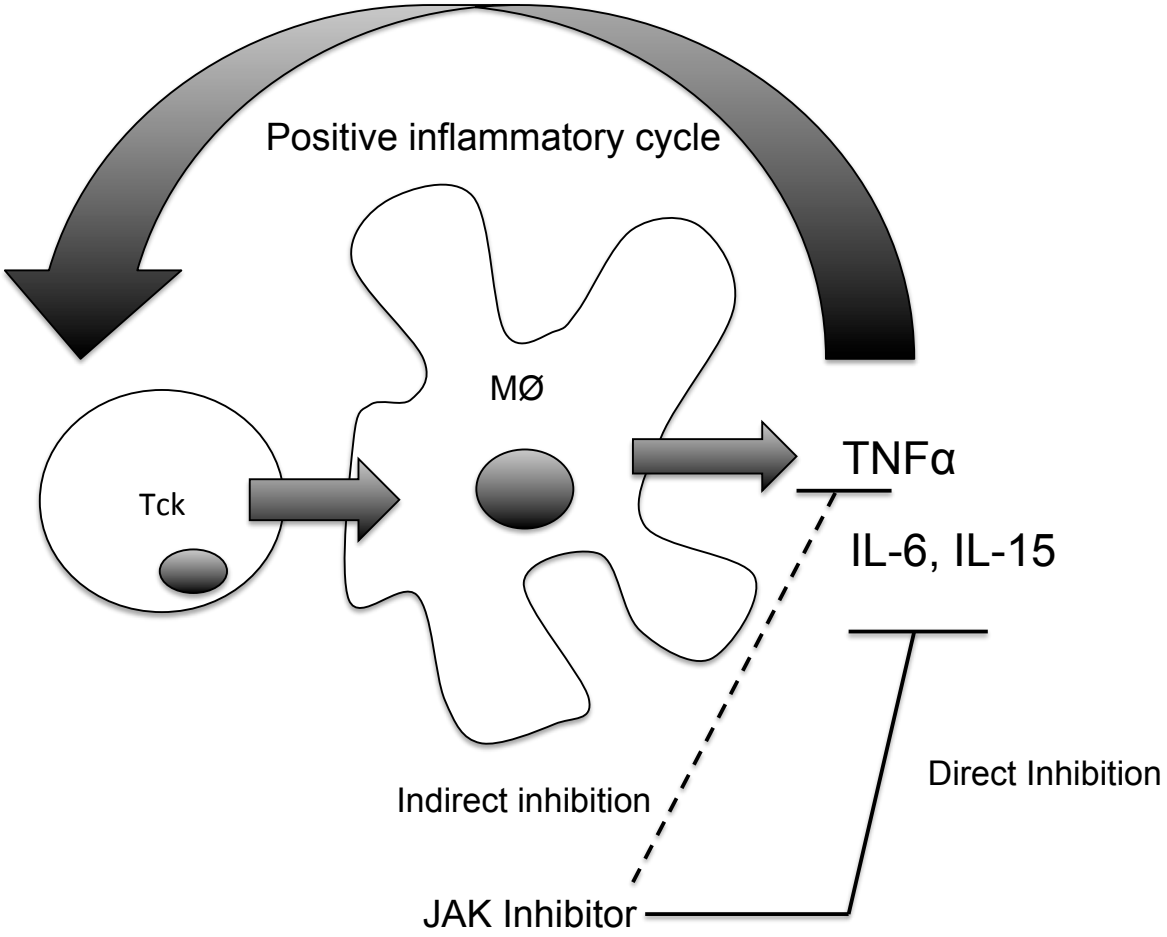


Figure 4-23 Breaking the inflammatory cycle. Cell contact activation of macrophages by T-cells results in the production of TNFα, IL-6 and IL-15 and these are reduced by JAK inhibition. Both tofacitinb and ruxolitinib decrease the levels of these cytokines and will directly affect signaling from IL-6 and IL-15. Furthermore, TNFα signaling, although not directly targeting by JAK inhibitors is disrupted in an indirect manner. Finally, production of chemokines that are associated with type I interferon such as IP-10 are also inhibited. The result of JAK inhibition, therefore, is more than inhibition of the JAK associated signaling cascades in each cell type alone.

Further investigation of the effect of JAK inhibition on the formation of Tck is required and in particular if surface expression of CD18, CD69 and CD49d is prevented if JAK inhibitors are used during Tck culture. Therefore examination of a time course investigating whether changes occur to the Tck following tofacitinib or ruxolitinib treatment would be a logical next step.

Also using fixed Tck or T-cell membrane fragments would allow me to investigate the role on macrophages alone and the latter would also facilitate the interrogation of the transcriptome of cell contact activated macrophages.

Utilising other compounds and tool compounds that are able to selectively inhibit JAK family members would increase our understanding of the whether differential inhibition has an effect on this system. In this same respect, neutralising antibodies to type I interferon, type II interferon and TNFα would help dissect primary from secondary effects.

Furthermore investigating this in patient samples would make the study translatable and I propose that we obtain peripheral blood cells from patients with RA to investigate whether this mechanism holds true in patients.

In conclusion, tofacitinib inhibits the production of pro-inflammatory cytokines which are generated in a macrophage:Tck co-culture assay, which is a model system for the chronic inflammatory process. Therefore in an established inflammatory arthritis, tofacitinib would be a useful therapeutic option given the ability to break the vicious cycle of macrophage and T-cell interaction. This has also been demonstrated clinically where tofacitinib was superior to methotrexate in improving joint signs and symptoms(69). However I have demonstrated that the effect of tofacitinib is manifest in both the T-cell and myeloid compartments and therefore JAK inhibition rather than single cytokine inhibition allows targeting of multiple cell types and therefore a better outcome for patients.

Chapter 5 Investigating whether JAK inhibition with tofacitinib inhibits synovial fluid stimulation of candidate leukocytes

Some phosphoFACs stimulation experiments were performed by Dr Moeed Akbar.

5.1 Introduction

I have demonstrated that an IL6R/JAK/STAT signature is evident in macrophages derived from RA synovial fluid. Furthermore, by using an *in vitro* macrophage: Tck cell contact activation assay, I have shown that tofacitinib and other JAK inhibitors can decrease the production of TNFα and other pro-inflammatory cytokines in a concentration dependent manner.

These cytokines and chemokines once secreted from macrophages act locally within the joint but are also secreted into synovial fluid. Synovial fluid from RA patients has been shown to contain high concentrations of IL-6, IL-21 and IL-23 (279) which can impact the activation of immune cells via JAK/STAT. Therefore I decided to determine whether tofacitinib would prevent the phospho stat stimulation of monocytes by soluble factors present in RA synovial fluid.

To achieve this I used a staining and stimulation protocol that had been developed as part of the Scottish Nested Arthritis Progression cohort (SNAP). The aim of SNAP was to take patients with RA and controls and intensely immunophenotype the peripheral blood cells and measure changes to intracellular phosphorylation of STAT1, 3 and 6 following activation with a cocktail of various stimuli such as cytokines, immunoglobulin, ionomycin, PMA and anti-CD3 and anti-CD28. I focused on STAT1, 3 and 6 because they would allow me to see signals from IFNγ (STAT1), IL-6 (STAT1 and 3) and IL-4 (STAT6)

In this chapter I set out with the following objectives:

1. Show that I can use flow cytometry to measure phosphorylation of STAT proteins following cytokine stimulation in THP-1 cells

2. To investigate whether tofacitinib can inhibit phosphorylation of STAT proteins in peripheral blood leucocytes from RA patients following stimulation with the SNAP stimulation cocktail

3. To investigate if synovial fluid from RA patients phosphorylates STAT proteins 1,3 and 6 in monocytes derived from RA patients

5.2 THP-1 cells phosphorylate STAT1 and 3 in a concentration dependent manner following stimulation with IFNγ and IL-6

To demonstrate that the SNAP staining protocol worked in my hands, I used a monocyte cell line, THP-1 to assess the effect of cytokine stimulation on specific phosphorylated STATs. THP-1 cells were stimulated with increasing concentrations of cytokine for fifteen minutes. Thereafter they were fixed and permeabilised prior to intracellular staining with phosphostat antibodies.

IFNγ signals through JAK1/2 to phosphorylate STAT 1 and we see that increasing concentrations of IFNγ results in a specific increase of phosphoSTAT1 but not phosphoSTAT3 or phosphoSTAT6 (figure 5-1).

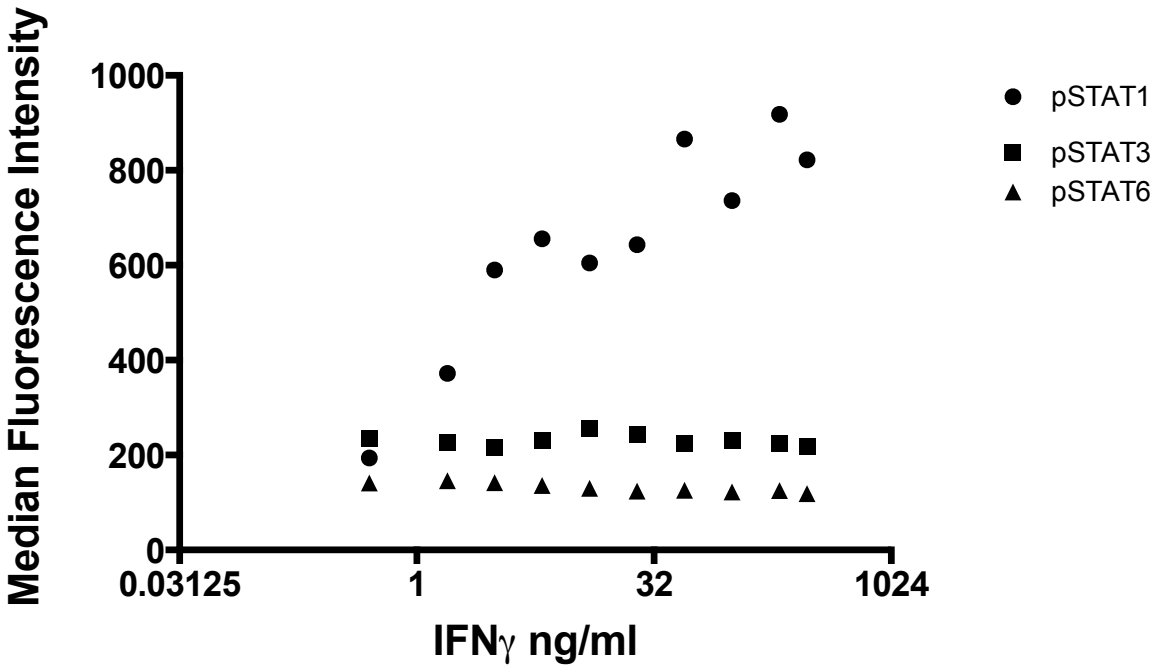


Figure 5-1 IFNγ phosphorylates STAT1 in a concentration dependent manner but has no effect on STAT3 or STAT6. Increasing concentrations of IFNγ (from 0.5ng/ml to 400ng/ml) were added to FACS tubes and 250,000 THP-1 cells added to each tube in 0.5ml of complete medium and incubated at 37C for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Phosphorylated STAT1 Mean Fluorescence Intensity increases with IFNγ stimulation in a concentration dependent manner. Phosphorylated STAT3 or 6 do not change with IFNγ stimulation. Data from one experiment.

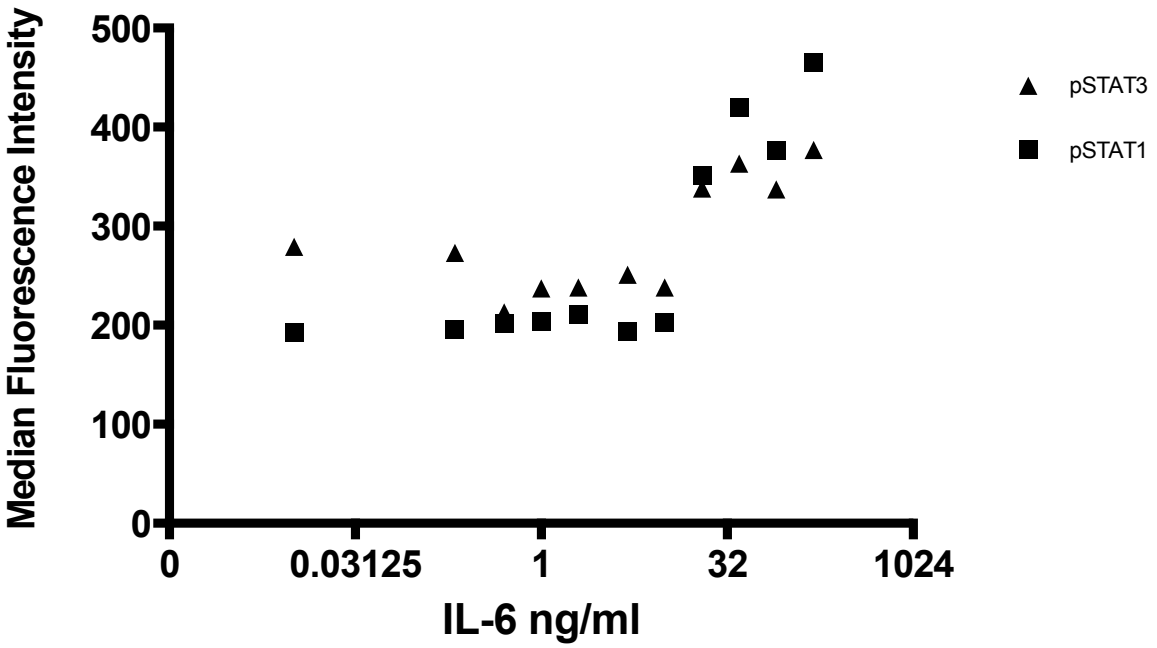


Figure 5-2 IL-6 phosphorylates STAT1 and STAT3 in a concentration dependent manner. Increasing concentrations of IL-6 (0.5ng/ml to 400ng/ml)were added to FACS tubes and 250,000 THP-1 cells added to each tube in 0.5ml of complete medium and incubated at 37C for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Phosphorylated STAT1 and phosphorylated STAT 3 Median Fluorescence Intensity increases with IL-6 stimulation in a concentration dependent manner. Phosphorylated STAT6 staining failed in this experiment. Data from one experiment.

IL-6 phosphorylates both STAT 1 and 3 in a concentration dependent manner (figure 5-2). PhosphoSTAT6 was not changed by IL-6 stimulation (data not shown). Therefore I have shown that the intracellular phosphostat staining protocol that was developed as part of the SNAP study is applicable to other cell types and I am able to perform the staining in my own hands. I therefore went forward to analyse the effect of JAK inhibition using tofacitinib on PBMC cultures that were derived from peripheral blood of patients with CCP positive RA.

5.3 Protocol of stimulation and gating strategy

Fresh blood was drawn into lithium heparin tubes from patients with CCP positive RA. PBMC were obtained by density centrifugation and cultured overnight with either vehicle control or 1000nM tofacitinib. Following culture, the cells were surface stained and stimulated for 15mins using the SNAP stimulation cocktail. This cocktail is outlined in chapter 2 but includes IFNγ, IL-6, IL-4, PMA, ionomycin and antibodies to activate CD3 and CD28. Cells were permeabilised using BD Perm Buffer III and stained with antibodies to pSTAT1, 3 and 6.

The gating strategy for both monocytes (figure 5-3) and lymphocytes (figure 5-4) is shown below:

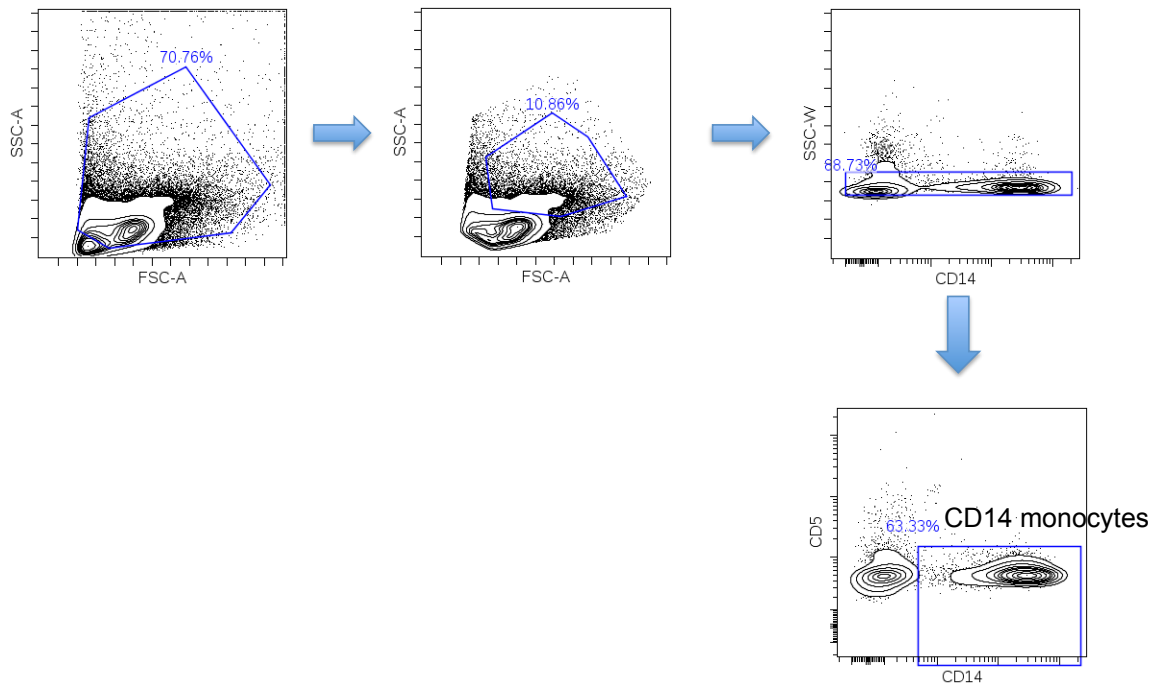


Figure 5-3 Monocyte gating strategy. Intact cells were gated from all events acquired. Monocytes were gated based on forward and side scatted characteristics. Doublets were excluded based on SSC-W and CD14+ CD5- cells were designated as monocytes.

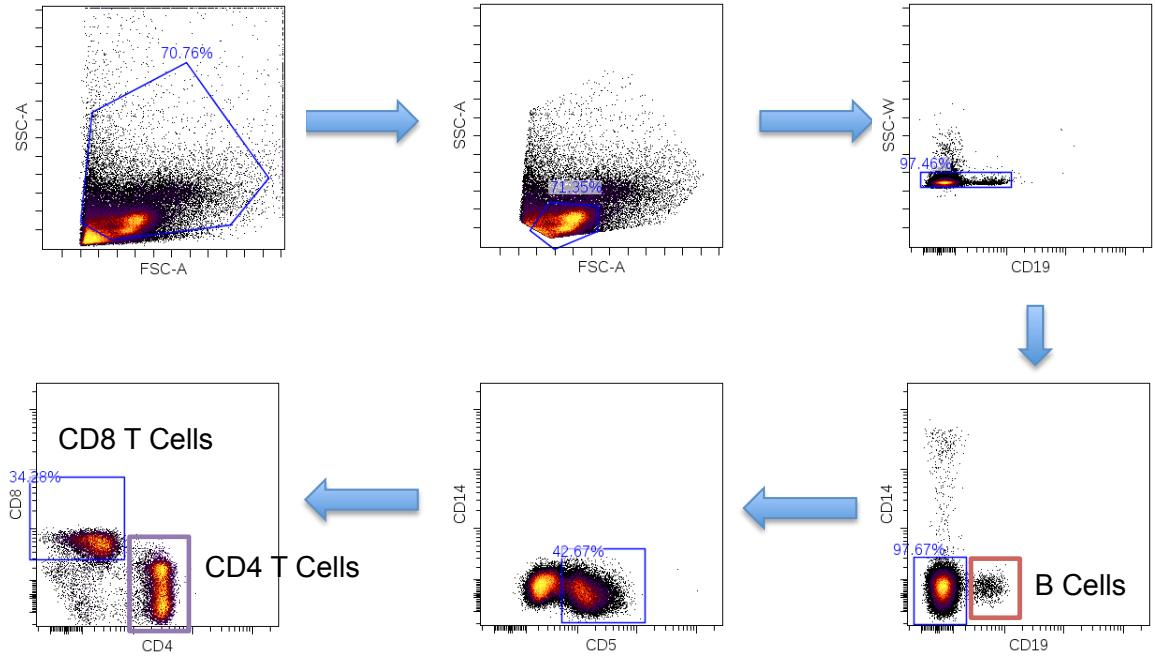


Figure 5-4 Lymphocyte gating strategy. Intact cells were gated from all events acquired. Lymphocytes were gated based on forward and side scatted characteristics. Doublets were excluded based on SSC-W. CD14-, CD19+ cells were designated B-cells. CD14- CD19- CD5 + cells were designated as T lymphocytes. CD4+ and CD8+ T lymphocytes were gated based on the last biplot.

5.4 Tofacitinib prevents STAT1 phosphorylation induced by a stimulation cocktail in CD14 Monocytes from RA patients

CD14 monocytes from RA patients phosphorylate STAT1 following stimulation with the SNAP cocktail; furthermore this is inhibited by tofacitinib (figure 5-5). Further monocytes also phosphorylate STAT6 in response to stimulation; although this is decreased by tofacitinib the result is not statistically significant. Finally monocytes do not phosphorylate STAT3 in response to the SNAP stimulation cocktail despite IL-6 being present in the stimulation cocktail.

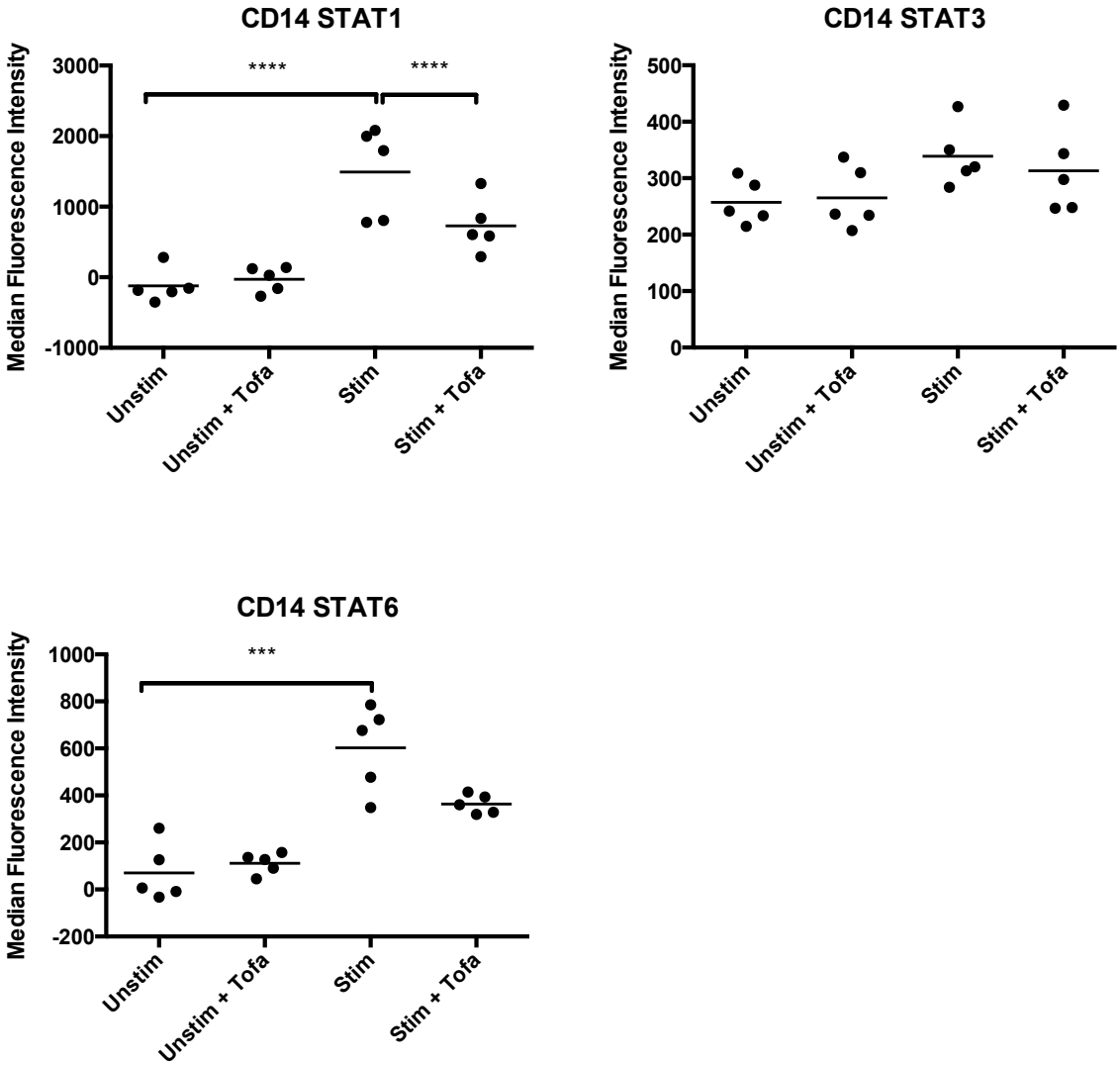


Figure 5-5 CD14 monocytes from RA patients phosphorylate STAT1 and STAT6 in response to stimulation with the SNAP stimulation cocktail and the phosphorylation of STAT1 is decreased by tofacitinib. Blood was obtained from CCP+ RA patients and PBMC obtained by density centrifugation. PBMC were cultured overnight in the presence of vehicle control or 1000nM tofacitinib, washed and simultaneously surface stained and stimulated with the SNAP cocktail for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Data was analysed using Cytobank. CD14 monocytes phosphorylate STAT1 and STAT6 following stimulation with SNAP cocktail. Tofacitinib prevents phosphorylation of STAT1 following stimulation. Two way repeated measures ANOVA was performed with Bonferroni's post-test. *** = $p<0.001$, **** = $p<0.0001$. (n=5 biological replicates)

5.5 Tofacitinib partially inhibits STAT1 and STAT6 phosphorylation induced by the SNAP stimulation cocktail in CD4 T cells from RA patients

In CD4 T cells STATs 1, 3 and 6 are phosphorylated in response to SNAP stimulation and tofacitinib partially inhibits phosphorylation of STAT1 and STAT6. Tofacitinib may inhibit STAT3 phosphorylation but this was not statistically significant and if the effect size if small, we may need more numbers to detect an effect.

Due to the constraints of surface staining within this experiment, I am unable to further characterise the CD4 T-cells and therefore do not know if this effect is global or there is differential responses within CD4 subsets.

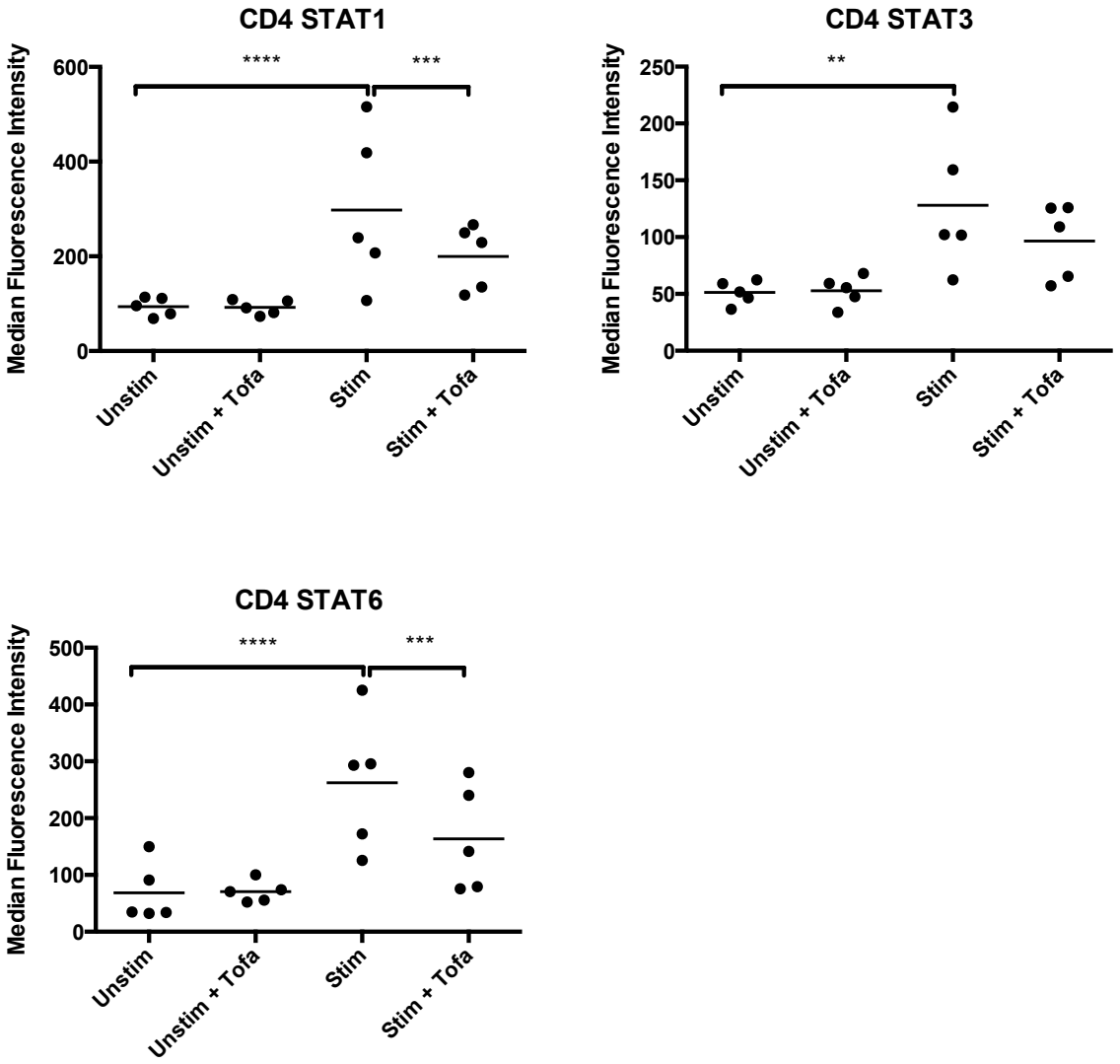


Figure 5-6 CD4 T-cells from RA patients phosphorylate STAT1, 3 and 6 in response to stimulation with the SNAP stimulation cocktail and the phosphorylation of STAT1 and STAT 6 is decreased by tofacitinib. Blood was obtained from CCP+ RA patients and PBMC obtained by density centrifugation. PBMC were cultured overnight in the presence of vehicle control or 1000nM tofacitinib, washed and simultaneously surface stained and stimulated with the SNAP cocktail for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Data was analysed using Cytobank. CD4 T-cells phosphorylate STAT1, 3 and 6 following stimulation with SNAP cocktail. Tofacitinib prevents phosphorylation of STAT1 and STAT6 following stimulation. Two way repeated measures ANOVA was performed with Bonferroni's post-test. ** = $p<0.01$, *** = $p<0.001$, **** = $p<0.0001$. (n=5 biological replicates)

5.6 Tofacitinib partially inhibits STAT1 and STAT6 phosphorylation induced by the SNAP stimulation cocktail in CD8 T cells from RA patients

CD8 T cells from RA patients also phosphorylate STAT1, 3 and 6 following stimulation (figure 5-7). Tofacitinib reverses the changes of STAT1 and STAT6 statistically but the absolute changes in MFI are small compared to the effect inCD4 T-cells or CD14 monocytes.

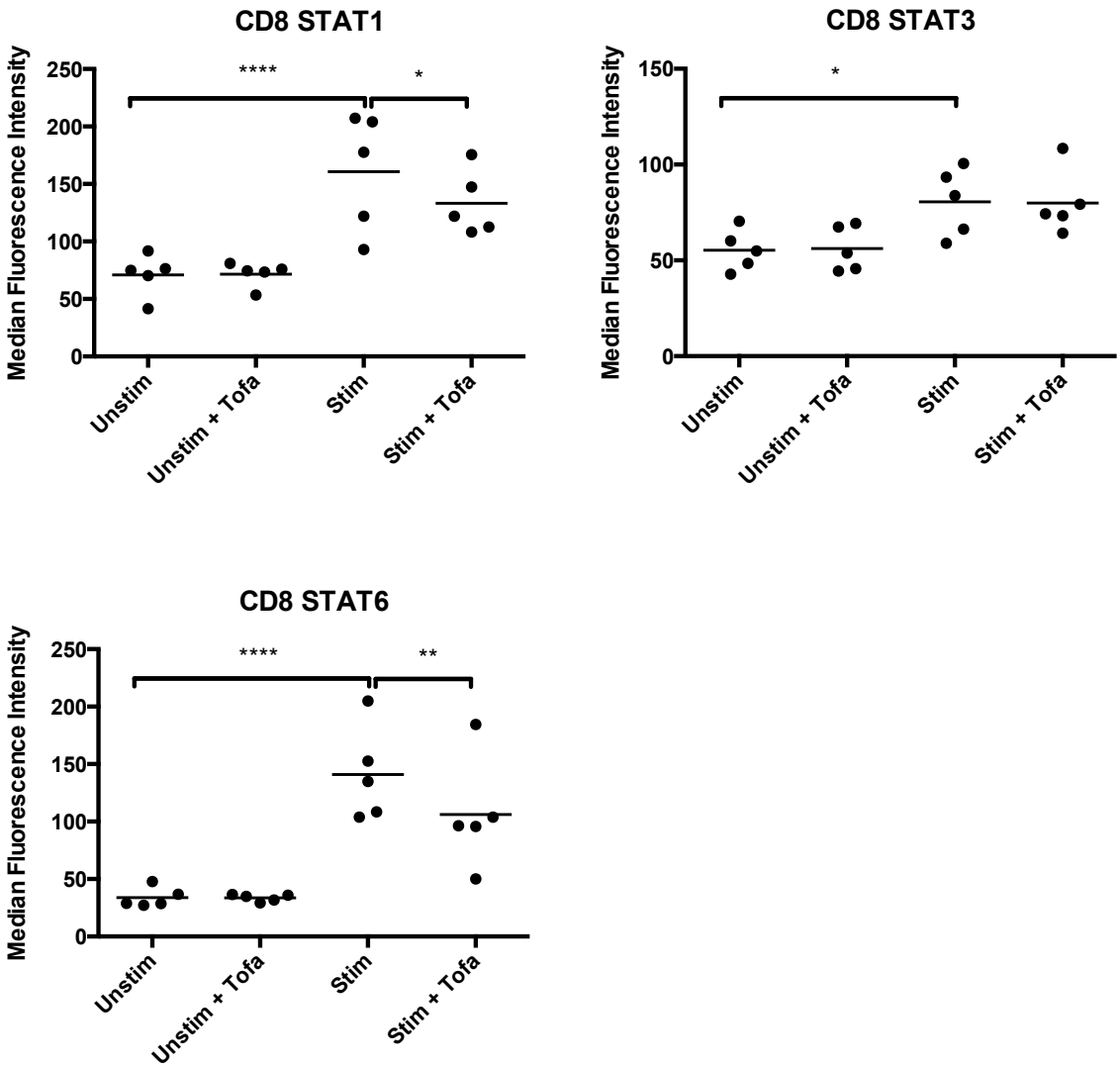


Figure 5-7 CD8 T-cells from RA patients phosphorylate STAT1, 3 and 6 in response to stimulation with the SNAP stimulation cocktail and the phosphorylation of STAT1 and STAT 6 is decreased by tofacitinib. Blood was obtained from CCP+ RA patients and PBMC obtained by density centrifugation. PBMC were cultured overnight in the presence of vehicle control or 1000nM tofacitinib, washed and simultaneously surface stained and stimulated with the SNAP cocktail for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Data was analysed using Cytobank. CD8 T-cells phosphorylate STAT1, 3 and 6 following stimulation with SNAP cocktail. Tofacitinib prevents phosphorylation of STAT1 and STAT6 following stimulation but the mean change in MFI is small. Two way repeated measures ANOVA was performed with Bonferroni's post-test. ** = $p<0.01$, *** = $p<0.001$, **** = $p<0.0001$. (n=5 biological replicates)

5.7 Tofacitinib reduces STAT1 and STAT6 phosphorylation induced by the SNAP stimulation cocktail in CD19 B cells from RA patients

The MFI of pSTAT1 and pSTAT6 is statistically increased by the SNAP stimulation cocktail and this is partially decreased by tofacitinib. Therefore B-cells from CCP+ RA patients respond to both IFNγ and IL-4 but not IL-6. Furthermore, both IFNγ and IL-4 employ JAK1 and therefore tofacitinib inhibited these as expected.

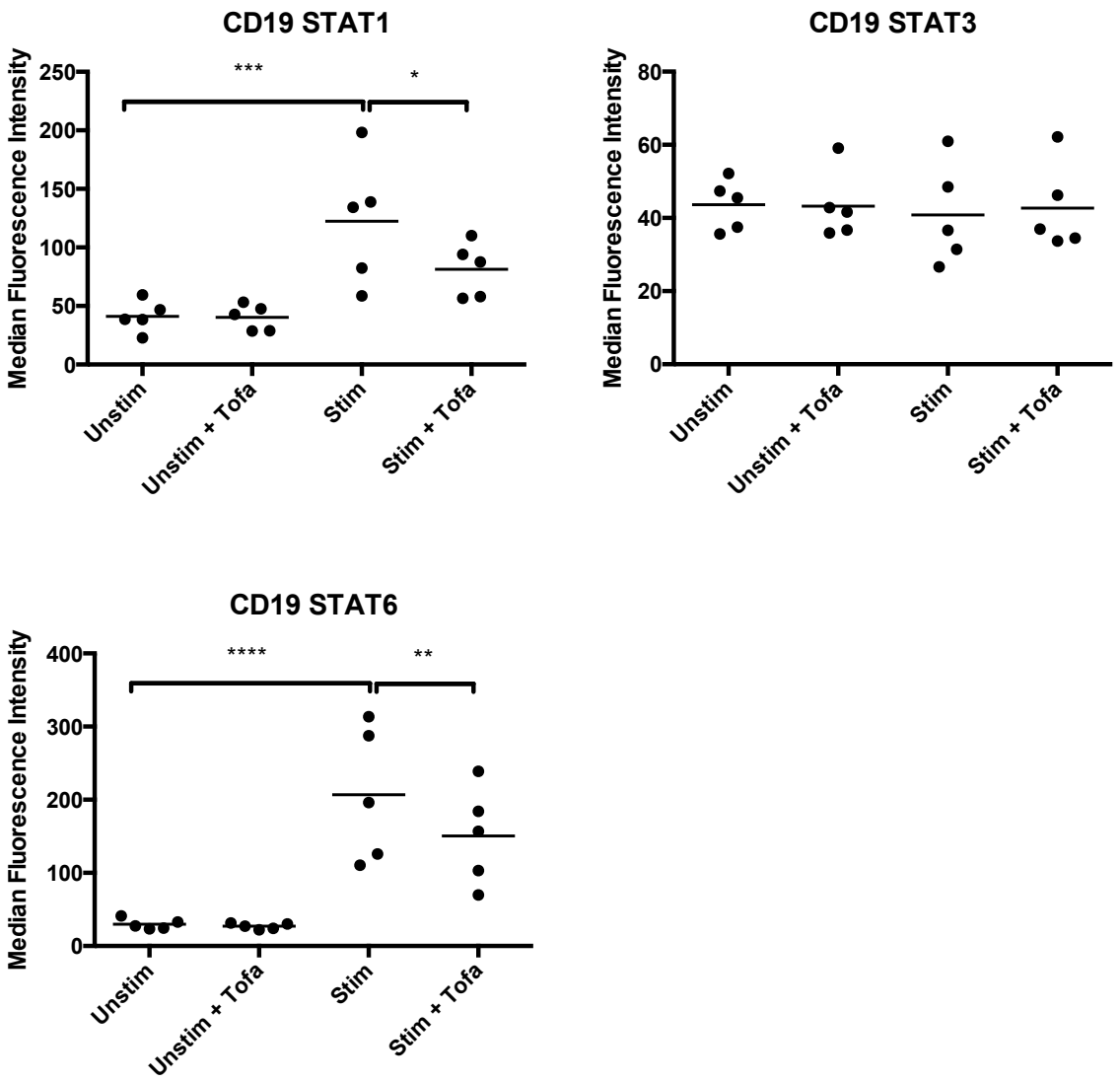


Figure 5-8 CD19 B-cells from RA patients phosphorylate STAT1 and 6 in response to stimulation with the SNAP stimulation cocktail and the phosphorylation of STAT1 and STAT 6 is decreased by tofacitinib. Blood was obtained from CCP+ RA patients and PBMC obtained by density centrifugation. PBMC were cultured overnight in the presence of vehicle control or 1000nM tofacitinib, washed and simultaneously surface stained and stimulated with the SNAP cocktail for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Data was analysed using Cytobank. CD19 B-cells phosphorylate STAT1 and 6 following stimulation with SNAP cocktail. Tofacitinib decreases phosphorylation of STAT1 and STAT6. Two way repeated measures ANOVA was performed with Bonferroni's post-test. ** = $p<0.01$, *** = $p<0.001$, **** = $p<0.0001$. (n=5 biological replicates)

In summary I have demonstrated the phosphorylation of STAT proteins in monocytes and lymphocytes from CCP+ RA patients occurs following stimulation with the SNAP cocktail. Furthermore, tofacitinib is able to decrease the phosphorylation of STATs but this is dependent on cell type. In B-cells and monocytes, STAT 1 and 6 are activated whereas in both CD4 and CD8 T-cells all the STATs are activated by stimulation. Furthermore the effect is partially reversed by tofacitinib at 1000nM but the median fluorescence intensities are not reduced to baseline values showing that phosphorylation is still occurring despite JAK inhibition. This may be due to incomplete inhibition of JAK by tofacitinib or STAT phosphorylation by mechanisms out with JAK signaling.

To achieve my final objective of determining if RA synovial fluid could phosphorylate STAT1, 3 and 6 in monocytes, I used RA synovial fluid as a stimulation condition and the SNAP cocktail as another. I pooled two samples of synovial fluid from patients with CCP+ RA that had high concentrations of antibodies against citrullinated peptides (ACPA) and also inflammatory cytokines. Table 5-1 shows the cytokines and chemokines that were found in the fluid.

Cytokine or Chemokine	Concentration (pg/ml)
IL-6	378.36
IL-12	16.16
IL1RA	129.37
sIL-2R	26.505
MCP-1/CCL2	188.18
IP-10/CXCL10	1045.58
MIG/CXCL9	118.25
IL-8/CXCL8	111.775

Table 5-1 Luminex analysis of pooled RA synovial fluid reveals high concentration of IL-6. Synovial fluid was obtained from RA patients undergoing joint aspiration. Fluid from two donors was pooled and added to complete medium to give a final synovial fluid concentration of 10%, aliquoted and frozen. Two samples were analysed by luminex analysis to determine cytokine and chemokine concentrations.

This particular pooled synovial fluid sample had high levels of IL-6 as the main pro-inflammatory cytokine.

In the particular synovial fluid there was no TNFα, IL-15 or IFNγ detectable although other groups have detected these cytokines. IL-23 also signals through STAT3 and this luminex did not contain the assay but I would explore this using an ELISA in the future.

Therefore I carried out a pilot experiment using two CCP+ RA patient samples to investigate whether synovial fluid would phosphorylate STATs in RA monocytes. I chose to investigate monocytes alone and therefore did not surface stain for lymphocytes but did employ CD5 to gate out any T-cells that express low levels of CD14.

5.8 CD14 monocytes do not phosphorylate STATs in response to stimulation with 10% RA Synovial fluid

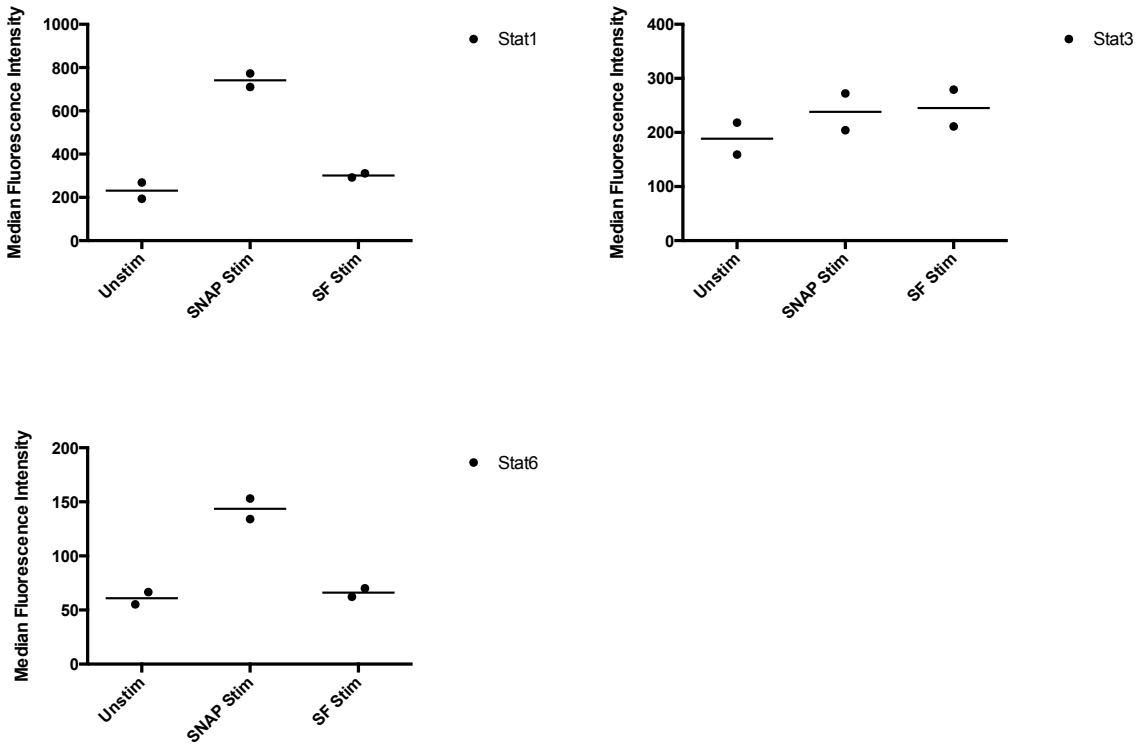


Figure 5-9 CD14 monocytes do not phosphorylate STATs following stimulation with synovial fluid. Blood was obtained from CCP+ RA patients and PBMC obtained by density centrifugation. PBMC were cultured overnight in complete medium, washed and simultaneously surface stained and stimulated with the SNAP cocktail or 10% synovial fluid for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Data was analysed using Cytobank. CD14 monocytes phosphorylate STAT1 and 6 following stimulation with SNAP cocktail. This result is in keeping with those seen in figure 5-5. 10% synovial fluid does not phosphorylate STATs. Data shows two biological replicates.

Monocytes phosphorylate STAT1 and 6 in response to the SNAP cocktail but do not phosphorylate STAT3, which is in keeping with finding from my previous experiment (figure 5-5). However, 10% synovial fluid does not phosphorylate STATS in monocytes and therefore tofacitinib would have no role in this system with regards to inhibiting the phosphorylation of STAT1, 3 and 6 by inhibiting JAKs.

I had used CD5 as a lymphocyte marker to gate out T-cell that express low levels of CD14 and therefore explored whether synovial fluid phosphorylated STATs in CD5 lymphocytes. Most of these cells will be CD4 T-cells and I have demonstrated that this cell is crucial in cell contact activation of macrophages and is affected by JAK inhibition.

5.9 CD5 lymphocytes phosphorylate STAT3 in response to synovial fluid stimulation

When I visualised the median fluorescent intensity of the CD5 population under stimulation conditions, using a histogram, I noticed that there were two peaks suggesting that two populations are present within the sample. When two populations are present we are unable to use MFI to compare groups and therefore I decided to visualise the phosphorylation of each STAT using forward scatter as per the gating strategy in figure 5-10. The STAT gate was set on the unstimulated condition having less that 0.5% cells and the percentage of cells in that gate under stimulated conditions is presented in figure 5-11.

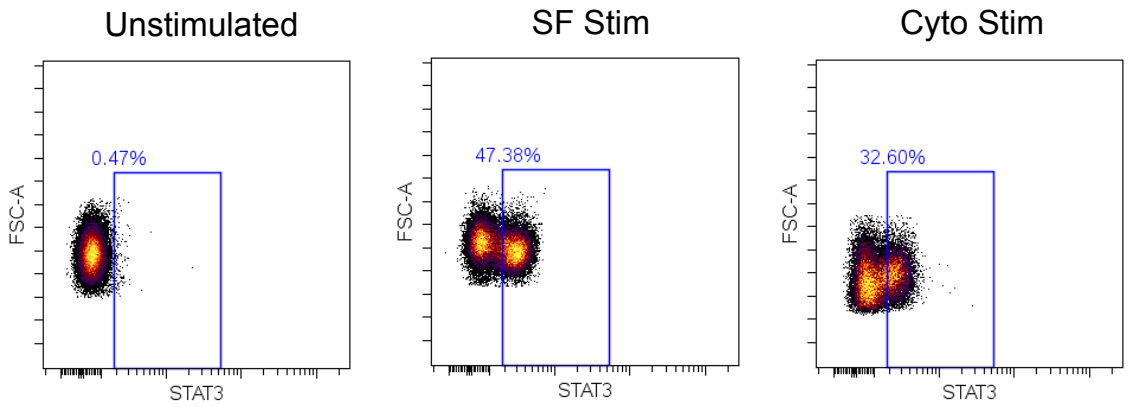


Figure 5-10 Gating strategy for phosphorylated STATs in synovial fluid stimulation pilot experiment. Blood was obtained from CCP+ RA patients and PBMC obtained by density centrifugation. PBMC were cultured overnight in complete medium, washed and simultaneously surface stained and stimulated with a cytokine cocktail of IFNγ, IL-6 and IL-4 or 10% synovial fluid for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Data was analysed using Cytobank. STAT gate was set by less 0.5% cells positive in the unstimulated condition and then applied to stimulation groups. There are two populations in both the cytokine and SF stimulated groups. This may be due to the stimulation protocol or staining protocol. Data shows two biological replicates.

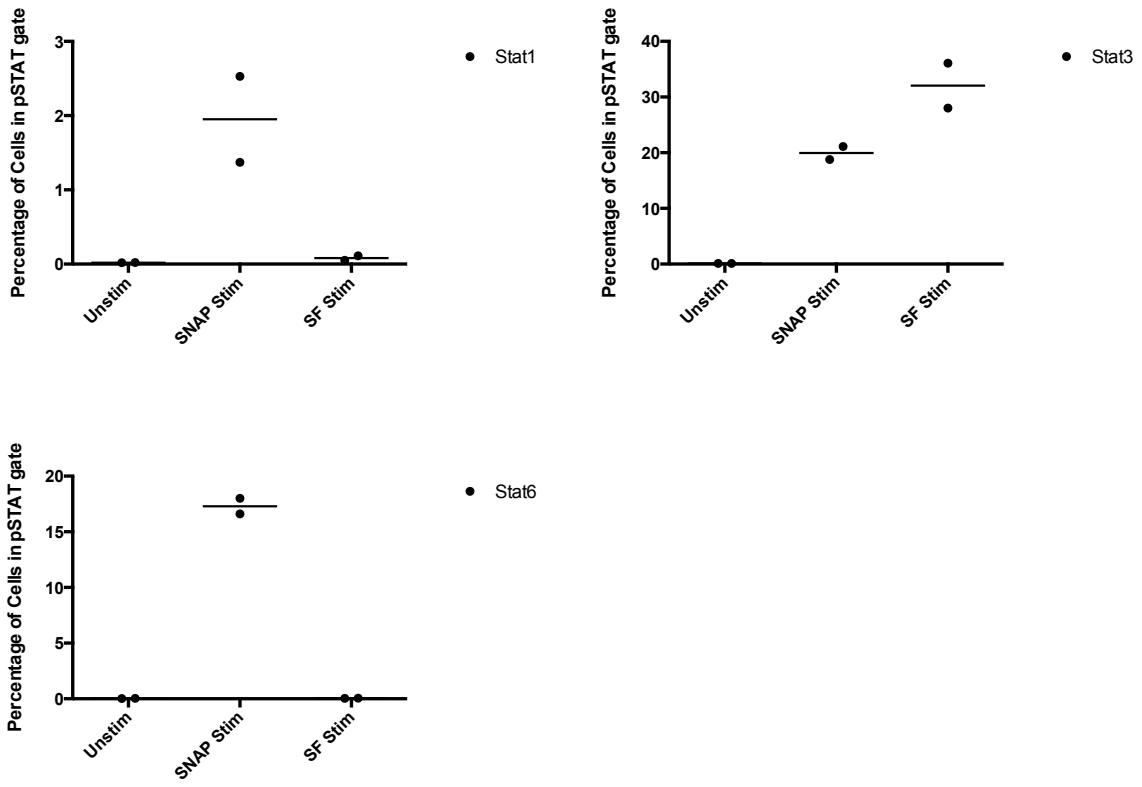


Figure 5-11 CD5 lymphocytes phosphorylate STAT3 in response to synovial fluid stimulation. Blood was obtained from CCP+ RA patients and PBMC obtained by density centrifugation. PBMC were cultured overnight in complete medium, washed and simultaneously surface stained and stimulated with the SNAP cocktail or 10% synovial fluid for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Data was analysed using Cytobank. CD5 lymphocytes from two different patients with CCP+ RA phosphorylate STAT3 in response to synovial fluid stimulation. The SNAP stimulation cocktail phosphorylates all STATs in keeping with previous results. Data shows two biological replicates.

Therefore we see that in keeping with the previously presented data, the SNAP stimulation cocktail stimulates the CD5 lymphocytes to phosphorylate STAT1, 3 and 6. However unlike the monocytes, synovial fluid from patients with CCP+ RA stimulates CD5 cells to phosphorylate STAT3 to a higher degree than the SNAP stimulation cocktail. It is also clear that there are two populations of cells ie that a percentage of the cells do not phosphorylate STAT3 within this group however the cells which do phosphorylate STAT3 makeup a reasonable percentage of the total CD5 lymphocytes.

STAT3 can be phosphorylated by many cytokines and growth factors including IL-6, GM-CSF and EGF and therefore given my previous luminex analysis of synovial fluid; it is likely that this effect is being mediated by IL-6.

To explore this finding, I collected further peripheral blood samples from patients with CCP+ RA to investigate whether this pilot finding could be replicated and also whether the phosphorylation occurred in CD4 or CD8 T-cells. Furthermore, instead of using the full SNAP stimulation cocktail, I decided to use IFNγ, IL-6 and IL-4 to specifically phosphorylate STATs and reduce secondary activation of STATs.

5.10 RA Synovial fluid stimulates CD4 T-cells and to a lesser extent CD8 T-cells to phosphorylate STAT3

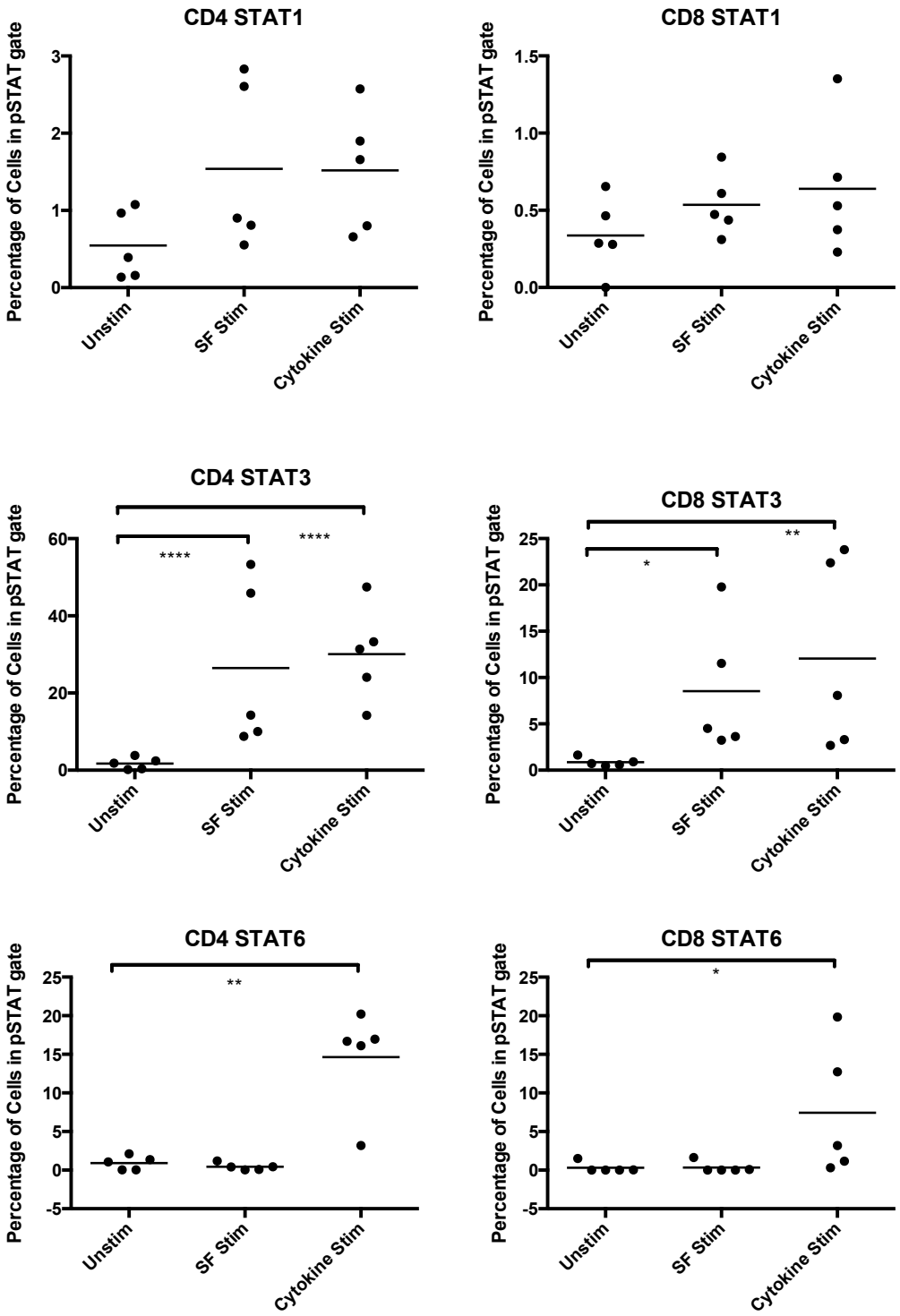


Figure 5-12 CD4 lymphocytes and to a lesser extent, CD8 lymphocytes phosphorylate STAT3 in response to synovial fluid stimulation. Blood was obtained from CCP+ RA patients and PBMC obtained by density centrifugation. PBMC were cultured overnight in complete medium, washed and simultaneously surface stained and stimulated with a cytokine cocktail or 10% synovial fluid for 15 minutes. Cells were washed in ice cold DPBS and fixed using BD Cytofix and washed three times. Cells were permeabilised using BD Perm Buffer III, washed and stained with intracellular antibodies to pSTAT1, pSTAT3 and pSTAT6. Following washing, data was acquired using an LSR II Instrument (BD). Data was analysed using Cytobank. Both CD4 and CD8 T-cells phosphorylate STAT 3 following stimulation with synovial fluid although this occurs to a higher degree in CD4 T-cells. Further, both CD4 and CD8 T-cells are stimulated by cytokine cocktail to phosphorylate STAT6 but this is not the case with synovial fluid. Two-way ANOVA with repeated measures and Bonferroni's post-test was used to calculate statistics. * = $p<0.05$, ** =$p<0.01$, *** =$p<0.001$, **** =$p<0.0001$. (n=5 biological replicates)

CD5 positive lymphocytes phosphorylate STAT3 in response to stimulation from IL-6 and also synovial fluid from CCP+ RA patients. When I investigated this further, I found both CD4 and CD8 T-cells phosphorylate STAT3 in response to synovial fluid and cytokine stimulation. Furthermore, neither CD4 nor CD8 T-cells phosphorylate STAT1 following stimulation with cytokine or synovial fluid and synovial fluid has no effect on phosphorylation of STAT6 in either cell type.

Therefore I conclude that the phosphorylation of STAT3, which is seen following synovial fluid stimulation, occurs in CD4 T-cells and to a lesser degree in CD8 T-cells. Furthermore I see a similar effect when a cytokine cocktail itself is used to stimulate these cells and IL-6 is likely to be responsible for some of this effect. Furthermore tofacitinib by inhibiting JAK1 will be able to prevent signal transduction through this pathway.

In conclusion, I have shown that RA synovial fluid does not phosphorylate STATs in monocytes and I have shown that synovial fluid is able to phosphorylate STAT3 in CD4 T-cells. Therefore, I decided to focus the next chapter of my thesis on RNA Sequencing CD4 T-cells following stimulation with RA synovial fluid and treatment with tofacitinib to explore this change on a global scale.

5.11 Discussion

I showed that STAT1 and STAT6 but not STAT3 are activated in monocytes and B cells upon stimulation with the SNAP stimulation cocktail that has a potential to stimulate all three STATs. In contrast, all three STATs were activated in CD4+ and CD8+ cells upon incubation with this cocktail. Tofacitinib was able to inhibit STAT1 but not STAT6 in monocytes; both STATs in B cells and STAT1 and STAT6 but not STAT3 in T cells. These suggest that distinct STAT pathways are engaged in the activation of myeloid and T/B cells and tofacitinib shows a cell specific effect

In addition, this small near patient study shows that soluble factors present in synovial fluid can phosphorylate STAT proteins. RA synovial fluid did not affect STATs activation in monocytes but induced STAT3 phosphorylation in CD4 and CD8 T cells. Previous data in the lab showed that RA SF activates monocytes to produces pro-inflammatory mediators and my data suggest that it is not through the JAK/STAT pathway.

There are limitations to my approach and the conclusions that I have drawn are specific to these experimental conditions.

Firstly, the number of patients involved in this study is small and furthermore I chose to limit the patients I used to those who had circulating antibodies to CCP. Most patients were on combinations of DMARDs and biologic drugs and furthermore I did not have measurements of their disease activity. These factors alone could account for the variation seen in my results, and therefore a larger study with more patients would help to clarify this.

Furthermore, I used a pool of two synovial fluids from patients with RA and therefore, this finding may be specific to these synovial fluids and therefore not generalisable to all RA synovial fluids. In addition, synovial fluids from other arthropathies may lead to the same effect and therefore investigating fluid from patients with PsA and osteoarthritis would be a logical next step.

Also although I determined that the fluid contained large concentrations of IL-6, other ligands such as G-CSF(153,156,157,280,281), Oncostatin M and Leukaemia Inhibitory Factor can phosphorylate STAT3. Finally, due to constraints with flurochrome channels, I

was unable to further characterise whether particular CD4 subsets had preferential phosphorylation. I have shown that there is a population of CD4 T-cells that do not phosphorylate STAT3 following stimulation with synovial fluid and IL-6 and so further targeted work into whether these cells are naïve, central or effector memory T-cells would be useful. In addition, measurement of surface IL-6 receptor on CD4 T-cells would help determine if the responding population has a higher expression of IL-6R. A complementary experiment would be to use tocilizumab to block IL-6 receptor prior to stimulation with synovial fluid to investigate if that prevents phosphorylation of STAT3.

Also, I did not measure total STAT protein and so during overnight culture, tofacitinib may be reducing total STAT levels thereby resulting in a lower MFI following stimulation. This is unlikely because I do not see a change in resting MFI when tofacitinib is used but this would need to be explored further.

Therefore my work extends that done by Pratt et al where they showed that transcriptomic analysis of blood CD4 T-cells from patients with early inflammatory arthritis showed changes in genes responsive to STAT3.

Pratt et al (282)used a biobank of CD4 T-cells from patients attending the early arthritis clinic to determine whether they could predict which patients with undifferentiated arthritis would progress on to rheumatoid arthritis. They employed Illumina whole genome bead arrays and machine learning in the form of support vector machines to determine a 12-gene signature that would predict progression to RA from undifferentiated arthritis.

However this metric was surpassed in the ACPA positive undifferentiated arthritis patients by their antibody status and the 12-gene metric added no further improvement over the predictive power of this. In the ACPA negative group, this metric was useful in predicting progression to RA and the genes were related to STAT3. This result was confirmed by qPCR and was also related to higher IL-6 levels in patients with ACPA negative RA and finally this was confirmed when IL-6 was used as a stimulus of healthy CD4 T cells.

Anderson et al validated a phosphoSTAT3 signature in CD4 T cells of patients with RA using a methodology similar to the one I employed(283). Furthermore they confirmed that when healthy CD4 T cells were cultured with 50ng/ml of IL-6, there was a significant fold change induction of SOCS3, BCL3 and SBNO2. The pSTAT3 signal was correlated with serum IL-6 levels and also disease activity. Finally, the ratio of basal pSTAT3 to pSTAT1

in CD4 T cells was determined as a factor that improved the Leiden risk metric (284,285)for progression to RA from undifferentiated arthritis.

Migita et al (153) showed that tofacitinib could prevent STAT phosphorylation following activation of CD4 T-cells with anti-CD3 antibody. They also demonstrated that CD4 T-cells produced less IL-4, IL-17, IL-22 and IFNγ following anti-CD3 stimulation when treated with tofacitinib.

Isomäki et al (155) showed that phosphor-STAT3 levels in peripheral blood T-cells and monocytes from patients with RA were higher than healthy controls. Furthermore, they correlated with systemic IL-6 levels that were higher than healthy controls. Furthermore, they suggested that in those patients who had circulating IL-6 levels, there was desensitisation of the IL-6 response in T-cells.

In summary I have found that synovial fluid from RA patients stimulates peripheral blood CD4 T-cells to phosphorylate STAT3. It is likely that this is due to high concentrations of IL-6 in the fluid but other ligands may be present and need to be investigated further. I was unable to characterise if a particular subset of CD4 cells was responsible for the two populations seen after stimulation but I am planning to work on this in the future.

To further explore the effect of synovial fluid stimulation on CD4 T cells I propose to perform a stimulation experiment using synovial fluid and tofacitinib as an inhibitor in a four-condition experiment. I will explore the transcriptome of the CD4 T cells using RNA sequencing to determine which genes are differentially expressed when stimulated and which are reversed by JAK1/3 inhibition to investigate whether I can discover new pathways or biomarkers of response to tofacitinib following stimulation.

Chapter 6 The effect of RA synovial fluid on gene expression of CD4 T cells

6.1 Introduction

Following my finding in chapter 5 that CD4+ T-cells from patients with CCP+ RA can be stimulated using synovial fluid from RA patients and phosphorylate STAT3, I went on to explore the downstream implications of this on gene transcription. My original aim had been to demonstrate that RA synovial fluid would stimulate RA patient derived monocytes but I was unable to show this. Therefore during my studies, I decided to change focus, driven by my findings, and investigate the RA CD4+ T-cell.

CD4+ T-cells are crucial in the pathogenesis of RA by providing help for B-cell and macrophage responses. Furthermore, in chapter 4, I demonstrated that cytokine activation of CD4+ T-cells results in large amounts of TNFα production in a macrophage cell contact activation assay. In summary, I was able to use tofacitinib to decrease the production of TNFα and other inflammatory cytokines in this assay and also prevent the formation of functional Tck.

Therefore, investigating if there are downstream transcriptional changes in CD4 T-cells following synovial fluid stimulation and whether this is prevented by tofacitinib is logical. Although the phospho-STAT3 effect is likely to be due to high concentrations of IL-6 in the RA synovial fluid, I decided to use synovial fluid instead of pure IL-6 stimulation. My reasons for doing so include: firstly, RA synovial fluid gives a more accurate simulation of the soluble factor microenvironment found in an inflamed joint than just cytokine stimulation alone and secondly, other soluble factors present in synovial fluid may be responsible for the phosphorylation of STAT3. Finally, tofacitinib in prolonged culture may affect the signaling of cytokines produced following stimulation with synovial fluid and therefore able to inhibit secondary signaling.

To investigate this, in a global fashion, I could use three different technologies: qRT-PCR based array systems such as Taqman TLDA, microarray either using an Affymetrix chip or Illumina Beadarray or RNA sequencing. RNA sequencing is an alternative method to microarrays for assessing global transcriptomic changes and has both advantages and disadvantages over array-based approaches.

A microarray relies on every gene or transcript of interest being represented on the array but in RNA-Seq, the entire transcriptome is sequenced including splice variants. RNA-Seq

can generally be done using two methods: polyadenylated (polyA) transcript selection or ribosomal depletion. The former uses the polyadenylated tag to pull down mRNA that is likely to be translated and the later involves removing ribosomal RNA (rRNA) that normally makes up 90% of the transcriptome of a cell. Briefly, once RNA is obtained it is fragmented, depleted of rRNA or polyA selected, reverse transcribed and then index libraries are built. At this stage, samples may be pooled using genetic barcodes to allow sample multiplexing.

By using ribosomal depletion as opposed to PolyA selection, you can sequence long non-coding RNA, which are novel RNA species whose role is currently being explored in the pathogenesis of inflammatory diseases(286,287). However microarray technology is more established, analysis pipelines are mature and newer arrays include variant transcripts. RNA sequencing requires complicated alignment of reads against a reference genome, if one exists, and thereafter calculation of differential gene, transcript and non-coding RNA expression using various software tools often in the command line. However, recently, Illumina developed Basespace, a cloud based method of analysing RNA sequencing data that made this novel technology accessible to the end user.

Furthermore, there is the question of depth and how often each base of a transcript is likely to or should be sequenced. ENCODE advises that for eukaryotic sequencing experiments, 33 million reads should be the minimum for polyadenylated pull down or 100-200 million for ribosomal reduction(288). The reason for the large number of reads required in the ribosomal reduction protocol is that many more reads are lost to common sequences such as mitochondrial or remaining ribosomal sequences due to protocol inefficiencies and long non-coding species take up a higher proportion of reads.

Finally, publicly available RNA-Seq data from RA patients is lacking and therefore this study will yield novel information about stimulated CD4 cells from RA patients and also the effect of tofacitinib in this system. I decided to use an Illumina NextSeq 500 platform with paired end reads with read length of 75bp and 20-25 million reads. This would allow me to determine splice variation and although this is less than that recommended by ENCODE, I was constrained by resources. Furthermore, although I would not be able to optimise the analysis pipeline as I had done for microarrays in chapter 3, Basespace would allow me to analyse this data in a time efficient manner while storing the raw data for subsequent in depth analyses.

Therefore, in this chapter I set out to seek new pathogenetic pathways and biomarkers of tofacitinib treatment response using RNA sequencing of CD4 T-cells from patients with RA. My objectives are to:

1. Demonstrate that pre-sequencing qRT-PCR of selected genes downstream of STAT3 are upregulated by stimulation with RA synovial fluid of CD4 T-cells and this is reversed by tofacitinib

2. Following this, show that ribosomal depletion and RNA Sequencing can be used to discover other genes that follow a similar pattern of upregulation with RA synovial fluid and inhbition with tofacitinib

3. To validate these genes by qRT-PCR and show that their levels are changed in keeping with the RNA Sequencing results

4. To use pathway analysis tools to predict whether networks of genes could act in concert and propose regulators of these networks that may be amenable to therapeutic targeting

6.2 Pre-sequencing validation of candidate genes show that tofacitinib prevents upregulation of SOCS3 and BCL6 following synovial fluid stimulation

Briefly, CD4 T-cells were obtained form patients with CCP+ RA by using density centrifugation of blood and positive selection of CD4 T-cells. Experimental conditions are outlined below but comprise a standard 2x2 design:

1. Vehicle Control
2. Tofacitinib Control
3. Synovial Fluid Stimulation
4. Synovial Fluid Stimulation + Tofacitinib

Cells were cultured for 24 hours in complete medium with their respective stimulants or inhibitors, harvested and RNA prepared using a column based method outlined in chapter 2. All samples were DNAse treated with the final RNA sample being split for cDNA production, described in chapter 2, and the remainder sent for quality control and RNA Sequencing. Samples were analysed with an Agilent BioAnalyser and only samples with an RNA Integrity Number (RIN) >8 were used in subsequent sequencing experiments.

In parallel, samples were stimulated with cytokines and 10% synovial fluid as per the protocol in chapter 5. Only samples that showed phosphorylation of STAT3 following synovial fluid stimulation were subsequently sent for RNA sequencing (data not shown).

In a hypothesis driven manner I decided to measure by qRT-PCR whether the levels of candidate genes downstream of STAT3 were increased by synovial fluid stimulation and if this was decreased by tofacitinib. I carried out qRT-PCR on the genes below:

- BCL2, BCL6, BCL2L1, MYC, SOCS3, 18S

The BCL family of genes, MYC and SOCS3 (156,159,289)have all been shown to be upregulated following activation of STAT3 in T-cells and 18S was chosen as a housekeeping gene.

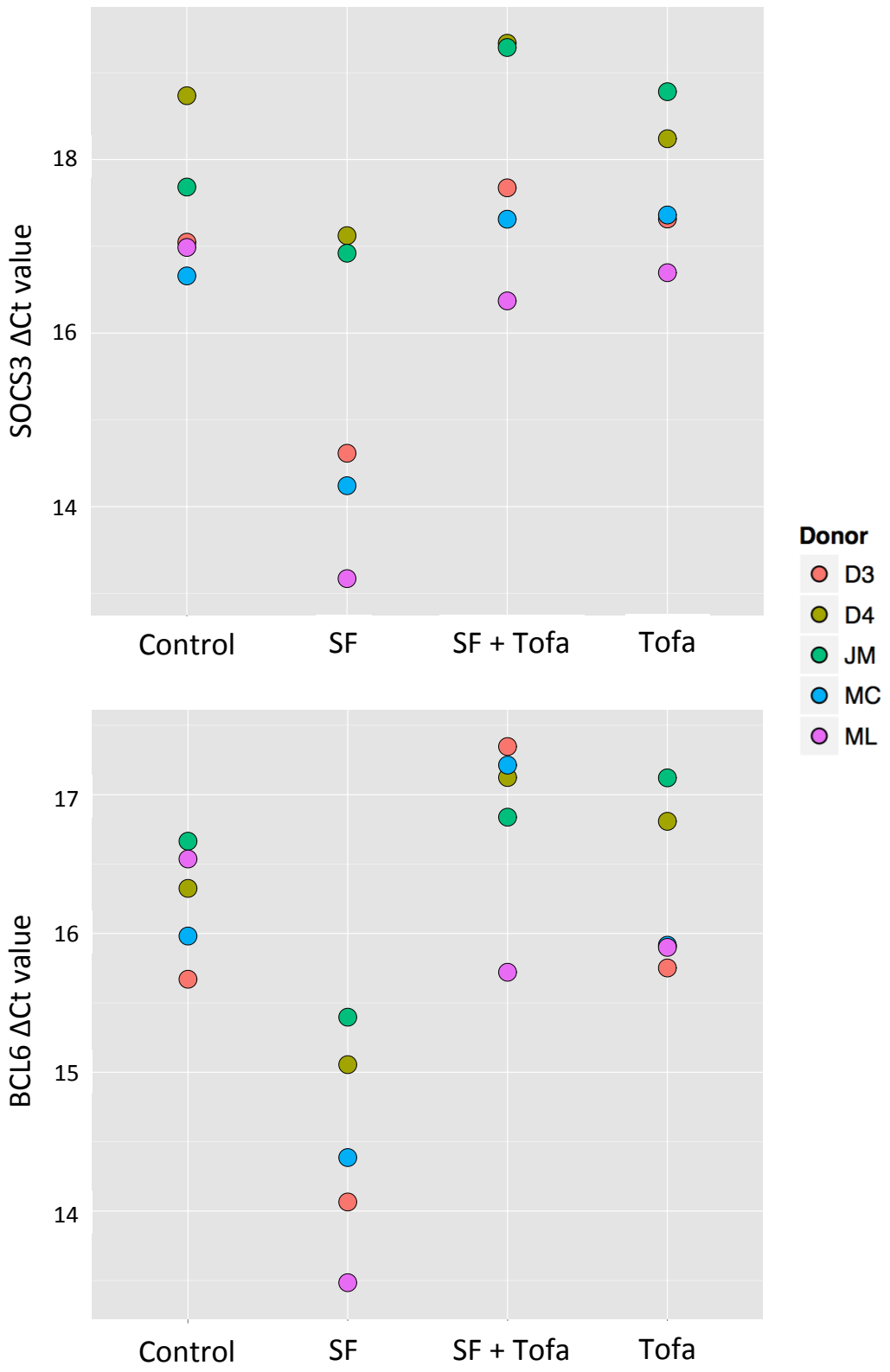


Figure 6-1: SOCS3 and BCL6 transcripts are increased following synovial fluid stimulation and this is reversed by tofacitinib. CD4 T-cells were obtained from blood of patients with CCP+ RA by density centrifugation to obtain PBMC. PBMC were then depleted of CD14+ cells using magnetic microbeads and CD4+ T-cells selected using positive selection from the CD14 negative fraction. 1.25 x 10^6 T-cells/well were cultured in 24 well plates in complete medium for 24 hours under the following conditions: Control – 0.001% DMSO, SF – 10% pooled RA synovial + 0.001% DMSO, SF + Tofa – 10% pooled RA synovial fluid + 1000nM tofacitinib and Tofa – 1000nM tofacitinib. Cells were harvested, RNA extracted and cDNA prepared as per previous protocols. qRT-PCR was performed using specific primers spanning exon junctions using SYBR green as a reporter dye. Ct values were calculated and deltaCt calculated by subtracting 18S Ct value from each sample. Both SOCS3 and BCL6 have lower deltaCt values following synovial fluid stimulation, demonstrating increased transcript, and this is reversed by tofacitinib. Figure produced in R following analysis of qRT-PCR data with HTqPCR package. (n=5 biological replicates)

Both BCL6 and SOCS3 are upregulated on synovial fluid stimulation and this is reversed with tofacitinib (figure 6-1). BCL2, BCL2L1 and MYC were not consistently changed. Therefore pre-sequencing qRT-PCR has shown that SOCS3 and BCL6 transcripts have changed in ways that I would expect following stimulation with synovial fluid and treatment with tofacitinib.

This approach means that when I move on to analysis of RNA sequencing data I can expect to see changes in these two transcripts and this will give me confidence that the sequencing itself has been successful and that the data analysis has been executed in the correct manner. In addition I believe that this method will also save resources because money will not be wasted on experiments that have failed because of technical reasons at the cell preparation stage.

I moved on to carrying out sequencing in two tranches:

1. An initial run of synovial fluid stimulated cells versus unstimulated control cells
2. A follow on experiment with the tofacitinib treated samples

I did this in case there was a technical issue with the sequencing that required sequencing to be repeated or if there was a complete failure to deliver results. In that case my plan was to analyse the problem and determine whether I had to repeat the experiment with fresh samples. On further analysis, Donor D4 was removed because although qRT-PCR results were encouraging, RNA integrity was not as high for this sample as for others, therefore I proceeded with sequencing 4 biological replicates.

6.3 Over 400 genes were differentially expressed in CD4 T-cells from RA patients treated with synovial fluid versus control cells

I chose to explore the peripheral blood CD4+ T cell transcriptome using synovial fluid as a stimulus as a surrogate for the synovial microenvironment. CD4+ T cells circulate and will enter the inflamed synovium in response to chemotactic stimuli and cellular adhesion molecule expression. Once there, they will be stimulated by cell contact events, inflammatory mediators such as cytokines and also TLR agonists such as CCP. Therefore by using peripheral CD4+ T cells from patients with RA and also synovial fluid derived from RA patients my aim was to recreate this environment *in vitro* and then determine the effect of JAK inhibition with tofacitinib in this system using RNA sequencing.

All samples from the initial run passed quality control measures within Basespace and this included FASTQC and analysis of the percentage of reads aligning to the genome and common sequences. This confirmed that ribosomal reduction had been successful and that there was good alignment to the reference genome.

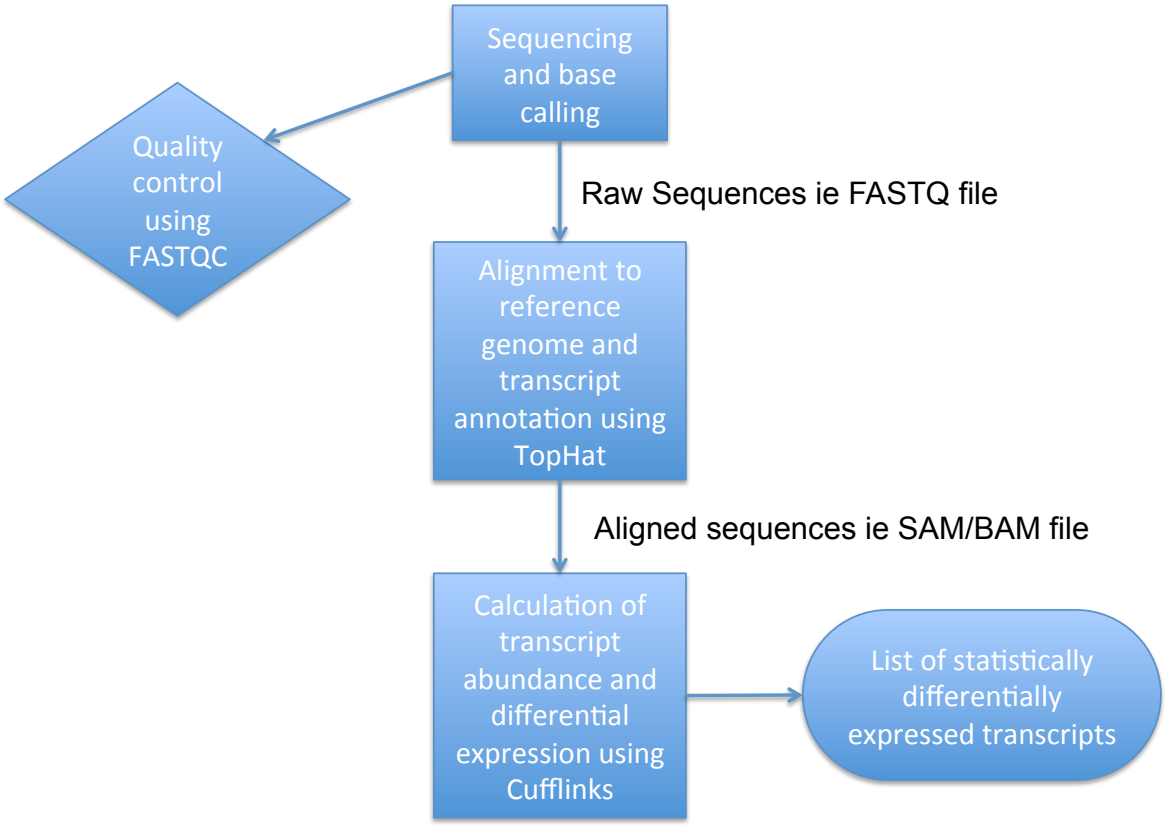


Figure 6-2 Overview of data processing in BaseSpace for RNA Sequencing data. As data is being generated on the sequencer, it simultaneously uploads to BaseSpace for quality control and analysis. Initial steps in the processing of RNA sequencing data are as for any 'omics experiment with investigation of data quality using programs such as FASTQC. Data is aligned to a reference genome using TopHat and alignment statistics interrogated for data quality control issues. Finally the aligned data is processed using Cufflinks to convert aligned reads to measures of gene abundance and also allows statistical testing of differential genes to give a final list of differentially expressed genes.

I have outlined the Tuxedo pipeline that uses Tophat(290-292) as an aligner and then Cufflinks(293) to measure differential expression (figure 6-2). Although TopHat provides information about differential transcript expression, on this first pass analysis, I decided to restrict myself to a gene level approach. Analysis transcript level data would take longer and also require more extensive validation that was outside the scope of this chapter.

For the same reasons, I decided not to discover potentially new transcripts and therefore restricted myself to genes that had been annotated in the reference genome. Although discovery of novel transcripts is exciting and necessary for novel therapeutic or biomarker development, it is also fraught with difficulties in validation. Therefore I made the decision to re-analyse the data for novel transcripts during my post-doctoral studies.

Reads were aligned to iGenome Human Genome hg19 (GRCh37) as a reference genome as this is available within BaseSpace. As part of follow-on work I would re-align this data to GRCh38, which is the latest genome to be released by the Genome Research Consortium with up to date annotations.

Following differential expression, a multiple testing corrected gene list was obtained with any gene with a corrected p value of <0.05 being deemed statistically different. Within BaseSpace, Cufflinks is only able to compare two conditions and therefore I performed differential gene expression analysis between the synovial fluid stimulated cells and control.

FPKM (fragments per kilobase of transcript per million mapped fragments) corresponds to the abundance of a transcript based on paired end reads mapping to that transcript and normalised for the length of the transcript and also depth of sequencing. Values differ between genes and there is no satisfactory value of how many FPKM amount to a transcript per cell but the value represents an estimate of gene level expression and therefore FPKM values can be compared in the same way as deltaCt values in qPCR.

6.4 Differential gene expression analysis of RNA Sequencing confirmed that SOCS3 and BCL6 gene levels are increased by synovial fluid stimulation

Complete lists of the differentially expressed genes are included in the appendix. The list includes both coding and non-coding species and is determined on a gene as opposed to a transcript level. Furthermore all genes in the list have been multiple testing corrected and have a p value <0.05.

Candidate genes that have been differentially expressed are shown in Table 6-1. I confirmed that SOCS3 and BCL-6 were upregulated in CD4 T-cells after synovial fluid stimulation. Furthermore, MYC levels were also statistically increased in the RNA sequencing data although this did not seem apparent on visual inspection of the pre-sequencing qRT-PCR data.

In addition genes associated with T cell activation such as CD69 were upregulated following synovial fluid stimulation demonstrating that synovial fluid is stimulating these cells. Also JAK3 and STAT3 were also upregulated following stimulation although this was not apparent in other JAK/STAT members

TNFSF8 is upregulated by synovial fluid and can stimulate T-cell proliferation(294). Also TNFAIP3 is also increased and this is a negative regulator of the TNF pathway and inhibits NF-κB activation(295).

Therefore I have demonstrated that RNA sequencing with ribosomal depletion can be used to measure differential gene and long-non-coding RNA expression in CD4 T-cells stimulated with synovial fluid. Furthermore, the data from RNA sequencing has confirmed findings that were suggested by the pre-sequencing qRT-PCR of SOCS3 and BCL6.

In conclusion, this gave me confidence that I should proceed with sequencing the tofacitinib treated samples and so narrow down the transcripts that are differentially expressed to those that are affected by JAK inhibition.

Gene	log2(control FPKM)	log2(SF FPKM)	log2(Ratio)	q Value
T Cell activation				
CD69	4.13	5.9	1.77	0.003
JAK/STAT signalling				
JAK3	4.64	5.53	0.89	0.003
STAT3	5.57	6.22	0.65	0.019
SOCS3	2.54	4.95	2.41	0.003
Transcription Factors				
BCL6	2.37	3.91	1.54	0.003
MYC	5.45	6.41	0.96	0.003
TNF related				
TNFSF8	4.2	6.96	2.76	0.003
TNFRSF25	7.71	6.52	-1.19	0.003
TNFAIP3	6.26	7.37	1.11	0.003

Table 6-1 Candidate genes including those from pre-sequencing qRT-PCR that are differentially expressed between synovial fluid treated versus control CD4+ T-cells from RA patients.

6.5 Over 100 genes are differentially expressed in CD4 T-cells treated with RA synovial fluid versus synovial fluid and tofacitinib treated cells

Sequencing of the tofacitinib treated samples had to be repeated because there was a failure of ribosomal depletion stage that resulted in totalRNA being sequenced 90% of which was ribosomal. Therefore the sequencing was repeated using a lower amount of input RNA and an alternative depletion kit was used, one that would deplete ribosomal and mitochondrial sequences.

The newer kit was used so that more reads would sequence the coding and non-coding sequences as opposed to remaining common sequences such as mitochondrial RNA. However the implications for this experiment are that in the future I will not be able to re-analyse the sequence for mitochondrial transcripts.

Furthermore, although a smaller volume of input RNA was used in sequencing, some samples did not have the required minimum amount and therefore this part of the sequencing experiment is subjected to a significant amount of technical noise because conditions are compared across sequencing runs as opposed to within them.

I proceeded to analyse the RNA sequencing data using the Tuxedo pipeline as previously and discovered that over 100 genes were differentially expressed with the majority being expressed at a higher level in the synovial fluid and tofacitinib treated cells compared to synovial fluid along alone. A list of the genes that were differentially expressed is included in the appendix.

SOCS3 and BCL6 were expressed at levels in keeping with pre-sequencing qRT-PCR (table 6-2).

Gene	log2(SF FPKM)	log2(SF + Tofa FPKM)	log2(Ratio)	q Value
JAK/STAT signalling				
SOCS3	4.98	2.45	-2.53	0.014
Transcription Factor				
BCL6	3.94	1.61	-2.33	0.014
TNF related				
TNFRSF25	6.55	7.59	1.04	0.04

Table 6-2 Candidate genes including those from pre-sequencing qRT-PCR that are differentially expressed between synovial fluid treated versus synovial fluid and tofacitinib treated CD4+ T-cells from RA patients.

At this stage because there had been problems with RNA sequencing of the second tranche of samples, I decided to determine a gene list that I could go on to validate using qRT-PCR using a Taqman Low Density Array (TLDA). I had prepared cDNA prior to RNA sequencing and therefore it would not be subject to the same batch effect as was present for the two sequencing tranches. Furthermore, a significant number of genes were differentially expressed in the synovial fluid versus synovial fluid and tofacitinib treated groups if I removed outlier samples that corresponded to samples that had less input RNA (data not shown).

I decided to use R and a bioinformatics package called CummeRbund to further analyse the data for two reasons: firstly based on a "candidate" gene expression profile such as SOCS3, I could ask the program to find other genes that followed a similar pattern and secondly it would allow me to visualise the raw expression or FPKM values. This would therefore allow me to achieve my objective of discovering other genes that followed a similar pattern of upregulation with synovial fluid and inhibition with tofacitinib treatment in this circumstance.

6.6 CummeRbund reveals genes that have a similar expression profile to SOCS3 and BCL6

Both BCL6 and SOCS3 showed the same predicted pattern in RNA sequencing as was determined on pre-sequencing qRT-PCR when visualised using CummeRbund(293). This demonstrates that both the wet lab preparation and dry lab data analysis have been successful because I have confirmed the same pattern using two different methods: qRT-PCR and RNA sequencing.

Figure 6-3 has been produced with cummeRbund and show individual replicate FPKM values along with a mean value and error bars which are based on the depth of the sequencing library and not just the standard deviation of the absolute FPKM values. I had to concatenate two datasets to produce the CummeRbund dataset and therefore occasionally only three FPKM replicate values are plotted as opposed to four. This does not detract from the validity of visualisation because the genes are differentially expressed on statistical testing. I therefore went on to use the "findSimilar" function in the CummeRbund package which allows you to find other genes within the data set that follow a particular pattern such as upregulated by synovial fluid and decreased with tofacitinib.

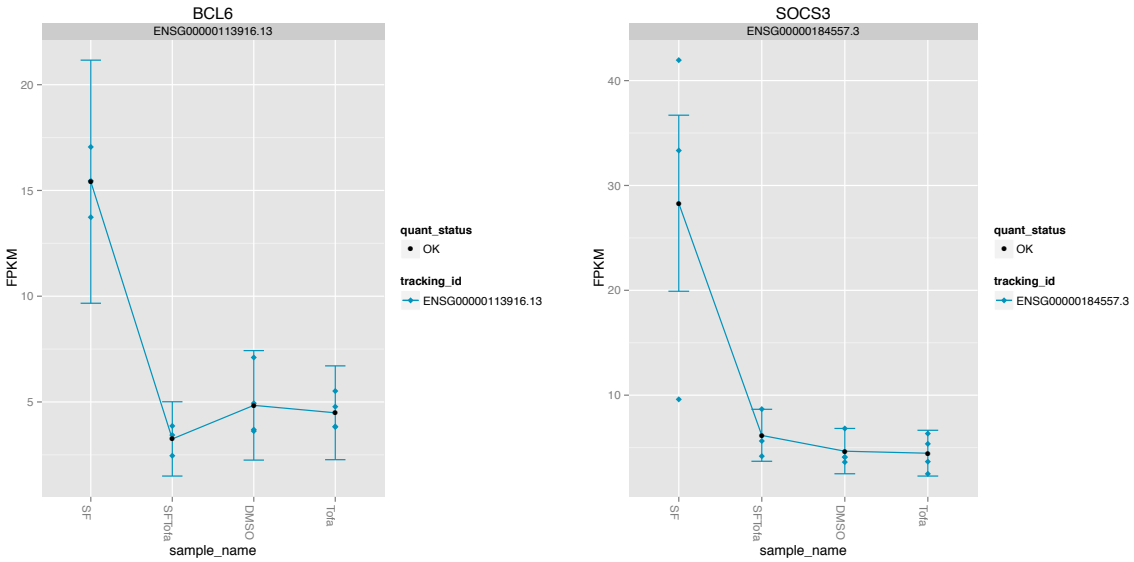


Figure 6-3 Plots of raw FPKM values for the BCL6 and SOCS3 genes across different conditions in CD4 T-cells. CD4 T-cells were isolated from the peripheral blood of patients with CCP+ RA by density centrifugation and positive selection using CD4 magnetic microbeads. 1.25 x10^6 per well were cultured in complete medium for 24 hours in a 24 well plate in four conditions: SF – 10% synovial fluid from RA patients, SFTofa – 10% synovial fluid from RA patients and 1000nM tofacitinib, DMSO – vehicle control condition, Tofa – 1000nM tofacitinib only. Cells were harvested, RNA extracted, quality controlled and sent for paired end ribosomal reduction RNA sequencing of length 75bp with 20-25M total reads. Raw data was aligned using TopHat and differential gene expression determined using Cuffdiff. FPKM plots were generated in CummeRbund following concatenation of two data sets. Error bars denote gene variance that is based on a combination of raw FPKM values and also sequencing depth. As demonstrated by differential gene expression analysis within Cufflinks, BCL6 and SOCS3 are upregulated by RA synovial fluid and this is reversed by tofacitinib treatment. (N=4 biological replicates)

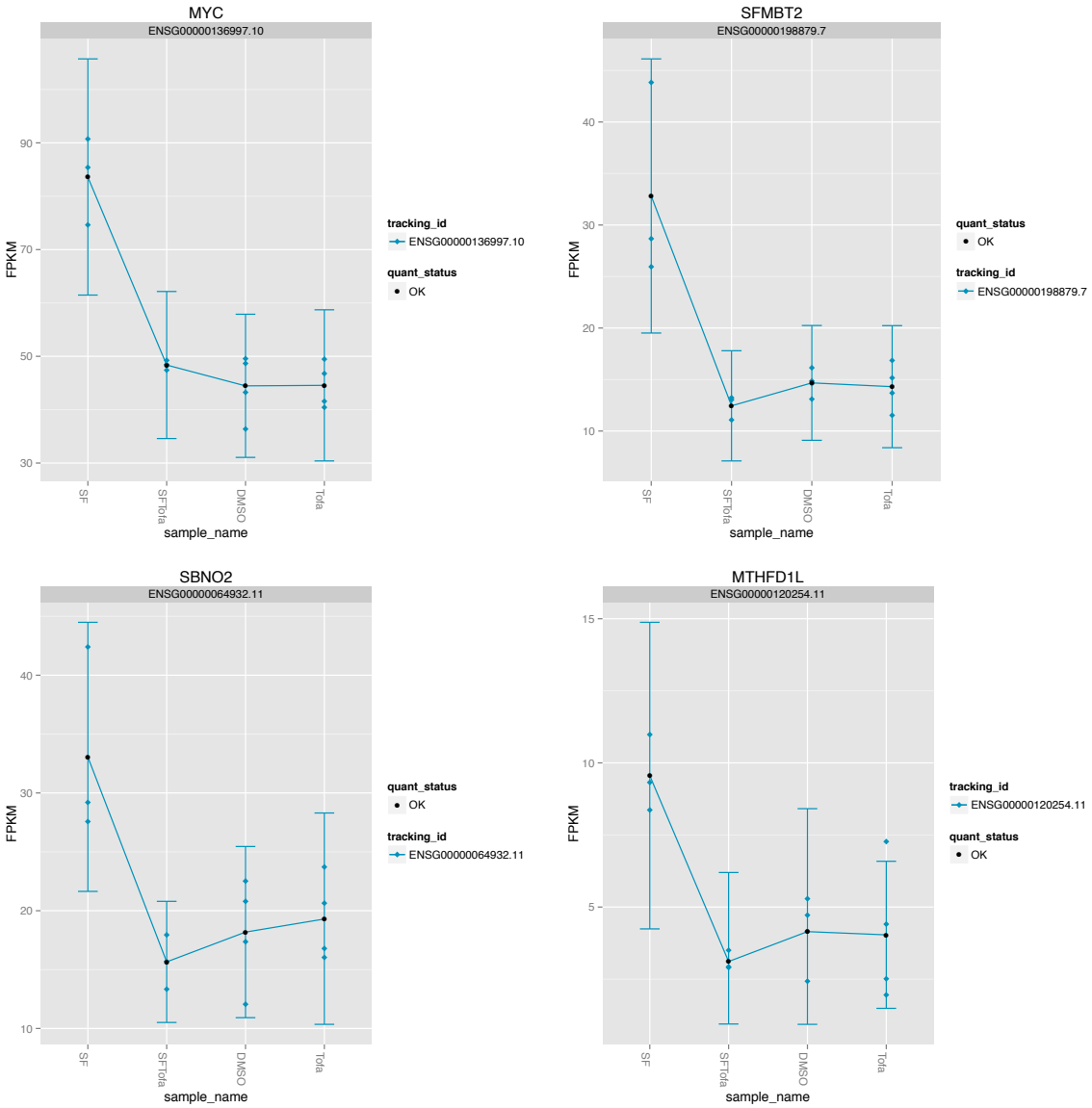


Figure 6-4 The findSimilar function in CummeRbund reveals genes that have a similar expression profile to BCL6 and SOCS3. CD4 T-cells were isolated from the peripheral blood of patients with CCP+ RA by density centrifugation and positive selection using CD4 magnetic microbeads. 1.25 x10^6 per well were cultured in complete medium for 24 hours in a 24 well plate in four conditions: SF – 10% synovial fluid from RA patients, SFTofa – 10% synovial fluid from RA patients and 1000nM tofacitinib, DMSO – vehicle control condition, Tofa – 1000nM tofacitinib only. Cells were harvested, RNA extracted, quality controlled and sent for paired end ribosomal reduction RNA sequencing of length 75bp with 20-25M total reads. Raw data was aligned using TopHat and differential gene expression determined using Cuffdiff. FPKM plots were generated in CummeRbund following concatenation of two data sets. Error bars denote gene variance that is based on a combination of raw FPKM values and also sequencing depth. The four genes above demonstrate similar expression profiles to that of SOCS3 and BCL6. (N=4 biological replicates)

MYC, SFMBT2, SBNO2 and MTHFD1L were shown to have similar expression profiles to BCL6 and SOCS3. Furthermore, each of these genes was statistically differentially expressed between synovial fluid treated CD4 T-cells and control cells. However, they did not achieve statistical significance between the synovial fluid treated cells and synovial fluid and tofacitinib treated cells in the RNA sequencing data although this may be due to technical aspects previously alluded to during the sequencing process.

As well as mRNA, ribosomal depletion allows the detection and quantification of long-non-coding RNA (lncRNA). The length of lncRNA means that coverage is less than the coding genome and therefore their FPKM levels are also less than other coding genes. Certain lncRNA have been annotated such as MIAT and others are still noted by an accession number.

The validation of lncRNA is complicated as some of them are made of one exon and therefore cannot be distinguished from genomic DNA when cDNA is prepared from RNA. In this situation a non-reverse transcribed control is often used but I was unable to do this because of lack of sample. Therefore I decided to validate by q-RT-PCR genes where probes spanned exons and so decided to focus on coding genes.

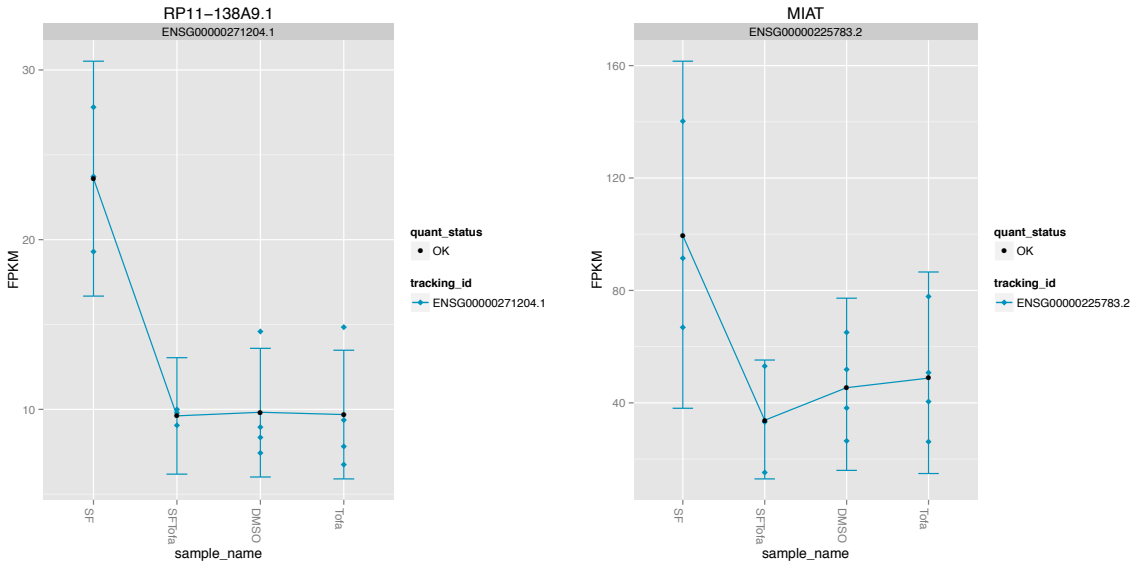


Figure 6-5 Long non coding RNA are differentially expressed between CD4 T-cells stimulated with RA synovial fluid and those treated with tofacitinib. CD4 T-cells were isolated from the peripheral blood of patients with CCP+ RA by density centrifugation and positive selection using CD4 magnetic microbeads. 1.25 x10^6 per well were cultured in complete medium for 24 hours in a 24 well plate in four conditions: SF – 10% synovial fluid from RA patients, SFTofa – 10% synovial fluid from RA patients and 1000nM tofacitinib, DMSO – vehicle control condition, Tofa – 1000nM tofacitinib only. Cells were harvested, RNA extracted, quality controlled and sent for paired end ribosomal reduction RNA sequencing of length 75bp with 20-25M total reads. Raw data was aligned using TopHat and differential gene expression determined using Cuffdiff. FPKM plots were generated in CummeRbund following concatenation of two data sets. Error bars denote gene variance that is based on a combination of raw FPKM values and also sequencing depth. Long non-coding RNA are also differentially expressed although FPKM values are low. (N=4 biological replicates)

6.7 CD69 and other genes are increased following synovial fluid stimulation and do not change with tofacitinib treatment

Genes such as CD69, a marker of lymphocyte activation, are low in control conditions but increase with synovial fluid treatment. Furthermore, these genes do not respond to treatment with tofacitinib suggesting that their expression is controlled by mechanisms out with the JAK/STAT pathway or that tofacitinib does not inhibit the appropriate JAK member such as JAK2 to a high enough degree to alter their expression level.

Many other genes followed a similar expression pattern to CD69 and these include transcription factors such as ZBTB16, proteins involved with calcineurin inhibition such as FKBP5 and also proteins with a role in redox balance such as TXNIP. Each of these genes was statistically differentially expressed between synovial fluid treated CD4 T-cells versus control cells.

In conclusion, I have demonstrated that RNA sequencing can be used to discover genes that show patterns of expression that are similar to candidate genes. Furthermore, by using ribosomal depletion I have shown that long non-coding RNA are upregulated by synovial fluid treatment and that this is reversed by tofacitinib.

Therefore, based on the differential gene expression results from Cuffdiff, the visual inspection of CummeRbund plots and also exploration of the differentially expressed genes using Ingenuity (data not shown), I decided to validate a list of genes using qRT-PCR. I designed a TLDA plate in conjunction with colleagues at Life Technologies to validate genes that had been differentially expressed. Furthermore, cDNA was created prior to RNA sequencing and therefore would not be subject to batch and technical effects experience during RNA sequencing.

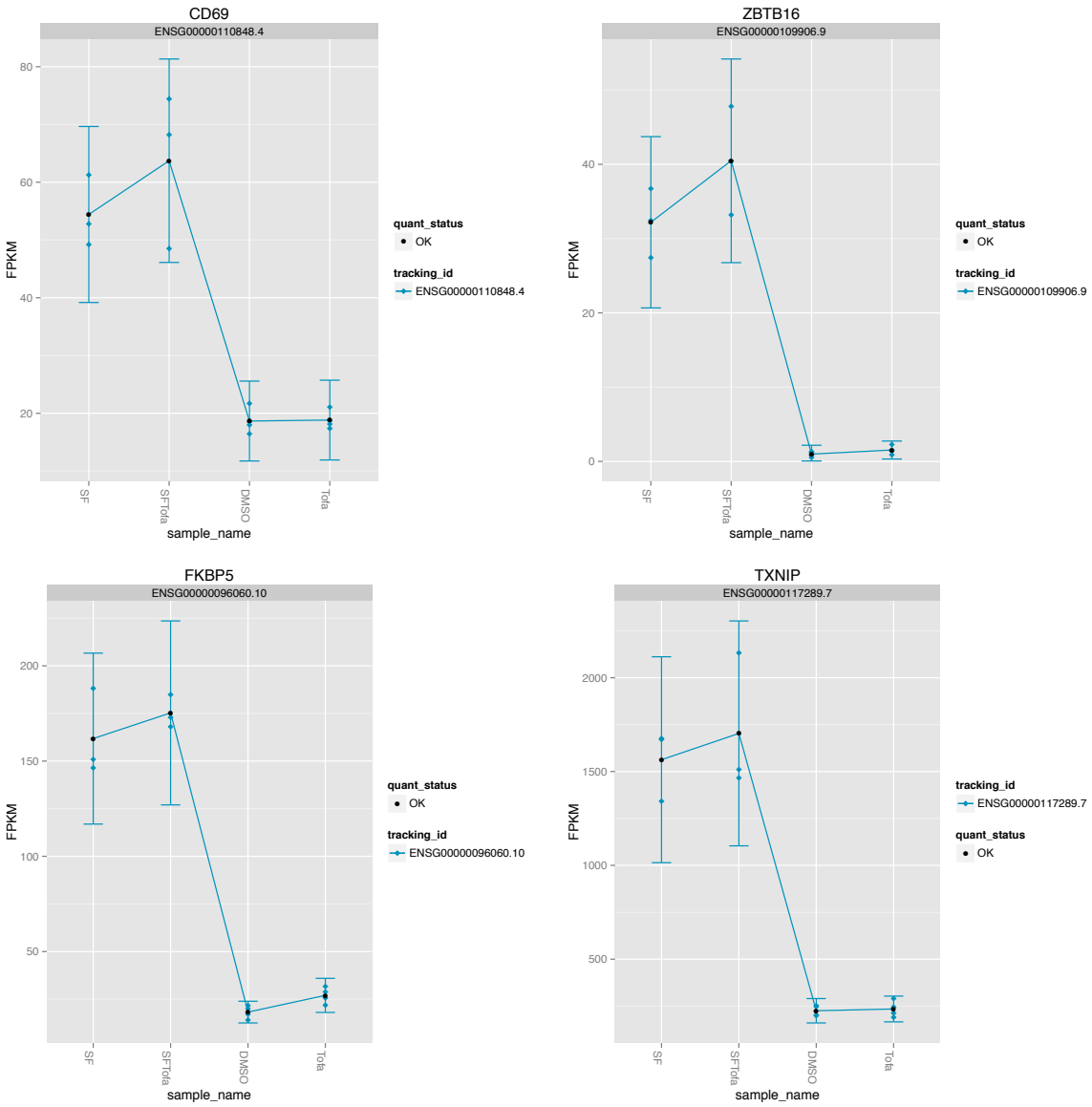


Figure 6-6 Genes such as CD69 and FKBP5 are upregulated following synovial fluid stimulation of CD4 T-cells and are not affected by tofacitinib. CD4 T-cells were isolated from the peripheral blood of patients with CCP+ RA by density centrifugation and positive selection using CD4 magnetic microbeads. 1.25 x10^6 per well were cultured in complete medium for 24 hours in a 24 well plate in four conditions: SF – 10% synovial fluid from RA patients, SFTofa – 10% synovial fluid from RA patients and 1000nM tofacitinib, DMSO – vehicle control condition, Tofa – 1000nM tofacitinib only. Cells were harvested, RNA extracted, quality controlled and sent for paired end ribosomal reduction RNA sequencing of length 75bp with 20-25M total reads. Raw data was aligned using TopHat and differential gene expression determined using Cuffdiff. FPKM plots were generated in CummeRbund following concatenation of two data sets. Error bars denote gene variance that is based on a combination of raw FPKM values and also sequencing depth. CD69, a marker of lymphocyte activation, is increased following synovial fluid stimulation and does not change with tofacitinib treatment. Other selected genes that change in a similar fashion are involved with calcineurin inhibition, redox or are transcription factors. (N=4 biological replicates)

6.8 Taqman Low Density Array data can be analysed using R and Bioconductor packages

To validate my findings from RNA sequencing, I custom designed a qPCR TLDA plate. I chose genes that were differentially expressed in the RNA-Sequencing experiment along with some housekeeping genes. I used cDNA that had ben prepared for the pre-sequencing qRT-PCR analysis of STAT3 related genes and therefore this had been prepared in one sitting by myself and therefore would not be subject to the technical effects experienced during sequencing.

TLDA data is often analysed in software packages such as Excel or Prism. Given my experience of R and bioinformatics packages, I chose to analyse the PCR data using R and Bioconductor. The package HTqPCR was developed by the European Bioinformatics Institute specifically for use with high throughput qPCR experiments. The advantage of using this method include the ability to annotate the data easily and carry out differential gene expression statistics with methods similar to microarray analysis.

Furthermore the package allows you to generate easy to interpret quality control plots and perform normalisation using the deltaCT method and plot differentially expressed genes easily. Finally the whole process of analysis can be written into a script and therefore the analysis of data should be reproducible and repeatable. The R script that I employed is included in the appendix.

Raw Ct values were inspected for spatial batch effects and there were none (appendix). Furthermore density plots of raw Ct and 18S normalised values showed good agreement between samples (figure 6-7). Also PCA analysis of the same data revealed two groups, one that had been treated with synovial fluid ie SF and SFTofa groups and the other that had not ie Control and Tofa groups (figure 6-8). These plots show that there is no obvious quality control issue and that there is likely to be a gross difference on differential gene expression between the synovial fluid and control cells.

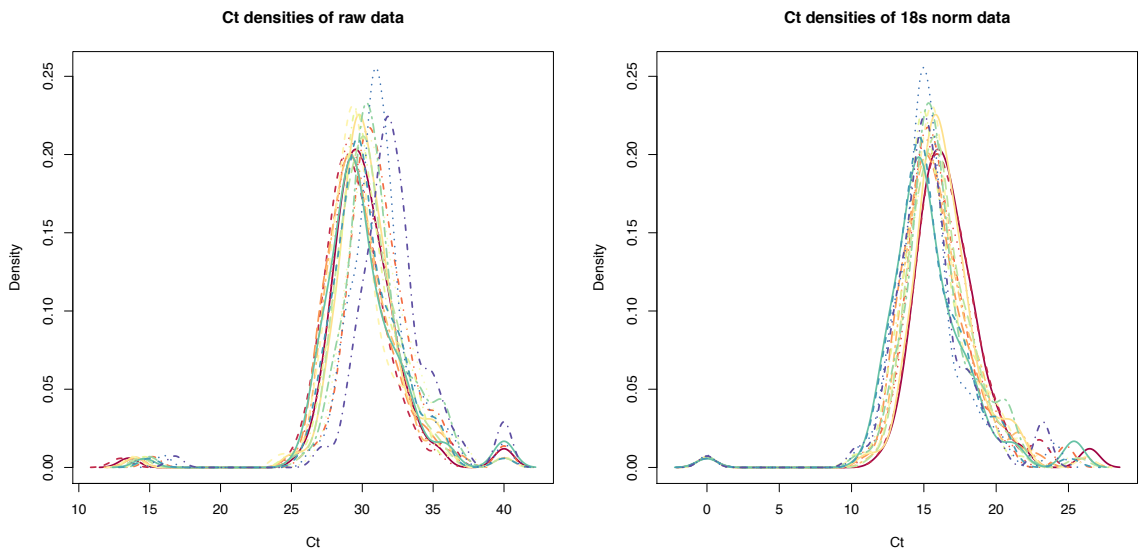


Figure 6-7 Density plot of raw Ct values shows reasonable agreement between samples and this improves when data is normalised to the 18S gene using the deltaCt method. CD4 T-cells were isolated from the peripheral blood of patients with CCP+ RA by density centrifugation and positive selection using CD4 magnetic microbeads. 1.25 x10^6 per well were cultured in complete medium for 24 hours in a 24 well plate in four conditions: SF – 10% synovial fluid from RA patients, SFTofa – 10% synovial fluid from RA patients and 1000nM tofacitinib, DMSO – vehicle control condition, Tofa – 1000nM tofacitinib only. Cells were harvested, RNA extracted, cDNA created and run on a TLDA plate. Smoothed density plots of Ct values were created using the HTqPCR package and show that there is reasonable agreement between the samples with most genes having a Ct value between 29 and 33. This improves with 18S deltaCt normalisation with average deltaCt values of 15. (N=4 biological replicates in 4 conditions)

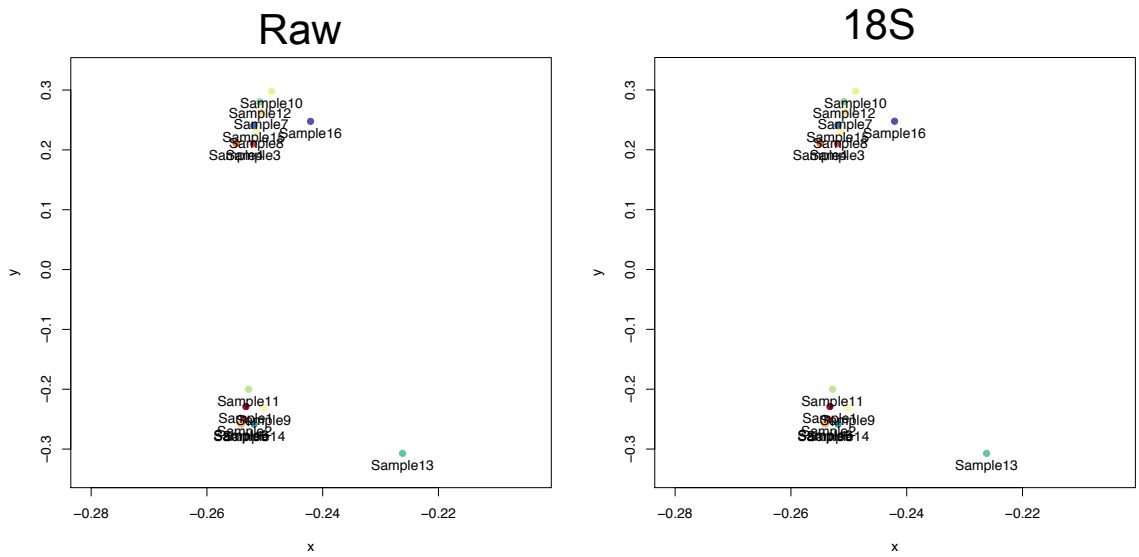


Figure 6-8 PCA of raw Ct and 18S deltaCt values shows two groups corresponding to presence and absence of synovial fluid stimulation with one outlier. CD4 T-cells were isolated from the peripheral blood of patients with CCP+ RA by density centrifugation and positive selection using CD4 magnetic microbeads. 1.25 x10^6 per well were cultured in complete medium for 24 hours in a 24 well plate in four conditions: SF – 10% synovial fluid from RA patients, SFTofa – 10% synovial fluid from RA patients and 1000nM tofacitinib, DMSO – vehicle control condition, Tofa – 1000nM tofacitinib only. Cells were harvested, RNA extracted, cDNA created and run on a TLDA plate. PCA plots of Ct values were created using the HTqPCR package and show that there are two groups present in the raw and normalised data, those that have been exposed to synovial fluid stimulation and those that have not. Sample 13 is an outlier and had a lower amount of input cDNA available due to experimental error. (N=4 biological replicates in 4 conditions)

I therefore went on to carry out limma based differential expression analysis within the HTqPCR package. A multiple correction adjustment method was not used in this case because these genes had been selected from a much larger population for validation. No differentially expressed genes were found between the tofacitinib treated and control cells and the other comparisons are presented below.

6.9 SBNO2 and MTHFD1L are confirmed to be upregulated in CD4 T-cells following RA synovial fluid stimulation and this is decreased by tofacitinib

The calibrator, or control, set of samples is set as the synovial fluid stimulated cells with the target samples set as the cells stimulated with synovial fluid and treated with tofacitinib. Therefore if a gene is downregulated on treatment with tofacitinib, it would result in a negative relative quantification (RQ) value. Furthermore I have also shown the mean deltaCt values in bar plots along with individual replicate values for each of the genes that were statistically different. To summarise, in a plot of deltaCt values, a lower deltaCt value corresponds to a higher transcript level.

BCL6, SOCS3 and MYC are confirmed and validated as differentially expressed, in these samples, between the two conditions (figure 6-9). Higher amounts of transcripts, for each of these genes, were found in the synovial fluid treated cells and this was decreased by tofacitinb. This result is in keeping with the pre-sequencing PCR and the RNA sequencing data giving me confidence that these genes are differentially expressed.

In addition, triple validating these genes gives me confidence that SBNO2, MTHFD1L and RELB are also differentially expressed. The RELB gene bar is hatched because in one replicate it was undetectable.

I went on to use a protein-protein interaction database, STRING 10 (296)to visualise whether these other genes are connected to each other and if they are known to form networks with JAK/STAT signaling.

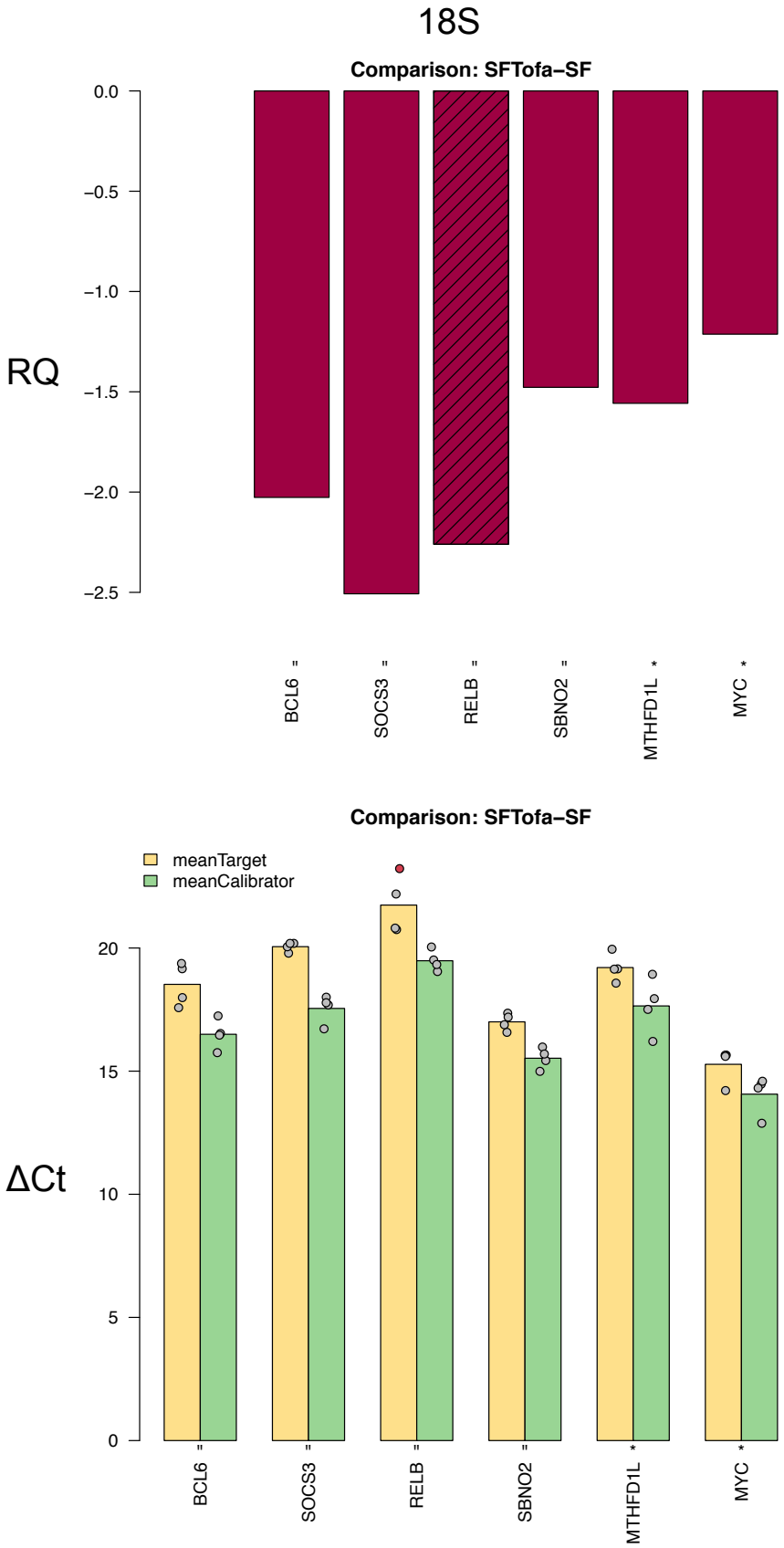


Figure 6-9 Bar chart of relative quantification and deltaCt values between CD4 T-cells treated with synovial fluid and tofacitinib (Target) and those treated with synovial fluid alone (Calibrator). CD4 T-cells were isolated from the peripheral blood of patients with CCP+ RA by density centrifugation and positive selection using CD4 magnetic microbeads. 1.25 x10^6 per well were cultured in complete medium for 24 hours in a 24 well plate in four conditions: SF – 10% synovial fluid from RA patients, SFTofa – 10% synovial fluid from RA patients and 1000nM tofacitinib. Cells were harvested, RNA extracted, cDNA created and run on a TLDA plate. RQ values were calculated using $2^{\Delta\Delta Ct}$ and 18S as a normalisation gene. Mean deltaCt values for target and calibrator samples are shown by boxplot with individual replicate values. Each gene is expressed at a higher level in synovial fluid treated CD4 T-cells and expression is reduced by tofacitinib treatment. Bayes moderated t-tests were used to test statistical difference. "=p<0.01, *=p<0.05. (N=4 biological replicates in 2 conditions)

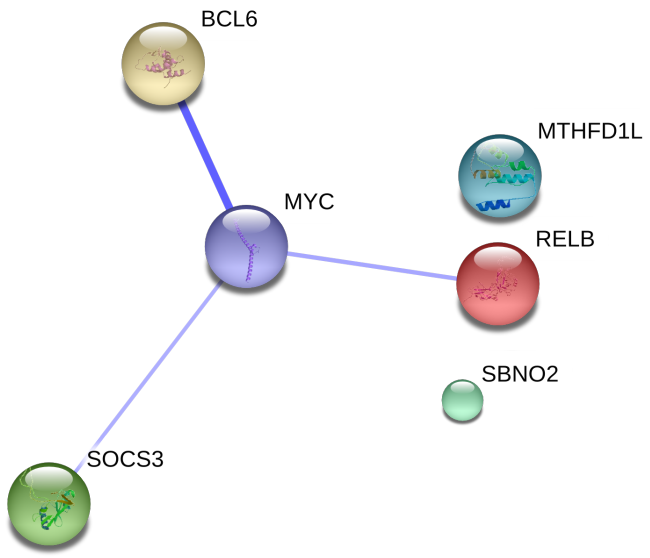


Figure 6-10 SBNO2 and MTHFD1L sit out with a network made of BCL6, SOCS3, MYC and RELB. Statistically different genes between CD4 T-cells that were stimulated with synovial fluid and treated with tofacitinib were analysed using the web version of STRING 10. Both SBNO2 and MTHFD1L do not have predicted interactions with other genes.

Four of the differentially expressed genes are related to each other (figure 6-10) and on growing the network, they are related to JAK signaling (figure 6-11). The thickness of the line between two proteins denotes the amount of evidence for that interaction. Evidence is derived from sources including databases of high-throughput experiment results, published literature or computational prediction.

To determine if there are networks that could act in concert I decided to us the MCL clustering algorithm in STRING(297,298). In this method you define the number of subnetworks you wish to have by setting a value called "inflation". As the value of the inflation increases, so do the number of clusters. This method of clustering relies on a global score for the gene that is assigned from the database based on the strength of evidence for the protein-protein interaction. This score is determined from high throughput experiments, text mining literature as well as predictive genome analysis.

When I performed clustering analysis on a network that had been enlarged (figure 6-11), I found three potential sub-networks: one related to MYC, one connecting JAK2 to RELB and SOCS3 and a final network connecting BCL6 to co-repressing genes. Furthermore, SBNO2 and MTHFD1L still remain out with this network.

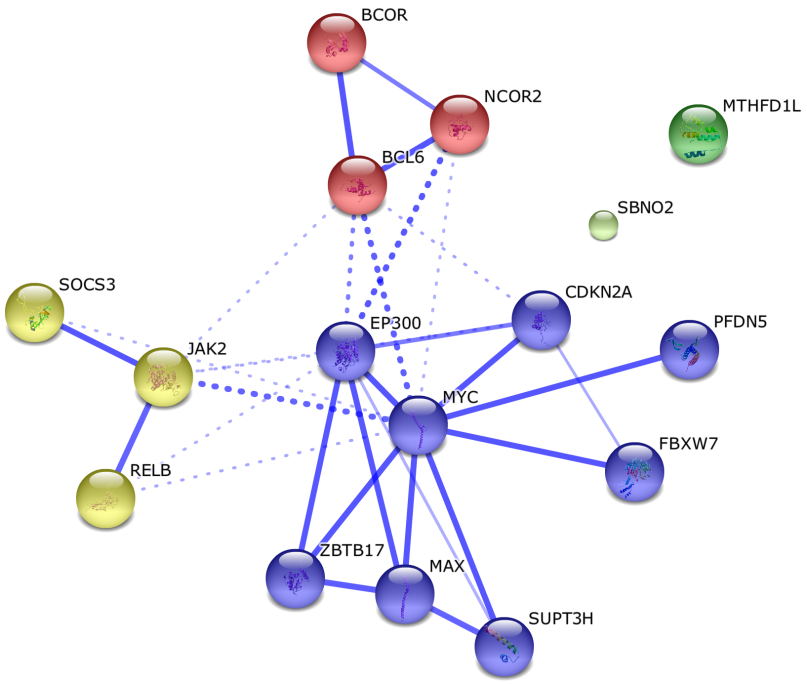


Figure 6-11 SBNO2 and MTHFD1L sit out with an enlarged network that has been subjected to clustering and demonstrates three subnetworks. Statistically different genes between CD4 T-cells that were stimulated with synovial fluid and treated with tofacitinib were analysed using the web version of STRING 10. The network was then grown and subjected to MCL clustering with an inflation value of four. Three subnetworks emerge but there are interactions between them with JAK2 directly connected to SOCS3, RELB, MYC and BCL6. Both SBNO2 and MTHFD1L do not have predicted interactions with other genes although this may be due to lack of public data.

6.10 RA synovial fluid stimulation of CD4 T-cells results in a large number of differentially expressed genes

When synovial fluid stimulated CD4 T-cells are compared to control, there are a large number of validated, statstically differentially expressed genes. In the RQ plot below (figure 6-12), the calibrator (or control) is DMSO and synovial fluid is the treatment. Therefore genes that are upregulated with synovial fluid stimulation have a positive RQ value.

Although many of these genes are upregulated, others are downregulated including LTB and CD82. Furthermore when we inspect the deltaCt plots (data not shown), many of these genes are changed following synovial fluid exposure but the transcript level is not altered by tofacitinb suggesting that other stimuli are present in RA synovial fluid and they signal via pathways other than JAK/STAT.

On initial analysis in STRING 10, two networks emerge: one which links IL23, SOCS3 and NFKBIA related genes and another involving ZBTB16, TXNIP and BCL6. This occurs without growing the network.

Unfortunately due to experimental error, supernatants for evaluation of secreted cytokines and growth factors were not available and therefore I was unable to confirm whether these cells were secreting IL23.

When the network was grown within STRING and then MCL clustered, three clusters emerge:

1. NF-κB related proteins outlined in blue which also have a link to the JAK/STAT pathway via RELA

2. JAK2, IL4, IL7 and EGFR related network outlined in yellow

3. BCL6, ZBTB16, TXNIP and DDIT4 outlined in red

Furthermore, many of the genes do not fit into a network but further work is required to elucidate the mechanism behind these changes and whether they are pathology related or reactive.

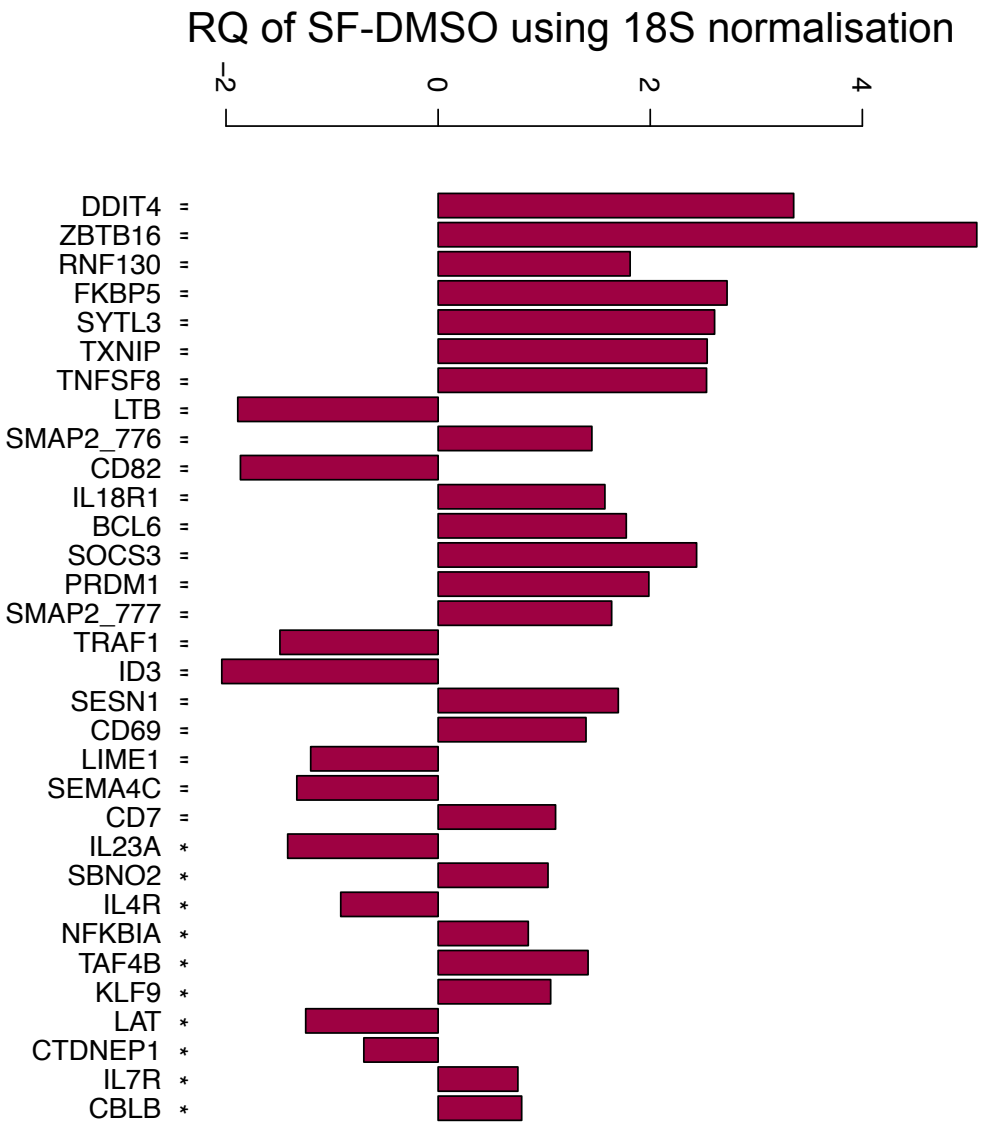


Figure 6-12 Bar chart of relative quantification between CD4 T-cells treated with synovial fluid (Target) and those treated with vehicle control alone (Calibrator). CD4 T-cells were isolated from the peripheral blood of patients with CCP+ RA by density centrifugation and positive selection using CD4 magnetic microbeads. 1.25 x10^6 per well were cultured in complete medium for 24 hours in a 24 well plate in four conditions: SF – 10% synovial fluid from RA patients, DMSO – 0.001% DMSO. Cells were harvested, RNA extracted, cDNA created and run on a TLDA plate. RQ values were calculated using $2^{\Delta\Delta Ct}$ and 18S as a normalisation gene. The genes shown above are statistically different between synovial fluid treated CD4 T-cells and control treated cell. Bayes moderated t-tests were used to test statistical difference. "=p<0.01, *=p<0.05. (N=4 biological replicates in 2 conditions)

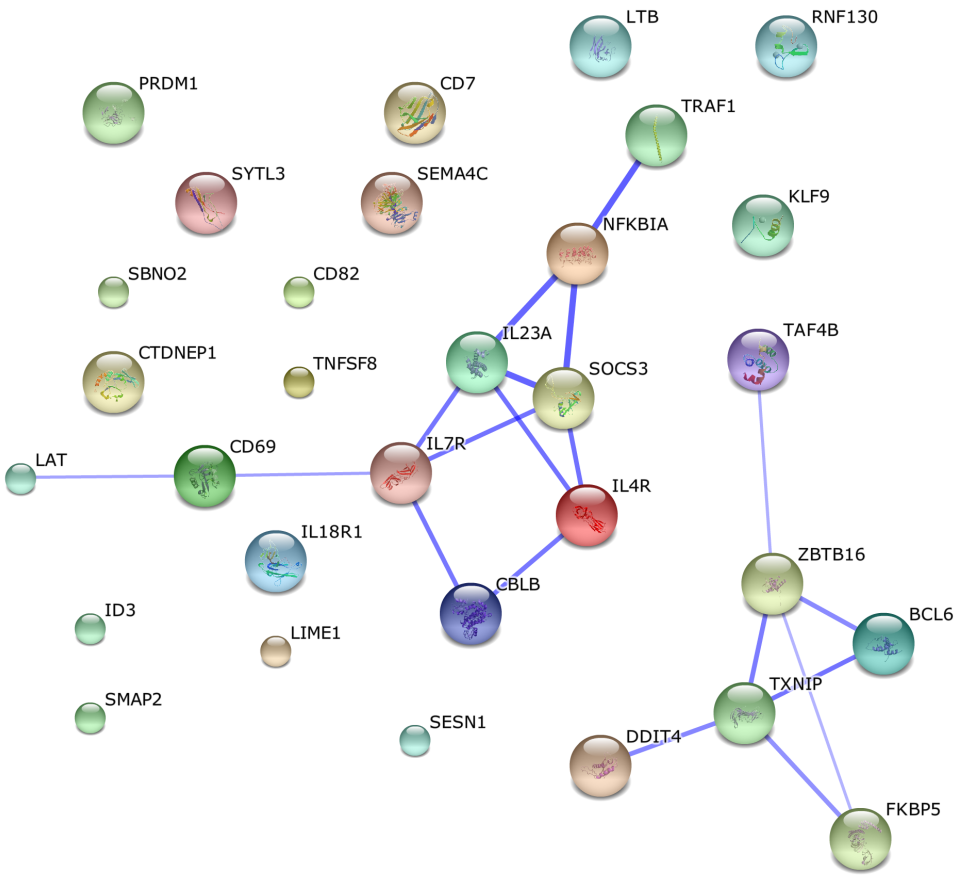


Figure 6-13 Two networks emerge following analysis of statistically differentially expressed genes between CD4 T-cells treated with synovial fluid and control cells. Statistically different genes between CD4 T-cells that were stimulated with synovial fluid and control cells were analysed using the web version of STRING 10. Two networks emerge based on known interactions between genes in the dataset.

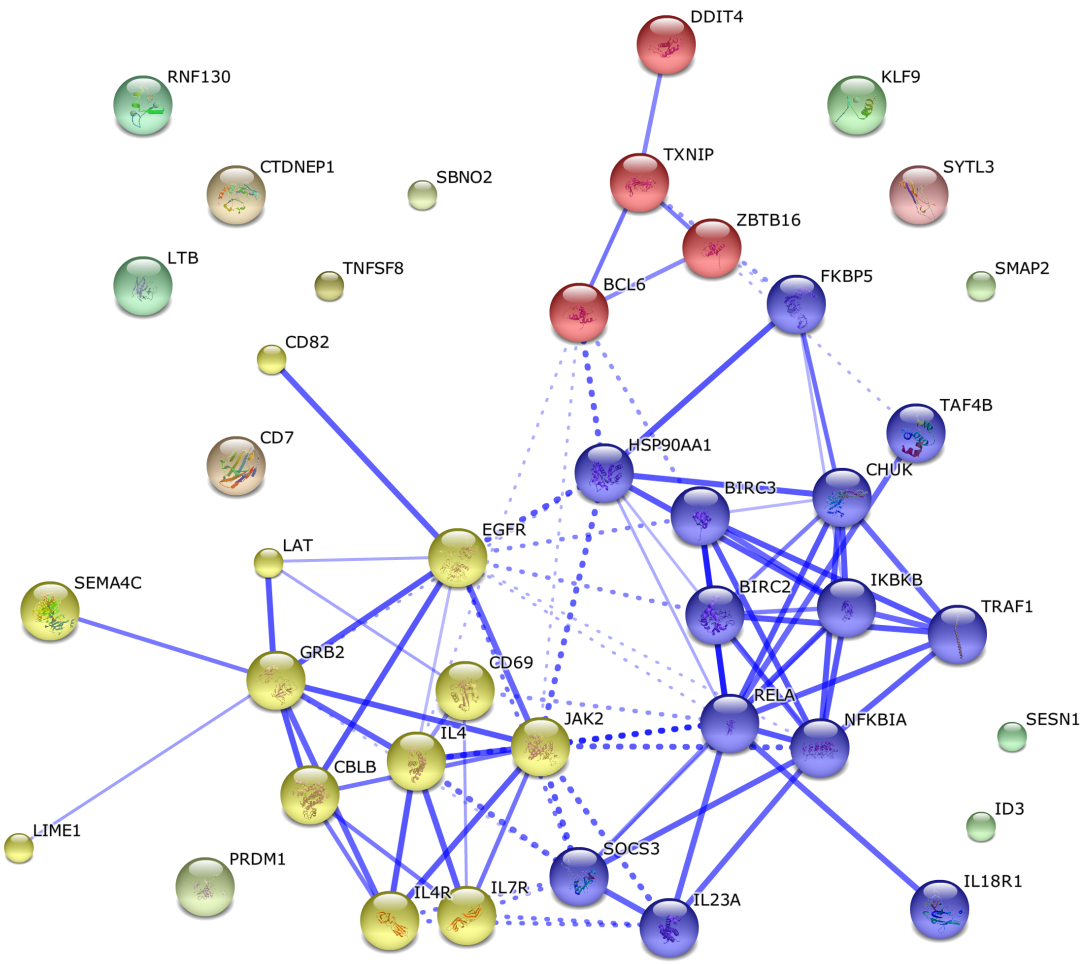


Figure 6-14 Three networks emerge following analysis of statistically differentially expressed genes between CD4 T-cells treated with synovial fluid and control cells. Statistically different genes between CD4 T-cells that were stimulated with synovial fluid and control cells were analysed using the web version of STRING 10. The network was grown and clustered using the MCL algorithm. Three networks emerge coloured in red, blue and yellow based on predicted protein-protein interaction. JAK signaling via JAK2 links the blue and yellow network although many other genes remain outside.

6.11 Discussion

In this chapter I set out to:

1. Demonstrate by pre-sequencing qRT-PCR that selected genes downstream of STAT3 are upregulated by stimulation with RA synovial fluid of CD4 T-cells and this is reversed by tofacitinib

2. Following this, show that ribosomal depletion and RNA Sequencing can be used to discover other genes that follow a similar pattern of upregulation with RA synovial fluid and reversal of this with tofacitinib

3. To validate these genes by qRT-PCR and show that their levels are changed in keeping with the RNA Sequencing results

4. To use pathway analysis tools to predict whether networks of genes could act in concert and propose regulators of these networks that may be amenable to therapeutic targeting

I confirmed that SOCS3 and BCL6 are upregulated following synovial fluid stimulation of CD4 T-cells and that this is reversed with tofacitinib treatment. Furthermore, RNA sequencing of these cells and bioinformatic analysis of data, reveals other transcripts such as SBNO2 that are differentially expressed in the same way. By using ribosomal depletion I have also found some long non-coding RNA that are differentially expressed in the same pattern although due to time and sample constraints I have not been able to validate these.

These three genes were validated using qRT-PCR as being differentially expressed and although SOCS3, BCL6 and MYC are related to STAT signaling, SBNO2 does not sit within this network when interrogated using protein-protein interaction databases. Finally synovial fluid causes the differential expression of many more genes that are not changed with tofacitinib treatment demonstrating that soluble factors signaling via pathways other than JAK/STAT are present.

Using a novel technology such as RNA sequencing can lead to experimental and data analysis difficulties and I experienced both. In terms of limitations of my work I will discuss these with regards to my experimental design, the RNA sequencing and limitations of data analysis.

Based on my findings in chapter 5 where CD4 T-cells responded to soluble factors in RA synovial fluid by phosphorylating STAT3, I decided to perform RNA sequencing in these conditions to further explore this on a global scale. Furthermore, in keeping with my exploration of JAK inhibitors, I decided to use tofacitinib as a treatment to determine if downstream transcript changes could be prevented.

In this way, although RNA sequencing could provide me with large amounts of data, I could remain focused with concentrating on genes that were changed following stimulation and drug treatment. Furthermore I decided to use RA synovial fluid from CCP+ patients to simulate the synovial microenvironment in terms of soluble factors. Although the fluid contained high concentrations of IL-6, which could be responsible for the phosphorylation of STAT3, other unclassified stimuli could also be present which limits the usefulness of this study.

Furthermore, as with experiments in chapter 5, my results are only applicable to the synovial fluid samples that I used although further experiments could be performed using other synovial fluid samples and also specific cytokine stimulation using IL-6. In addition, I have generated a custom TLDA card and therefore subsequent experiments would not need to be sequenced if I wished to restrict myself to genes that are differentially expressed in the current experiment.

Finally a significant limitation of my work is the lack of supernatants, due to experimental error, from CD4 T-cell culture under various conditions. In particular measurement of T-cell related cytokines such as interferons, IL4, IL-21 or IL-10 would help to determine if the naïve cells were being driven down a particular differentiation path(299) by synovial fluid stimulation and whether this was reversed by tofacitinib.

In terms of execution of RNA sequencing, the pre-sequencing qRT-PCR screen helped to detect at an early stage whether changes in genes such as SOCS3 and BCL6, which are increased following activation of STAT3, were detectable prior to the expense of RNA sequencing.

Furthermore, I decided to split my experiment into two sequencing runs and the second was complicated by a failure of ribosomal depletion. This meant that the second set of samples, which were tofacitinib treated with and without synovial fluid, had a lower and inconsistent amount of input RNA and a different kit was used to deplete the sample of common ribosomal and mitochondrial sequences. This de facto has led to a significant batch effect that cannot be corrected and therefore meant that results that were based on experiments across sequencing runs were subject to bias. This may account for why so few genes were validated by TLDA qRT-PCR between the synovial fluid and synovial fluid with tofacitinib treatment conditions.

I decided to use ribosomal reduction at the library preparation stage for RNA sequencing to assess the coding and long non-coding species. This would not have been possible had I used a polyA selection method. However I have not been able to validate any non-coding species in the study to date because of lack of sample and this would be necessary future work.

In addition, although it is suggested to use 50 Million reads in RNA sequencing experiments of eukaryotic cells when using ribosomal depletion, I was only able to use 20-25 million reads because of resource constraints. I was able to detect statistically differentially expressed genes using these parameters, however it is possible that both coding and non-coding RNA that are expressed at low copy numbers would not be detected in my experiment.

Processing of RNA sequencing data can be a daunting concept for most researchers without dedicated dry lab bioinformatics facilities. Initial sequencing experiments required heavy infrastructure in terms of sequencers, computing power in the form of processors, short term swap storage and also long term storage as well as custom software to align and build the sequence from fragments. This whole process has been significantly shortened and also made easier by three factors:

1. The improvement of computing resources according to Moore's law
2. The development of cloud based sequencing analysis pipelines such as Illumina Basespace and Galaxy projects
3. The research and development of newer and more memory and processor efficient algorithms for sequence alignment and differential gene expression

I decided to use Illumina Basespace because it is a cloud-based application that acts as a data repository and allows simple analysis in an app-based manner. The main advantages of this system are that it is flexible because more computing power can be brought online as required, the data is backed up and held securely and data analysis is reproducible because most of the pipeline is decided for the uses. The disadvantages are that customisation of the pipeline is almost absent, newer genomes and tools take a while to propagate to the platform and the data can be prohibitively large to download to a local machine.

Alternatives to using Basespace include other graphical user interface (GUI) based methods such as the Galaxy project or using command line tools. The former is possible but without a local Galaxy instance, the data transfer to a server overseas is a bottleneck. Furthermore, although the raw data for this experiment is in the region of 80GB, almost ten times this amount is required for analysis as a minimum once all alignment files have been generated. Thereafter, if you wished to carry out a full analysis with iterative changes to analysis parameters, approximately 4-6 TB of storage space would be required. This is not an option given that most public Galaxy instances give at the maximum, 250GB of storage space.

Command line tools can be used on a local machine or server and these allow optimization of the analysis pipeline, however they also require local installation of a reference genome and knowledge of command line tools that often have a steep learning curve. Finally RNA-Seq analysis can be performed within R but in this case I decided to use Basespace because it uses a well-tested analysis pipeline and is also able to process the samples quickly without having to download the raw data.

Limitations to using this approach, for me, included the lack of access to the newest human genome, GRCh38 that was initially released in 2013. Basespace utilises GRCh37 that was released in 2009 and therefore has an older human reference genome installed. The newer genome has fewer gaps, and includes up to date annotations as well as mitochondrial sequences.

Furthermore, Basespace utilises either TopHat as an aligner or STAR(300,301) as part of the RNA Express pipeline. I decided to use TopHat because it was part of the Tuxedo pipeline and also for ease of use within BaseSpace. However more efficient and faster

aligners have been developed along with newer methods of calculating differentially expressed genes.

In general Tophat alignment of reads to a reference genome takes approximately 24 hours using Basespace prior to running a differential expression program. If an experiment has to be optimised this can be cumbersome if local computing power is limited.

There are newer algorithms available such as HISAT(302) and kallisto(303) which offer genome alignment at much higher speed. In terms of speed, an experiment that takes almost 2000 minutes to analyse in Tophat and Cufflinks is processed in just over 5 minutes by kallisto. This allows experiments with larger numbers of replicates to be carried out but also means that bioinformaticians can execute optimisation of analysis pipelines without significant time expense.

In summary, although there are significant limitations in the experimental setup and data analysis, I believe that this study does add to the knowledge of the CD4 T-cell transcriptome and also the effect of tofacitinib on gene expression.

There are very few publicly available studies using RNA sequencing of primary cells in inflammatory arthritis. The repositories of GEO and ArrayExpress have four such studies that are described further in the context of my results and three of these studies concentrate on CD4 T-cells with two of the studies involving in vitro drug treatment of the cells.

The latest study from Peeters et al (304) investigated the role of drug inhibition of super-enhancers in synovial fluid derived CD4 T-cells and the effect on gene expression. Enhancers are normally a few hundred base pairs in size and regulate DNA transcription by binding transcription factors and co-factors and can be analysed based on Histone 3 Lysine 27 acetylation (H3K27ac).
Super enhancers are longer lengths of DNA that can affect cell identity and disease specific genes and can be affected by drug treatment. Vahedi et al (305) investigated the role of super-enhancers and found that of the 98 currently known genetic variants in RA (90), more than half were lined to super-enhancer regions in CD4 T-cells. These SNPs often fell within intergenic regions and were associated with non-coding RNA. Furthermore, the group went on to use tofacitinib as an in vitro treatment of CD4 T-cells that had been activated by anti CD3/CD28 and found that this drug had a preferential effect on super-enhancer sites in CD4 T-cells. Finally they found that SNPs of other conditions

such as type 2 diabetes and cancer, where CD4 T-cells do not have a large role, were not over represented in CD4 T-cell super-enhancer regions suggesting that disease pathobiology is crucial in super-enhancer biology.

Peeters et al went on to use publicly available data to compare the enhancer and super-enhancer profiles of naïve CD4 T-cells (CD4+, CD45RA+, CD45RO-) and effector/memory cells (CD4+, CD45RA-, CD45RO+)(304). Following this they then went on to assess the super-enhancer landscape within these subsets that had been derived from healthy controls and activated by in vitro stimulation with anti CD3/CD28 beads and compared this to synovial fluid derived T-cells from patients with juvenile idiopathic arthritis. They discovered a disease specific super-enhancer signature that was only evidence in synovial fluid cells, furthermore they went on to treat synovial fluid derived effector/memory T-cells with a bromodomain inhibitor and found that it inhibited immune related super-enhancers and also decreased disease associated genes such as cytokines.

Their study was based on a small number of patient samples (n=4) but they were activated or treated with drug to investigate a specific hypothesis; that super-enhancers in primary site-derived cells would offer more insight into disease pathogenesis than healthy activated cells. It went further by looking at synovial fluid derived T-cells and also sorting them into naïve and effector/memory cells that I did not do. In addition, the novelty of combining an epigenetic approach by using chromatin immunoprecipitation of primary cells and then overlaying publicly available SNP data from GWAS is appealing and an approach I will employ in subsequent studies.

Jiang et al (306) investigated the role of non-coding RNA in JIA by again combining an epigenetic approach with RNA sequencing. Their reasoning for this was again data from GWAS of JIA showing that most of the risk loci associated with disease are located in non-coding regions of the genome. Therefore they carried out RNA sequencing to determine differential expression of long non-coding RNA in CD4 T-cells and neutrophils from the blood of patients with JIA. This was combined with public and researcher generated data of Histone 3 Lysine 4 monomethylation (H3K4me1) and H3K27ac to find areas of functional relevance within the non-coding genome. In doing so they discovered two new long non-coding RNA within neutrophils but when validated in a clinical cohort, did not discriminate between disease and healthy. This group sequenced neutrophils from 16 patients, 8 with untreated disease and 8 in remission on drugs and combined this with public epigenetic data on CD4 T-cells and also neutrophil epigenetic data they generated

themselves. This particular study does not look at samples from patients with RA but does show that an approach of combining data from GWAS ie SNPs and then merging this with epigenetic data such as H3K4me1 or K3K27ac to identify areas of functional relevance in the non-coding genome is valid. Furthermore, I could employ this same approach in further work to investigate whether long non-coding RNA that I identified are associated with RA risk loci to prioritise investigation of the sequencing data.

Rochman et al (307) performed RNA sequencing on resting, anergic and effector CD4 T-cells. These cells were isolated from healthy peripheral blood and generated in vitro and with anergic cells exposed to abatacept. They demonstrated that human data did not agree with mouse models of anergy and confirmed, using western blots, that $p27^{kip}$ cyclin-dependent kinase inhibitor was associated with abatacept induced hyporesponsiveness in anergic cells. This study does not add directly to my own but does demonstrate that investigation of cell subsets is crucial as $p27^{kip}$ protein levels were significantly different between naïve, effector memory like and anergic T-cells.

Finally Donlin et al (308) carried out RNA sequencing of one biological sample subjected to four conditions: macrophage alone, macrophage and fibroblast co-culture, macrophage and TNFα and macrophage with fibroblast in co-culture and TNFα. A large number of genes were differentially expressed in each condition but due to the lack of replicates and hence statistics, it is difficult to comment on whether these results are reproducible. Furthermore, RNA sequencing of co-cultured cells leads to the transcriptome of two completely different cells being mixed and an approach employing single cell sequencing may provide clearer results.

Three out of four studies investigated the role of CD4 T-cells including subsets thereof and two of these studies involved a drug. They demonstrated that combining genomics, epigenetics and RNA sequencing data in a polyomic fashion is useful when employing a hypothesis-based approach. Furthermore, cell subsets show differential responses to stimulation and using a drug helps to focus the experimental analysis.

Therefore my study adds to the field in two ways: firstly it addresses RA and secondly it looks at the effect of JAK inhibition in the context of synovial fluid stimulation to simulate the RA synovial microenvironment. These two factors make it novel. Furthermore, compared to the publicly available data, my study uses more biological replicates therefore improving the reliability. The study from Jiang et al used 16 samples but unfortunately

these are not publicly available. Synovial fluid as opposed to cytokines alone were chosen because many other unclassified factors may be present in the fluid itself and therefore exposing the CD4+ T-cells from patients with seropositive RA to synovial fluid from patients with seropositive RA allowed me to simulate the synovial microenvironment and the milieu in a synovitic joint.

In addition to my pre-selected genes: SOCS3, BCL6 and MYC, SBNO2 was also increased following synovial fluid treatment of CD4 T-cells and this was reduced by tofacitinib. SBNO2 or strawberry notch homologue 2 has been shown to have transcriptional repression activities in macrophages(309). El Kasmi et al showed that IL-10 mediated STAT3 activation led to increased levels of SBNO2 in human macrophages. Furthermore, they went on to knockdown STAT3 and BTK using siRNA and demonstrated that the effect was STAT3 dependent. Knockdown of BTK had no effect on IL-10 mediated SBNO2 production. In addition, by using luciferase constructs in 293T cells, this group demonstrated that SBNO2 inhibited NF-κB but not IRF7 mediated transcription.

Furthermore, in a murine model, sbno2 as been shown to regulate osteoclast fusion with sbno2 deficient mice having increased bone mass(310). Therefore there is a role for SBNO2 in myeloid lineage cells and this may be related to IL-10 mediated STAT3 activation. However in CD4 T-cells from RA patients following STAT3 activation, I see upregulation of SBNO2 and this is reversed by tofacitinib. Furthermore, I was not able to detect IL-10 in the synovial fluid pool that I used to stimulate the CD4 T-cells. This suggests that STAT3 activation with other ligands such as IL-6 may have a similar effect.

Sbno2 is also expressed by astrocytes in the murine CNS (311)and is stimulated by endotoxin and also IL-6 suggesting that both the innate and adaptive immune system are involved. Other STAT3 ligands such as OSM, LIF and IL-11 could also increase sbno2 levels and so this work showed that quantifying these other ligands, in synovial fluid, is an essential next step in my work.

TNFα and IL-1β stimulation of astrocytes also led to increased sbno2 but this was independent of STAT3 phosphorylation showing that sbno2 transcript is governed by a STAT dependent and independent mechanism. This with the finding from El Kasmi suggests that SBNO2 may be able to link STAT3 and NK-κB signaling and may be a therapeutically tractable target.

The role of SBNO2 in CD4 T cells in Rheumatoid arthritis has been explored in clinical samples from the Newcastle early RA clinic by Pratt et al (282). This group examined peripheral blood CD4 T–cells from patients and compared the transcriptome using Illumina Beadarrays from patients with RA to those who had spondyloarthropathy, self-limiting arthritis and osteoarthritis. The group found that in their study genes that were associated with STAT3 activation such as SBNO2 were expressed at a higher level in the CD4+ T cells of those patients with CCP negative RA. This is in contrast to what I found because I chose to derive cells from patients who were seropositive for CCP and also the synovial fluid samples were from patients who had high levels of anti CCP antibodies.

Furthermore they showed a correlation between serum IL-6 concentration and SOCS3, SBNO2, BCL3 and PIM1 expression levels. However this correlation was stronger in patients who were CCP negative as opposed to seropositive RA. Therefore my study adds to this in that I found a similar finding in patients with seropositive RA. These results were validated in a further cohort by Anderson et al (283) who showed that a gene signature that included SBNO2 was useful in predicting progression from undifferentiated arthritis to Rheumatoid arthritis in CCP negative patients. In conclusion, SBNO2 is a novel gene that has an as yet unknown role in CD4 T cells in RA and I have found that it is increased following STAT3 activation of CD4 T-cells by RA synovial fluid and decreased by tofacitinib treatment. The ligand responsible for this effect is likely IL-6 but this requires clarification and evidence suggests that SBNO2 may regulate and be regulated by NF-κB therefore suggesting that it may link STAT3 and NF-κB mediated signaling.

With regards to future work, in data processing, I believe that optimising alignment and normalisation methods will help prioritise the differentially expressed genes. As previously eluded to, I plan to use my local computing resources to align my data to the latest human reference genome GRCh38. Finally, by using de novo assembly methods and also techniques to discover novel genes I plan to fully explore this novel dataset.

Combining publicly available GWAS and epigenetic datasets to investigate the role of lncRNA is an area of biology that I would explore. Non-coding RNA mainly in the form of microRNA are important in the pathogenesis of inflammatory arthritis (232,312-314)and given that most of the SNPs associated with disease in RA lie in non-coding regions, this method will help to determine if lncRNA that are differentially expressed in my dataset could be important.

Further wet lab experimental work could include using synovial fluids from other arthropathies and specific cytokines such as IL-6 and others that activate STAT3 to determine whether my findings are unique to this synovial fluid or generalisable. Furthermore, blockade of IL-6 receptor using tocilizumab to prevent IL-6 signaling would help to this end. Finally, JAK1/2 inhibitors such as ruxolitinib and baricitinib may be more potent in inhibiting transcripts such as SBNO2 given that the IL-6 receptor also utilises JAK2.

Although there is an association between SBNO2, IL-6 and disease activity, this has not been explored in a mechanistic fashion in RA and I would propose to knockout this gene in fibroblasts or cell lines such as THP-1 or Jurkatt T cells using CRISPR technology to allow me to functionally characterise this gene.

Finally, it is clear from other work that single cell sequencing is maturing as a technology and providing insight into disease pathogenesis with increasing fidelity(315,316). Publicly available RNA sequencing studies have shown that CD4 T-cell subsets respond differently to the same stimulus and therefore investigating this at a single cell resolution would help to investigate risk loci in RA that are currently in non-coding regions by utilising new methods in single cell epigenetics and RNA sequencing.

Chapter 7 General Discussion

In my studies presented herein, I used bioinformatic tools to re-analyse a microarray dataset of synovial fluid macrophages in RA to determine if a JAK/STAT signature was present and unique to RA. I decided to do this to fully exploit the data because packages such as Genespring still rely on traditional statistics such as t-tests and ANOVA but it has been shown that hierarchical models such as those implemented in the Bioconductor package limma are more suited to analysing data where we have a large number of observations and only a few samples the so called "large p and small n" problem.

Therefore within this package, I compared methods of normalisation of microarray data and discovered that RMA normalisation was a reasonable method of normalisation if there was no evidence of batch effect. I also importantly discovered that there is no way of measuring without performing a complete validation experiment whether one method of normalisation is superior to another. Assessment of quality control can help a user decide between one or the other but unfortunately as with many 'omics technologies, there is no "right" answer. I also compared methods of differential gene expression and settled on an analysis pipeline to further examine the dataset.

Within my data, I discovered that there are differences between RA and PsA synovial fluid derived CD14 macrophages. Furthermore, a JAK/STAT signature is present and differs between the two arthropathies with IL-6R and STAT3 signaling more prominent in RA and IL21R upregulated in PsA. This demonstrated that the signature was present, as others have shown, and that there were subtle differences between the two diseases despite the same cell type being isolated from a similar environment.

Following demonstration of this signature in RA, I decided to use the JAK1/3 inhibitor, tofacitinib to investigate the effect of JAK inhibition on an in vitro model of macrophage cell contact activation by cytokine stimulated T-cells(121). I demonstrated that both tofacitinb and ruxolitinib, a JAK 1/2 inhibitor decreased the production of TNFα in this assay. Furthermore I also demonstrated that the pro-inflammatory cytokines IL-6 and IL-15 are also decreased. Finally chemokines were also affected and the interferon responsive chemokines MIG and IP-10 were decreased not only in the cell contact stimulation assay but also following LPS mediated stimulation of macrophages.

This suggests but does not prove that secondary interferon signaling is affected by JAK inhibition although similar results have been found in TNFα stimulated macrophages (267)and also a fibroblast and macrophage co-culture assay(74). Further work is required

and the use of fixed Tck or cell membrane fragments along with interferon neutralising antibodies would help to explore this mechanism further.

I then went on to explore whether RA synovial fluid would stimulate RA patient derived monocytes and if this stimulation could be inhibited by tofacitinb. I employed phospho FACS analysis of pSTAT1, pSTAT3 and pSTAT6 using a stimulation cocktail with and without tofacitinib treatment to demonstrate that JAK inhibition prevented STAT phosphorylation.

RA patient monocytes phosphorylated STAT1 and STAT6 in response to the stimulation cocktail and this was decreased by tofacitinb. However, RA synovial fluid did not cause phosphorylation of STAT1, STAT3 or STAT6 in RA monocytes despite high levels of IL-6 in RA synovial fluid. As there was no phosphorylation of STATs, tofacitinib would not have a direct "on-target" effect in this system. In summary, I demonstrated that in the RA synovial fluid that I used, there was no phosphorylation of STATs 1, 3 or 6 in RA monocytes.

However, CD4 T-cells did phosphorylate STAT3 following synovial fluid stimulation to a degree similar to cytokine stimulation alone. IL-6 should phosphorylate both STAT1 and STAT3 but I only saw phosphorylation of the later. This may have been due to the length of time the cells spent in culture or a disease specific effect. This finding requires further investigation using cells from healthy donors and also other arthropathies such as PsA.

Tofacitinib has been shown to inhibit IL-6 mediated STAT3 signaling and therefore I went on to explore this further using RNA sequencing. I chose to use RNA sequencing instead of microarray technology to examine the effect of synovial fluid stimulation on CD4 T-cells on a global scale and with high fidelity. I went on to use ribosomal reduction as opposed to polyA selection to investigate if long non-coding RNA were changed following stimulation and treatment with tofacitinib.

I demonstrated that a pre-sequencing qRT-PCR screen of SOCS3, BCL6 and MYC showed that these genes were upregulated by synovial fluid treatment and that this was decreased by tofacitinb in cells derived from four patients with RA. These genes were also confirmed as differentially expressed by RNA sequencing and were validated again using a qRT-PCR TLDA array.

Using further bioinformatics tools in R, I showed that SBNO2 follows a similar pattern to SOCS3 and BCL6 in that it is upregulated following synovial fluid stimulation and is decreased by tofacitinib. Furthermore, this gene has been identified in peripheral blood CD4 T-cells from RA patients and has been used to develop a gene metric to predict the progression from undifferentiated arthritis to RA(282,283). There is also evidence from other cell types that this gene is upregulated following IL-6 stimulation but the exact role is unclear. Therefore this gene may represent a biomarker of response to tofacitinib, systemic IL-6 levels or could be crucial to the crosstalk between the JAK/STAT and NF-κB pathway as it can repress NF-κB activity in vitro(309).

Finally, long- non-coding RNA were upregulated by synovial fluid stimulation in CD4 T-cells and some were decreased by tofacitinib. However due to time, resource and sample constraints I was unable to validate these. In further work, I would go on to prioritise non-coding species in areas with SNPs conferring risk to RA and where epigenetic markers suggest functional relevance and validate these in cells derived from patients. This approach of combining epigenetics, genetics and transcriptomics has led to novel findings in pathogenesis of RA(304,306,317).

This progress would not have been possible without the "Big Data" available from public consortia, the bioinformatic tools that are available and also patient derived samples that are interrogated using high throughput technology such as RNA sequencing. With cell subset interrogation becoming more important to fully understand disease pathogenesis, I believe that single cell analysis using all three 'omics described above will lead to society achieving the goal of getting the right drug to the right patient at the right time.

In conclusion, tofacitinib reduces pro-inflammatory cytokine secretion in a macrophage cell contact activation assay suggesting that it affects both the innate and adaptive immune system. Furthermore, soluble factors present in RA synovial fluid stimulate CD4 T-cells and this has an effect on genes that are regulated by STAT3. Finally, RNA sequencing shows that genes such as SBNO2 are differentially expressed in a fashion similar to STAT3 associated genes and that long non-coding RNA may have an important role as biomarkers of response or in the pathogenesis of RA.

References

1. Crowson CS, Matteson EL, Myasoedova E, Michet CJ, Ernste FC, Warrington KJ, et al. The lifetime risk of adult-onset rheumatoid arthritis and other inflammatory autoimmune rheumatic diseases. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2011 Mar;63(3):633–9.

2. Matcham F, Rayner L, Steer S, Hotopf M. The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis. Rheumatology (Oxford). Oxford University Press; 2013 Dec;52(12):2136–48.

3. Tobón GJ, Youinou P, Saraux A. The environment, geo-epidemiology, and autoimmune disease: Rheumatoid arthritis. Autoimmun Rev. 2010 Mar;9(5):A288–92.

4. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. Arthritis Res. 2002;4 Suppl 3:S265–72.

5. Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2008 Jan;58(1):15–25.

6. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. 2014 Feb 20;506(7488):376–81.

7. Stastny P. Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. N Engl J Med. 1978 Apr 20;298(16):869–71.

8. Klareskog L, Forsum U, Scheynius A, Kabelitz D, Wigzell H. Evidence in support of a self-perpetuating HLA-DR-dependent delayed-type cell reaction in rheumatoid arthritis. Proc Natl Acad Sci USA. National Academy of Sciences; 1982 Jun;79(11):3632–6.

9. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum. 1987 Nov;30(11):1205–13.

10. Klareskog L, Stolt P, Lundberg K, Källberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2006 Jan;54(1):38–46.

11. Klareskog L, Malmström V, Lundberg K, Padyukov L, Alfredsson L. Smoking, citrullination and genetic variability in the immunopathogenesis of rheumatoid arthritis. Semin Immunol. 2011 Apr;23(2):92–8.

12. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. N Engl J Med. 2011 Dec 8;365(23):2205–19.

13. Morgan AW, Thomson W, Martin SG, Yorkshire Early Arthritis Register

Consortium, Carter AM, UK Rheumatoid Arthritis Genetics Consortium, et al. Reevaluation of the interaction between HLA-DRB1 shared epitope alleles, PTPN22, and smoking in determining susceptibility to autoantibody-positive and autoantibody-negative rheumatoid arthritis in a large UK Caucasian population. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2009 Sep;60(9):2565–76.

14. Costenbader KH, Feskanich D, Mandl LA, Karlson EW. Smoking intensity, duration, and cessation, and the risk of rheumatoid arthritis in women. Am J Med. Elsevier; 2006 Jun;119(6):503.e1–9.

15. Demoruelle MK, Weisman MH, Simonian PL, Lynch DA, Sachs PB, Pedraza IF, et al. Brief report: airways abnormalities and rheumatoid arthritis-related autoantibodies in subjects without arthritis: early injury or initiating site of autoimmunity? Arthritis Rheum. 2012 Jun;64(6):1756–61.

16. Blanc PD, Järvholm B, Torén K. Prospective Risk of Rheumatologic Disease Associated with Occupational Exposure in a Cohort of Male Construction Workers. Am J Med. 2015 Oct;128(10):1094–101.

17. Schreiber J, Koschel D, Kekow J, Waldburg N, Goette A, Merget R. Rheumatoid pneumoconiosis (Caplan's syndrome). Eur J Intern Med. Elsevier; 2010 Jun;21(3):168–72.

18. Reynisdottir G, Karimi R, Joshua V, Olsen H, Hensvold AH, Harju A, et al. Structural changes and antibody enrichment in the lungs are early features of anti-citrullinated protein antibody-positive rheumatoid arthritis. Arthritis & rheumatology (Hoboken, NJ). 2014 Jan;66(1):31–9.

19. Abhishek A, Butt S, Gadsby K, Zhang W, Deighton CM. Anti-TNF-alpha agents are less effective for the treatment of rheumatoid arthritis in current smokers. J Clin Rheumatol. 2010 Jan;16(1):15–8.

20. McEntegart A, Morrison E, Capell HA, Duncan MR, Porter D, Madhok R, et al. Effect of social deprivation on disease severity and outcome in patients with rheumatoid arthritis. Ann Rheum Dis. BMJ Group; 1997 Jul;56(7):410–3.

21. Harrison MJ, Tricker KJ, Davies L, Hassell A, Dawes P, Scott DL, et al. The relationship between social deprivation, disease outcome measures, and response to treatment in patients with stable, long-standing rheumatoid arthritis. The Journal of Rheumatology. 2005 Dec;32(12):2330–6.

22. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum. 1988 Mar;31(3):315–24.

23. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2010 Sep;69(9):1580–8.

24. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts

EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis. BMJ Group; 1990 Nov;49(11):916–20.

25. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology. 2003 Feb;42(2):244–57.

26. Aletaha D, Nell VPK, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther. BioMed Central Ltd; 2005;7(4):R796–806.

27. Aletaha D, Smolen JS. The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. Best Pract Res Clin Rheumatol. Elsevier; 2007 Aug;21(4):663–75.

28. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum. 1995 Jun;38(6):727–35.

29. Van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. Arthritis Rheum. 1996 Jan;39(1):34–40.

30. Wegner N, Lundberg K, Kinloch A, Fisher B, Malmström V, Feldmann M, et al. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. Immunol Rev. Blackwell Publishing Ltd; 2010 Jan;233(1):34–54.

31. Majka DS, Deane KD, Parrish LA, Lazar AA, Barón AE, Walker CW, et al. Duration of preclinical rheumatoid arthritis-related autoantibody positivity increases in subjects with older age at time of disease diagnosis. Ann Rheum Dis. 2008 Jun;67(6):801–7.

32. Brink M, Hansson M, Mathsson L, Jakobsson P-J, Holmdahl R, Hallmans G, et al. Multiplex analyses of antibodies against citrullinated peptides in individuals prior to development of rheumatoid arthritis. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2013 Apr;65(4):899–910.

33. Kokkonen H, Brink M, Hansson M, Lassen E, Mathsson-Alm L, Holmdahl R, et al. Associations of antibodies against citrullinated peptides with human leukocyte antigen-shared epitope and smoking prior to the development of rheumatoid arthritis. Arthritis Res Ther. BioMed Central Ltd; 2015;17(1):125.

34. Harre U, Georgess D, Bang H, Bozec A, Axmann R, Ossipova E, et al. Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin. J Clin Invest. American Society for Clinical Investigation; 2012 May;122(5):1791–802.

35. Sokolove J, Zhao X, Chandra PE, Robinson WH. Immune complexes containing citrullinated fibrinogen costimulate macrophages via Toll-like receptor 4 and Fcγ receptor. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2011 Jan;63(1):53–62.

36. Thabet MM, Huizinga TWJ, van der Heijde DM, van der Helm-van Mil AHM. The prognostic value of baseline erosions in undifferentiated arthritis. Arthritis Res Ther. BioMed Central Ltd; 2009;11(5):R155.

37. Kennedy T, McCabe C, Struthers G, Sinclair H, Chakravaty K, Bax D, et al. BSR guidelines on standards of care for persons with rheumatoid arthritis. Vol. 44, Rheumatology. 2005. pp. 553–6.

38. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. Lancet. 2004;364(9430):263–9.

39. Landewé RBM, Boers M, Verhoeven AC, Westhovens R, van de Laar MAFJ, Markusse HM, et al. COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. Arthritis Rheum. 2002 Feb;46(2):347–56.

40. O'Dell JR, Curtis JR, Mikuls TR, Cofield SS, Bridges SL, Ranganath VK, et al. Validation of the methotrexate-first strategy in patients with early, poor-prognosis rheumatoid arthritis: results from a two-year randomized, double-blind trial. Arthritis Rheum. 2013 Aug;65(8):1985–94.

41. Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, et al. Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. ATTRACT Study Group. The Lancet. 1999 Dec 4;354(9194):1932–9.

42. Weinblatt ME, Kremer JM, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Fox RI, et al. A trial of etanercept, a recombinant tumor necrosis factor receptor:Fc fusion protein, in patients with rheumatoid arthritis receiving methotrexate. N Engl J Med. 1999 Jan 28;340(4):253–9.

43. Elliott MJ, Maini RN, Feldmann M, Kalden JR, Antoni C, Smolen JS, et al. Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor alpha (cA2) versus placebo in rheumatoid arthritis. The Lancet. 1994 Oct 22;344(8930):1105–10.

44. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. Lancet. 2004 Feb 28;363(9410):675–81.

45. O'Dell JR, Mikuls TR, Taylor TH, Ahluwalia V, Brophy M, Warren SR, et al. Therapies for active rheumatoid arthritis after methotrexate failure. N Engl J Med. 2013 Jul 25;369(4):307–18.

46. Galloway JB, Mercer LK, Moseley A, Dixon WG, Ustianowski AP, Helbert

M, et al. Risk of skin and soft tissue infections (including shingles) in patients exposed to anti-tumour necrosis factor therapy: results from the British Society for Rheumatology Biologics Register. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2013 Feb;72(2):229–34.

47. Mercer LK, Lunt M, Low ALS, Dixon WG, Watson KD, Symmons DPM, et al. Risk of solid cancer in patients exposed to anti-tumour necrosis factor therapy: results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2015 Jun;74(6):1087–93.

48. Arkema EV, Jonsson J, Baecklund E, Bruchfeld J, Feltelius N, Askling J, et al. Are patients with rheumatoid arthritis still at an increased risk of tuberculosis and what is the role of biological treatments? Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2015 Jun;74(6):1212–7.

49. Graudal NA, Svenson M, Tarp U, Garred P, Jurik A-G, Bendtzen K. Autoantibodies against interleukin 1alpha in rheumatoid arthritis: association with long term radiographic outcome. Ann Rheum Dis. BMJ Group; 2002 Jul;61(7):598–602.

50. Cohen SB, Moreland LW, Cush JJ, Greenwald MW, Block S, Shergy WJ, et al. A multicentre, double blind, randomised, placebo controlled trial of anakinra (Kineret), a recombinant interleukin 1 receptor antagonist, in patients with rheumatoid arthritis treated with background methotrexate. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2004 Sep;63(9):1062–8.

51. So A, De Meulemeester M, Pikhlak A, Yücel AE, Richard D, Murphy V, et al. Canakinumab for the treatment of acute flares in difficult-to-treat gouty arthritis: Results of a multicenter, phase II, dose-ranging study. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2010 Oct;62(10):3064–76.

52. Emery P, Keystone E, Tony HP, Cantagrel A, Van Vollenhoven R, Sanchez A, et al. IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2008 Nov;67(11):1516–23.

53. Gabay C, Emery P, van Vollenhoven R, Dikranian A, Alten R, Pavelka K, et al. Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (ADACTA): a randomised, double-blind, controlled phase 4 trial. Lancet. 2013 May 4;381(9877):1541–50.

54. Genovese MC, McKay JD, Nasonov EL, Mysler EF, da Silva NA, Alecock E, et al. Interleukin-6 receptor inhibition with tocilizumab reduces disease activity in rheumatoid arthritis with inadequate response to disease-modifying antirheumatic drugs: the tocilizumab in combination with traditional disease-modifying antirheumatic drug therapy study. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2008 Oct;58(10):2968–80.

55. Noss EH, Nguyen HN, Chang SK, Watts GFM, Brenner MB. Genetic polymorphism directs IL-6 expression in fibroblasts but not selected other cell types. Proc Natl Acad Sci USA. National Acad Sciences; 2015 Dec 1;112(48):14948–53.

56. Muraguchi A, Hirano T, Tang B, Matsuda T, Horii Y, Nakajima K, et al. The essential role of B cell stimulatory factor 2 (BSF-2/IL-6) for the terminal differentiation of B cells. J Exp Med. The Rockefeller University Press; 1988 Feb 1;167(2):332–44.

57. Cheung MC, Haynes AE, Meyer RM, Stevens A, Imrie KR, Members of the Hematology, Disease Site Group of the Cancer Care Ontario Program in Evidence-Based Care. Rituximab in lymphoma: a systematic review and consensus practice guideline from Cancer Care Ontario. Cancer Treat Rev. 2007 Apr;33(2):161–76.

58. Cohen SB, Emery P, Greenwald MW, Dougados M, Furie RA, Genovese MC, et al. Rituximab for rheumatoid arthritis refractory to anti-tumor necrosis factor therapy: Results of a multicenter, randomized, double-blind, placebo-controlled, phase III trial evaluating primary efficacy and safety at twenty-four weeks. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2006 Sep;54(9):2793–806.

59. Edwards JCW, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR, et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med. 2004 Jun 17;350(25):2572–81.

60. Kremer JM, Genant HK, Moreland LW, Russell AS, Emery P, Abud-Mendoza C, et al. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: a randomized trial. Ann Intern Med. 2006 Jun 20;144(12):865–76.

61. Baslund B, Tvede N, Danneskiold-Samsøe B, Larsson P, Panayi G, Petersen J, et al. Targeting interleukin-15 in patients with rheumatoid arthritis: a proof-of-concept study. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2005 Sep;52(9):2686–92.

62. Burmester GR, Weinblatt ME, McInnes IB, Porter D, Barbarash O, Vatutin M, et al. Efficacy and safety of mavrilimumab in subjects with rheumatoid arthritis. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2013 Sep 1;72(9):1445–52.

63. Banna GL, Collovà E, Gebbia V, Lipari H, Giuffrida P, Cavallaro S, et al. Anticancer oral therapy: emerging related issues. Cancer Treat Rev. 2010 Dec 1;36(8):595–605.

64. O'Brien SG, Guilhot F, Larson RA, Gathmann I, Baccarani M, Cervantes F, et al. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. N Engl J Med. 2003 Mar 13;348(11):994–1004.

65. Blay J-Y, Sayadi El H, Thiesse P, Garret J, Ray-Coquard I. Complete response to imatinib in relapsing pigmented villonodular synovitis/tenosynovial giant

cell tumor (PVNS/TGCT). Ann Oncol. 2008 Apr 1;19(4):821–2.

66. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science. 2002 Dec 6;298(5600):1912–34.

67. Ghoreschi K, Laurence A, O'Shea JJ. Selectivity and therapeutic inhibition of kinases: to be or not to be? Nat Immunol. 2009 Apr 1;10(4):356–60.

68. Chartier M, Chénard T, Barker J, Najmanovich R. Kinome Render: a stand-alone and web-accessible tool to annotate the human protein kinome tree. PeerJ. 2013.

69. Fleischmann R, Kremer J, Cush J, Schulze-Koops H, Connell CA, Bradley JD, et al. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. N Engl J Med. 2012 Aug 9;367(6):495–507.

70. Genovese MC, Kavanaugh A, Weinblatt ME, Peterfy C, DiCarlo J, White ML, et al. An oral Syk kinase inhibitor in the treatment of rheumatoid arthritis: a three-month randomized, placebo-controlled, phase II study in patients with active rheumatoid arthritis that did not respond to biologic agents. Arthritis Rheum. 2011 Feb 1;63(2):337–45.

71. Nijjar JS, Tindell A, McInnes IB, Siebert S. Inhibition of spleen tyrosine kinase in the treatment of rheumatoid arthritis. Rheumatology (Oxford). 2013 Sep;52(9):1556–62.

72. Karaman MW, Herrgard S, Treiber DK, Gallant P, Atteridge CE, Campbell BT, et al. A quantitative analysis of kinase inhibitor selectivity. Nat Biotechnol. 2008;26(1):127–32.

73. Ghoreschi K, Jesson MI, Li X, Lee JL, Ghosh S, Alsup JW, et al. Modulation of innate and adaptive immune responses by tofacitinib (CP-690,550). The Journal of Immunology. 2011 Apr 1;186(7):4234–43.

74. Rosengren S, Corr M, Firestein GS, Boyle DL. The JAK inhibitor CP-690,550 (tofacitinib) inhibits TNF-induced chemokine expression in fibroblast-like synoviocytes: autocrine role of type I interferon. Ann Rheum Dis. 2012 Mar;71(3):440–7.

75. van Beuningen HM, de Vries-van Melle ML, Vitters EL, Schreurs W, van den Berg WB, van Osch GJVM, et al. Inhibition of TAK1 and/or JAK Can Rescue Impaired Chondrogenic Differentiation of Human Mesenchymal Stem Cells in Osteoarthritis-Like Conditions. Tissue Eng Part A. 2014 Mar 25;:140325123546005.

76. Kubo S, Yamaoka K, Kondo M, Yamagata K, Zhao J, Iwata S, et al. The JAK inhibitor, tofacitinib, reduces the T cell stimulatory capacity of human monocyte-derived dendritic cells. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2014 Dec;73(12):2192–8.

77. Moisan A, Lee Y-K, Zhang JD, Hudak CS, Meyer CA, Prummer M, et al. White-to-brown metabolic conversion of human adipocytes by JAK inhibition. Nat Cell Biol. 2015 Jan;17(1):57–67.

78. Keystone EC, Taylor PC, Drescher E, Schlichting DE, Beattie SD, Berclaz P-Y, et al. Safety and efficacy of baricitinib at 24 weeks in patients with rheumatoid arthritis who have had an inadequate response to methotrexate. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2015 Feb;74(2):333–40.

79. Fleischmann RM, Damjanov NS, Kivitz AJ, Legedza A, Hoock T, Kinnman N. A Randomized, Double-Blind, Placebo-Controlled, Twelve-Week, Dose-Ranging Study of Decernotinib, an Oral Selective JAK-3 Inhibitor, as Monotherapy in Patients With Active Rheumatoid Arthritis. Arthritis & Rheumatology. 2015 Jan 28;67(2):334–43.

80. Whang JA, Chang BY. Bruton's tyrosine kinase inhibitors for the treatment of rheumatoid arthritis. Drug Discovery Today. 2014 Aug;19(8):1200–4.

81. Rossi AG, Sawatzky DA, Walker A, Ward C, Sheldrake TA, Riley NA, et al. Cyclin-dependent kinase inhibitors enhance the resolution of inflammation by promoting inflammatory cell apoptosis. Nat Med. Nature Publishing Group; 2006 Sep;12(9):1056–64.

82. Bartok B, Hammaker D, Firestein GS. Phosphoinositide 3-kinase δ regulates migration and invasion of synoviocytes in rheumatoid arthritis. The Journal of Immunology. American Association of Immunologists; 2014 Mar 1;192(5):2063–70.

83. Camps M, Rückle T, Ji H, Ardissone V, Rintelen F, Shaw J, et al. Blockade of PI3Kgamma suppresses joint inflammation and damage in mouse models of rheumatoid arthritis. Nat Med. Nature Publishing Group; 2005 Sep;11(9):936–43.

84. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet. Nature Publishing Group; 2010 Jun;42(6):508–14.

85. Eyre S, Bowes J, Diogo D, Lee A, Barton A, Martin P, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet. Nature Publishing Group; 2012 Dec;44(12):1336–40.

86. Aho K, Koskenvuo M, Tuominen J, Kaprio J. Occurrence of rheumatoid arthritis in a nationwide series of twins. The Journal of Rheumatology. 1986 Oct;13(5):899–902.

87. Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A, et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol. 1993 Oct;32(10):903–7.

88. Carlton VEH, Hu X, Chokkalingam AP, Schrodi SJ, Brandon R, Alexander HC, et al. PTPN22 genetic variation: evidence for multiple variants associated with rheumatoid arthritis. Am J Hum Genet. 2005 Oct;77(4):567–81.

89. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med. 2007 Sep 6;357(10):977–86.

90. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. 2013.

91. Vassallo R, Luckey D, Behrens M, Madden B, Luthra H, David C, et al. Cellular and humoral immunity in arthritis are profoundly influenced by the interaction between cigarette smoke effects and host HLA-DR and DQ genes. Clin Immunol. 2014 May;152(1-2):25–35.

92. Adachi M, Okamoto S, Chujyo S, Arakawa T, Yokoyama M, Yamada K, et al. Cigarette smoke condensate extracts induce IL-1-beta production from rheumatoid arthritis patient-derived synoviocytes, but not osteoarthritis patient-derived synoviocytes, through aryl hydrocarbon receptor-dependent NF-kappa-B activation and novel NF-kappa-B sites. Journal of Interferon & Cytokine Research. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA; 2013 Jun;33(6):297–307.

93. Couderc T, Lecuit M. Chikungunya virus pathogenesis: From bedside to bench. Antiviral Res. 2015 Sep;121:120–31.

94. Rothschild BM, Woods RJ, Rothschild C, Sebes JI. Geographic distribution of rheumatoid arthritis in ancient North America: implications for pathogenesis. Semin Arthritis Rheum. 1992 Dec;22(3):181–7.

95. Wegner N, Wait R, Sroka A, Eick S, Nguyen K-A, Lundberg K, et al. Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and α-enolase: implications for autoimmunity in rheumatoid arthritis. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2010 Sep;62(9):2662–72.

96. Taurog JD, Richardson JA, Croft JT, Simmons WA, Zhou M, Fernández-Sueiro JL, et al. The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. J Exp Med. The Rockefeller University Press; 1994 Dec 1;180(6):2359–64.

97. Goto Y, Panea C, Nakato G, Cebula A, Lee C, Diez MG, et al. Segmented filamentous bacteria antigens presented by intestinal dendritic cells drive mucosal Th17 cell differentiation. Immunity. 2014 Apr 17;40(4):594–607.

98. Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. Nat Med. 2015.

99. Dennis G, Holweg CT, Kummerfeld SK, Choy DF, Setiadi AF, Hackney JA, et al. Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics. Arthritis Res Ther. BioMed Central Ltd; 2014 Apr 30;16(2):R90.

100. Tak PP, Smeets TJ, Daha MR, Kluin PM, Meijers KA, Brand R, et al. Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. Arthritis Rheum. 1997 Feb;40(2):217–25.

101. Haringman JJ, Gerlag DM, Zwinderman AH, Smeets TJM, Kraan MC, Baeten D, et al. Synovial tissue macrophages: a sensitive biomarker for response to treatment in patients with rheumatoid arthritis. Ann Rheum Dis. 2005 Jun;64(6):834–8.

102. Geissmann F, Jung S, Littman DR. Blood monocytes consist of two principal subsets with distinct migratory properties. Immunity. 2003 Jul;19(1):71–82.

103. Kawanaka N, Yamamura M, Aita T, Morita Y, Okamoto A, Kawashima M, et al. CD14+,CD16+ blood monocytes and joint inflammation in rheumatoid arthritis. Arthritis Rheum. 2002 Oct;46(10):2578–86.

104. Cairns AP, Crockard AD, Bell AL. The CD14+ CD16+ monocyte subset in rheumatoid arthritis and systemic lupus erythematosus. Rheumatol Int. 2002 Mar;21(5):189–92.

105. van Furth R, Cohn ZA. The origin and kinetics of mononuclear phagocytes. J Exp Med. The Rockefeller University Press; 1968 Sep 1;128(3):415–35.

106. van Furth R, Cohn ZA, Hirsch JG, Humphrey JH, Spector WG, Langevoort HL. The mononuclear phagocyte system: a new classification of macrophages, monocytes, and their precursor cells. Bull World Health Organ. 1972;46(6):845–52.

107. Stout RD, Jiang C, Matta B, Tietzel I, Watkins SK, Suttles J. Macrophages sequentially change their functional phenotype in response to changes in microenvironmental influences. J Immunol. 2005 Jul 1;175(1):342–9.

108. Hoeffel G, Wang Y, Greter M, See P, Teo P, Malleret B, et al. Adult Langerhans cells derive predominantly from embryonic fetal liver monocytes with a minor contribution of yolk sac-derived macrophages. J Exp Med. Rockefeller Univ Press; 2012 Jun 4;209(6):1167–81.

109. Yona S, Kim K-W, Wolf Y, Mildner A, Varol D, Breker M, et al. Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. Immunity. Elsevier; 2013 Jan 24;38(1):79–91.

110. Jongbloed SL, Benson RA, Nickdel MB, Garside P, McInnes IB, Brewer JM. Plasmacytoid dendritic cells regulate breach of self-tolerance in autoimmune arthritis. J Immunol. 2009 Jan 15;182(2):963–8.

111. Lebre MC, Tak PP. Dendritic cells in rheumatoid arthritis: Which subset should be used as a tool to induce tolerance? Hum Immunol. 2009 May 1;70(5):321–4.

112. Leung BP, Conacher M, Hunter D, McInnes IB, Liew FY, Brewer JM. A novel dendritic cell-induced model of erosive inflammatory arthritis: distinct roles for dendritic cells in T cell activation and induction of local inflammation. J Immunol. 2002 Dec 15;169(12):7071–7.

113. Walker JG, Ahern MJ, Coleman M, Weedon H, Papangelis V, Beroukas D, et al. Characterisation of a dendritic cell subset in synovial tissue which strongly expresses Jak/STAT transcription factors from patients with rheumatoid arthritis. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2007 Aug;66(8):992–9.

114. Steinman RM, Nussenzweig MC. Avoiding horror autotoxicus: the importance of dendritic cells in peripheral T cell tolerance. Proc Natl Acad Sci USA. National Acad Sciences; 2002 Jan 8;99(1):351–8.

115. Vital EM, Emery P. Abatacept in the treatment of rheumatoid arthritis. Ther Clin Risk Manag. Dove Press; 2006 Dec;2(4):365–75.

116. Brennan FM, Smith NMG, Owen S, Li C, Amjadi P, Green P, et al. Resting CD4+ effector memory T cells are precursors of bystander-activated effectors: a surrogate model of rheumatoid arthritis synovial T-cell function. Arthritis Res Ther. 2008;10(2):R36.

117. Brennan FM, Hayes AL, Ciesielski CJ, Green P, Foxwell BMJ, Feldmann M. Evidence that rheumatoid arthritis synovial T cells are similar to cytokine-activated T cells: involvement of phosphatidylinositol 3-kinase and nuclear factor kappaB pathways in tumor necrosis factor alpha production in rheumatoid arthritis. Arthritis Rheum. John Wiley & Sons, Inc; 2002 Jan;46(1):31–41.

118. Humby F, Bombardieri M, Manzo A, Kelly S, Blades MC, Kirkham B, et al. Ectopic lymphoid structures support ongoing production of class-switched autoantibodies in rheumatoid synovium. Huizinga T, editor. PLoS Med. 2009 Jan 13;6(1):e1.

119. Miossec P, Korn T, Kuchroo VK. Interleukin-17 and type 17 helper T cells. N Engl J Med. 2009 Aug 27;361(9):888–98.

120. Nadkarni S, Mauri C, Ehrenstein MR. Anti-TNF-alpha therapy induces a distinct regulatory T cell population in patients with rheumatoid arthritis via TGF-beta. J Exp Med. Rockefeller Univ Press; 2007 Jan 22;204(1):33–9.

121. McInnes IB, Leung BP, Liew FY. Cell-cell interactions in synovitis. Interactions between T lymphocytes and synovial cells. Arthritis Res. BioMed Central Ltd; 2000;2(5):374–8.

122. Carvalheiro H, Duarte C, Silva-Cardoso S, da Silva JAP, Souto-Carneiro MM. CD8+ T Cell Profiles in Patients With Rheumatoid Arthritis and Their Relationship to Disease Activity. Arthritis & Rheumatology. 2015 Jan 28;67(2):363–71.

123. Yeo L, Lom H, Juarez M, Snow M, Buckley CD, Filer A, et al. Expression of FcRL4 defines a pro-inflammatory, RANKL-producing B cell subset in rheumatoid arthritis. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2015 May;74(5):928–35.

124. Cui D, Zhang L, Chen J, Zhu M, Hou L, Chen B, et al. Changes in regulatory B cells and their relationship with rheumatoid arthritis disease activity. Clin Exp Med. Springer International Publishing; 2015 Aug;15(3):285–92.

125. Raza K, Scheel-Toellner D, Lee C-Y, Pilling D, Curnow SJ, Falciani F, et al. Synovial fluid leukocyte apoptosis is inhibited in patients with very early rheumatoid arthritis. Arthritis Res Ther. BioMed Central Ltd; 2006;8(4):R120.

126. Ross EA, Douglas MR, Wong SH, Ross EJ, Curnow SJ, Nash GB, et al. Interaction between integrin alpha9beta1 and vascular cell adhesion molecule-1 (VCAM-1) inhibits neutrophil apoptosis. Blood. American Society of Hematology; 2006 Feb 1;107(3):1178–83.

127. Buckley CD, Ross EA, McGettrick HM, Osborne CE, Haworth O, Schmutz C,

et al. Identification of a phenotypically and functionally distinct population of long-lived neutrophils in a model of reverse endothelial migration. Journal of Leukocyte Biology. Society for Leukocyte Biology; 2006 Feb;79(2):303–11.

128. Khandpur R, Carmona-Rivera C, Vivekanandan-Giri A, Gizinski A, Yalavarthi S, Knight JS, et al. NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis. Sci Transl Med. 2013 Mar 27;5(178):178ra40.

129. Aggarwal A, Sharma A, Bhatnagar A. Bi(o)communications among peripheral blood fractions: a focus on NK and NKT cell biology in rheumatoid arthritis. Autoimmunity. 2013 Jun;46(4):238–50.

130. Filer A. The fibroblast as a therapeutic target in rheumatoid arthritis. Curr Opin Pharmacol. 2013 Jun;13(3):413–9.

131. Shigeyama Y, Pap T, Kunzler P, Simmen BR, Gay RE, Gay S. Expression of osteoclast differentiation factor in rheumatoid arthritis. Arthritis Rheum. John Wiley & Sons, Inc; 2000 Nov;43(11):2523–30.

132. Li BT, Zhang FZ, Xu TS, Ding R, Li P. Increasing production of matrix metalloproteinases, tumor necrosis factor-α, vascular endothelial growth factor and prostaglandin E2 in rheumatoid arthritis synovial fibroblasts by different adiponectin isoforms in a concentration-dependent manner. Cell Mol Biol (Noisy-le-grand). 2015;61(7):27–32.

133. McGettrick HM, Buckley CD, Filer A, Rainger GE, Nash GB. Stromal cells differentially regulate neutrophil and lymphocyte recruitment through the endothelium. Immunology. Blackwell Publishing Ltd; 2010 Nov;131(3):357–70.

134. Lefèvre S, Knedla A, Tennie C, Kampmann A, Wunrau C, Dinser R, et al. Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. Nat Med. 2009 Dec;15(12):1414–20.

135. Sakkas LI, Bogdanos DP, Katsiari C, Platsoucas CD. Anti-citrullinated peptides as autoantigens in rheumatoid arthritis-relevance to treatment. Autoimmun Rev. 2014 Nov;13(11):1114–20.

136. Diarra D, Stolina M, Polzer K, Zwerina J, Ominsky MS, Dwyer D, et al. Dickkopf-1 is a master regulator of joint remodeling. Nat Med. 2007 Feb;13(2):156–63.

137. de Rooy DPC, Yeremenko NG, Wilson AG, Knevel R, Lindqvist E, Saxne T, et al. Genetic studies on components of the Wnt signalling pathway and the severity of joint destruction in rheumatoid arthritis. Ann Rheum Dis. 2013 Apr 4;72(5):769–75.

138. Tetlow LC, Woolley DE. Effect of histamine on the production of matrix metalloproteinases-1, -3, -8 and -13, and TNFalpha and PGE(2) by human articular chondrocytes and synovial fibroblasts in vitro: a comparative study. Virchows Arch. 2004 Nov;445(5):485–90.

139. Tetlow LC, Adlam DJ, Woolley DE. Matrix metalloproteinase and proinflammatory cytokine production by chondrocytes of human osteoarthritic

cartilage: associations with degenerative changes. Arthritis Rheum. John Wiley & Sons, Inc; 2001 Mar;44(3):585–94.

140. Woolley DE, Tetlow LC. Observations on the microenvironmental nature of cartilage degradation in rheumatoid arthritis. Ann Rheum Dis. 1997 Mar;56(3):151–61.

141. Masters SL, Simon A, Aksentijevich I, Kastner DL. Horror autoinflammaticus: the molecular pathophysiology of autoinflammatory disease (*). Annu Rev Immunol. Annual Reviews; 2009;27(1):621–68.

142. McInnes IB, Schett G. Cytokines in the pathogenesis of rheumatoid arthritis. Nat Rev Immunol. 2007 Jun 1;7(6):429–42.

143. Wallace DJ, Gavin IM, Karpenko O, Barkhordar F, Gillis BS. Cytokine and chemokine profiles in fibromyalgia, rheumatoid arthritis and systemic lupus erythematosus: a potentially useful tool in differential diagnosis. Rheumatol Int. Springer Berlin Heidelberg; 2015 Jun;35(6):991–6.

144. Zhao B, Grimes SN, Li S, Hu X, Ivashkiv LB. TNF-induced osteoclastogenesis and inflammatory bone resorption are inhibited by transcription factor RBP-J. J Exp Med. 2012 Feb 13;209(2):319–34.

145. Bruns H, Meinken C, Schauenberg P, Härter G, Kern P, Modlin RL, et al. Anti-TNF immunotherapy reduces CD8+ T cell-mediated antimicrobial activity against Mycobacterium tuberculosis in humans. J Clin Invest. American Society for Clinical Investigation; 2009 May;119(5):1167–77.

146. Mease PJ, Goffe BS, Metz J, VanderStoep A, Finck B, Burge DJ. Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. The Lancet. Elsevier; 2000 Jul 29;356(9227):385–90.

147. Hanauer SB, Sandborn WJ, Rutgeerts P, Fedorak RN, Lukas M, MacIntosh D, et al. Human anti-tumor necrosis factor monoclonal antibody (adalimumab) in Crohn's disease: the CLASSIC-I trial. Gastroenterology. Elsevier; 2006 Feb;130(2):323–33–quiz591.

148. Gattorno M, Federici S, Pelagatti MA, Caorsi R, Brisca G, Malattia C, et al. Diagnosis and management of autoinflammatory diseases in childhood. J Clin Immunol. 2008 May 1;28 Suppl 1:S73–83.

149. Cornish AL, Campbell IK, McKenzie BS, Chatfield S, Wicks IP. G-CSF and GM-CSF as therapeutic targets in rheumatoid arthritis. Nat Rev Rheumatol. 2009 Oct;5(10):554–9.

150. Meier FMP, McInnes IB. Small-molecule therapeutics in rheumatoid arthritis: scientific rationale, efficacy and safety. Best Pract Res Clin Rheumatol. 2014 Aug;28(4):605–24.

151. van der Heijde D, Tanaka Y, Fleischmann R, Keystone E, Kremer J, Zerbini C, et al. Tofacitinib (CP-690,550) in patients with rheumatoid arthritis receiving methotrexate: twelve-month data from a twenty-four-month phase III randomized radiographic study. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2013 Mar;65(3):559–70.

152. Kremer JM, Cohen S, Wilkinson BE, Connell CA, French JL, Gomez-Reino J, et al. A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. Arthritis Rheum. 2012 Apr 1;64(4):970–81.

153. Migita K, Komori A, Torigoshi T, Maeda Y, Izumi Y, Jiuchi Y, et al. CP690,550 inhibits oncostatin M-induced JAK/STAT signaling pathway in rheumatoid synoviocytes. Arthritis Res Ther. 2011;13(3):R72.

154. Walker JG, Ahern MJ, Coleman M, Weedon H, Papangelis V, Beroukas D, et al. Expression of Jak3, STAT1, STAT4, and STAT6 in inflammatory arthritis: unique Jak3 and STAT4 expression in dendritic cells in seropositive rheumatoid arthritis. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2006 Feb;65(2):149–56.

155. Isomäki P, Junttila I, Vidqvist K-L, Korpela M, Silvennoinen O. The activity of JAK-STAT pathways in rheumatoid arthritis: constitutive activation of STAT3 correlates with interleukin 6 levels. Rheumatology (Oxford). 2015 Jun;54(6):1103–13.

156. Levy DE, Darnell JE. Stats: transcriptional control and biological impact. Nat Rev Mol Cell Biol. 2002 Sep;3(9):651–62.

157. Levy DE, Marié IJ. STATus report on tetramers. Immunity. Elsevier; 2012 Apr 20;36(4):553–5.

158. Ghoreschi K, Gadina M. Jakpot! new small molecules in autoimmune and inflammatory diseases. Exp Dermatol. 2014 Jan;23(1):7–11.

159. Carow B, Rottenberg ME. SOCS3, a Major Regulator of Infection and Inflammation. Front Immunol. Frontiers; 2014;5:58.

160. Liu B, Liao J, Rao X, Kushner SA, Chung CD, Chang DD, et al. Inhibition of Stat1-mediated gene activation by PIAS1. Proc Natl Acad Sci USA. National Academy of Sciences; 1998 Sep 1;95(18):10626–31.

161. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science. American Association for the Advancement of Science; 2006 Dec 1;314(5804):1461–3.

162. Wellcome Trust Case Control Consortium, Australo-Anglo-American Spondylitis Consortium (TASC), Burton PR, Clayton DG, Cardon LR, Craddock N, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet. Nature Publishing Group; 2007 Nov;39(11):1329–37.

163. McInnes IB, Kavanaugh A, Gottlieb AB, Puig L. Efficacy and safety of ustekinumab in patients with active psoriatic arthritis: 1 year results of the phase 3, multicentre, double-blind, placebo-controlled PSUMMIT 1 trial. The Lancet. 2013.

164. Macchi P, Villa A, Giliani S, Sacco MG, Frattini A, Porta F, et al. Mutations of Jak-3 gene in patients with autosomal severe combined immune deficiency

(SCID). Nature. Nature Publishing Group; 1995 Sep 7;377(6544):65–8.

165. Levine RL, Wadleigh M, Cools J, Ebert BL, Wernig G, Huntly BJP, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell. 2005 Apr;7(4):387–97.

166. Baxter EJ, Scott LM, Campbell PJ, East C, Fourouclas N, Swanton S, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. The Lancet. 2005 Mar;365(9464):1054–61.

167. Wollenhaupt J, Silverfield J, Lee EB, Curtis JR, Wood SP, Soma K, et al. Safety and Efficacy of Tofacitinib, an Oral Janus Kinase Inhibitor, for the Treatment of Rheumatoid Arthritis in Open-label, Longterm Extension Studies. The Journal of Rheumatology. The Journal of Rheumatology; 2014 Apr 1;:jrheum.130683.

168. Sonomoto K, Yamaoka K, Kubo S, Hirata S, Fukuyo S, Maeshima K, et al. Effects of tofacitinib on lymphocytes in rheumatoid arthritis: relation to efficacy and infectious adverse events. Rheumatology (Oxford). 2014 May;53(5):914–8.

169. Charles-Schoeman C. Cardiovascular disease and rheumatoid arthritis: an update. Curr Rheumatol Rep. 2012 Oct 1;14(5):455–62.

170. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. Nature Publishing Group. Nature Publishing Group; 2013 Jul 19;12(8):581–94.

171. Costa FF. Big data in biomedicine. Drug Discovery Today. 2014.

172. WATSON JD, CRICK FH. Genetical implications of the structure of deoxyribonucleic acid. Nature. 1953 May 30;171(4361):964–7.

173. WATSON JD, CRICK FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953 Apr 25;171(4356):737–8.

174. FRANKLIN RE, GOSLING RG. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. Nature. 1953 Jul 25;172(4369):156–7.

175. FRANKLIN RE, GOSLING RG. Molecular configuration in sodium thymonucleate. Nature. 1953 Apr 25;171(4356):740–1.

176. Nirenberg M, Leder P, Bernfield M, Brimacombe R, Trupin J, Rottman F, et al. RNA codewords and protein synthesis, VII. On the general nature of the RNA code. Proc Natl Acad Sci USA. National Academy of Sciences; 1965 May;53(5):1161–8.

177. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology. 1975 May 25;94(3):441–8.

178. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA. National Academy of Sciences; 1977 Dec;74(12):5463–7.

179. Trenkmann M, Brock M, Ospelt C, Gay S. Epigenetics in rheumatoid arthritis. Clin Rev Allergy Immunol. 2010 Aug 1;39(1):10–9.

180. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society Series B (…. 1995.

181. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. Nature Publishing Group; 2009 Jan;10(1):57–63.

182. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. Oxford University Press; 2016 May 2;:gkw343.

183. Team RC. R: A Language and Environment for Statistical Computing

. Vienna, Austria. Available from: https://www.R-project.org

184. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004;5(10):R80.

185. Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J. shiny: Web Application Framework for R. 2015 Aug 5. Available from: http://CRAN.R-project.org/package=shiny

186. Bantscheff M, Lemeer S, Savitski M, Kuster B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Anal Bioanal Chem. 2012;404(4):939–65.

187. Stalmach A, Johnsson H, McInnes IB, Husi H, Klein J, Dakna M, et al. Identification of urinary peptide biomarkers associated with rheumatoid arthritis. PLoS ONE. 2014;9(8):e104625.

188. Weljie AM, Dowlatabadi R, Miller BJ, Vogel HJ, Jirik FR. An inflammatory arthritis-associated metabolite biomarker pattern revealed by 1H NMR spectroscopy. J Proteome Res. 2007 Sep 1;6(9):3456–64.

189. Kapoor SR, Filer A, Fitzpatrick MA, Fisher BA, Taylor PC, Buckley CD, et al. Metabolic Profiling Predicts Response to Anti-Tumor Necrosis Factor α Therapy in Patients With Rheumatoid Arthritis. Arthritis Rheum. 2013 Jun 1;65(6):1448–56.

190. Priori R, Scrivo R, Brandt J, Valerio M, Casadei L, Valesini G, et al. Metabolomics in rheumatic diseases: The potential of an emerging methodology for improved patient diagnosis, prognosis, and treatment efficacy. Autoimmun Rev. Elsevier B.V; 2013 Aug 1;12(10):1022–30.

191. Yang R, Chiang N, Oh SF, Serhan CN. Metabolomics-lipidomics of eicosanoids and docosanoids generated by phagocytes. Curr Protoc Immunol. 2011 Nov 1;Chapter 14:Unit14.26.

192. Menni C, Kastenmuller G, Petersen AK, Bell JT, Psatha M, Tsai PC, et al.

Metabolomic markers reveal novel pathways of ageing and early development in human populations. International Journal of Epidemiology. 2013 Sep 3;42(4):1111–9.

193. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, Bouatra S, et al. The human serum metabolome. Flower D, editor. PLoS ONE. Public Library of Science; 2011;6(2):e16957.

194. Giera M, Ioan-Facsinay A, Toes R, Gao F, Dalli J, Deelder AM, et al. Lipid and lipid mediator profiling of human synovial fluid in rheumatoid arthritis patients by means of LC-MS/MS. Biochim Biophys Acta. 2012 Nov 1;1821(11):1415–24.

195. O'Neill LAJ. A critical role for citrate metabolism in LPS signalling. Biochem J. 2011 Sep 15;438(3):e5–6.

196. Creek DJ, Jankevics A, Burgess KEV, Breitling R, Barrett MP. IDEOM: an Excel interface for analysis of LC-MS-based metabolomics data. Bioinformatics. 2012 Apr 1;28(7):1048–9.

197. Scheltema RA, Jankevics A, Jansen RC, Swertz MA, Breitling R. PeakML/mzMatch: a file format, Java library, R library, and tool-chain for mass spectrometry data analysis. Analytical chemistry. 2011 Apr 1;83(7):2786–93.

198. Xia J, Mandal R, Sinelnikov IV, Broadhurst D, Wishart DS. MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis. Nucleic Acids Res. 2012 Jul 1;40(Web Server issue):W127–33.

199. Wang Z, Chen Z, Yang S, Wang Y, Yu L, Zhang B, et al. (1)H NMR-based metabolomic analysis for identifying serum biomarkers to evaluate methotrexate treatment in patients with early rheumatoid arthritis. Exp Ther Med. 2012 Jul 1;4(1):165–71.

200. Lauridsen MB, Bliddal H, Christensen R, Danneskiold-Samsøe B, Bennett R, Keun H, et al. 1H NMR spectroscopy-based interventional metabolic phenotyping: a cohort study of rheumatoid arthritis patients. J Proteome Res. 2010 Sep 3;9(9):4545–53.

201. Gligorijević V, Malod-Dognin N, Pržulj N. Integrative methods for analyzing big data in precision medicine. Dunn MJ, editor. Proteomics. 2016 Mar;16(5):741–58.

202. Chen B, Butte AJ. Leveraging big data to transform target selection and drug discovery. Clin Pharmacol Ther. 2016 Mar;99(3):285–97.

203. de Lemos JA, Rohatgi A, Ayers CR. Applying a Big Data Approach to Biomarker Discovery: Running Before We Walk? Circulation. Lippincott Williams & Wilkins; 2015 Dec 15;132(24):2289–92.

204. Rozman D, Acimovic J, Schmeck B. Training in Systems Approaches for the Next Generation of Life Scientists and Medical Doctors. Methods Mol Biol. New York, NY: Springer New York; 2016;1386(Chapter 5):73–86.

205. Deo RC. Machine Learning in Medicine. Circulation. Lippincott Williams &

Wilkins; 2015 Nov 17;132(20):1920–30.

206. Lum PY, Singh G, Lehman A, Ishkanov T, Vejdemo-Johansson M, Alagappan M, et al. Extracting insights from the shape of complex data using topology. Sci Rep. Nature Publishing Group; 2013;3:1236.

207. Nicolau M, Levine AJ. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. 2011.

208. Nielson JL, Paquette J, Liu AW, Guandique CF, Tovar CA, Inoue T, et al. Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury. Nat Commun. 2015;6:8581.

209. Gladman DD, Antoni C, Mease P, Clegg DO, Nash P. Psoriatic arthritis: epidemiology, clinical features, course, and outcome. Ann Rheum Dis. 2005 Mar;64 Suppl 2:ii14–7.

210. Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P, Mielants H, et al. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum. 2006 Aug;54(8):2665–73.

211. Clegg DO, Reda DJ, Mejias E, Cannon GW, Weisman MH, Taylor T, et al. Comparison of sulfasalazine and placebo in the treatment of psoriatic arthritis. A Department of Veterans Affairs Cooperative Study. Arthritis Rheum. 1996 Dec;39(12):2013–20.

212. Coates LC, Moverley AR, McParland L, Brown S, Navarro-Coy N, O'Dwyer JL, et al. Effect of tight control of inflammation in early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. Lancet. 2015 Dec 19;386(10012):2489–98.

213. McInnes IB, Mease PJ, Kirkham B, Kavanaugh A. Secukinumab, a human anti-interleukin-17A monoclonal antibody, in patients with psoriatic arthritis (FUTURE 2): a randomised, double-blind, placebo-controlled, …. The Lancet. 2015.

214. Kavanaugh A, Mease PJ, Gomez-Reino JJ, Adebajo AO, Wollenhaupt J, Gladman DD, et al. Treatment of psoriatic arthritis in a phase 3 randomised, placebo-controlled trial with apremilast, an oral phosphodiesterase 4 inhibitor. Ann Rheum Dis. 2014 Jun;73(6):1020–6.

215. Wong AK, You M. Entropy and distance of random graphs with application to structural pattern recognition. IEEE Trans Pattern Anal Mach Intell. 1985 May;7(5):599–609.

216. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 19;43(7):e47–7.

217. Ji H, Liu XS. Analyzing 'omics data using hierarchical models. Nat Biotechnol. 2010 Apr;28(4):337–40.

218. Cope LM, Irizarry RA, Jaffee HA, Wu Z, Speed TP. A benchmark for Affymetrix GeneChip expression measures. Bioinformatics. Oxford University Press; 2004 Feb 12;20(3):323–31.

219. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. Oxford University Press; 2003 Apr;4(2):249–64.

220. McCall MN, Bolstad BM, Irizarry RA. Frozen robust multiarray analysis (fRMA). Biostatistics. 2010 Apr;11(2):242–53.

221. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. Oxford University Press; 2007 Jan;8(1):118–27.

222. Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. 2004. Bunce C, editor. Vol. 3, Statistical Applications in Genetics and Molecular Biology.

223. Murie C, Woody O, Lee AY, Nadon R. Comparison of small n statistical tests of differential expression applied to microarrays. BMC Bioinformatics. 2009;10(1):45–18.

224. Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett. 2004 Aug 27;573(1-3):83–92.

225. Breitling R, Herzyk P. Rank-based methods as a non-parametric alternative of the T-statistic for the analysis of biological microarray data. J Bioinform Comput Biol. 2005 Oct;3(5):1171–89.

226. Hong F, Breitling R. A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. Bioinformatics. 2008 Feb 1;24(3):374–82.

227. Miller AM, Gilchrist DS, Nijjar J, Araldi E, Ramirez CM, Lavery CA, et al. MiR-155 Has a Protective Role in the Development of Non-Alcoholic Hepatosteatosis in Mice. Federici M, editor. PLoS ONE. 2013 Aug 21;8(8):e72324–10.

228. Vardhanabhuti S, Blakemore SJ, Clark SM, Ghosh S, Stephens RJ, Rajagopalan D. A comparison of statistical tests for detecting differential expression using Affymetrix oligonucleotide microarrays. OMICS. 2006;10(4):555–66.

229. Jeffery IB, Higgins DG, Culhane AC. Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. BMC Bioinformatics. 2006;7:359.

230. Lim WK, Wang K, Lefebvre C, Califano A. Comparative analysis of microarray normalization procedures: effects on reverse engineering gene networks. Bioinformatics. 2007 Jul 1;23(13):i282–8.

231. Cantini F, Niccoli L, Nannini C, Kaloudi O, Bertoni M, Cassarà E. Psoriatic arthritis: a systematic review. Int J Rheum Dis. 2010 Oct 1;13(4):300–17.

232. Kurowska-Stolarska M, Alivernini S, Ballantine LE, Asquith DL, Millar NL, Gilchrist DS, et al. MicroRNA-155 as a proinflammatory regulator in clinical

and experimental arthritis. Proc Natl Acad Sci USA. 2011 Jun 20.

233. Li J, Hsu H-C, Mountz JD. The Dynamic Duo-Inflammatory M1 macrophages and Th17 cells in Rheumatic Diseases. J Orthop Rheumatol. 2013 Nov 1;1(1):4.

234. Hashimoto-Kataoka T, Hosen N, Sonobe T, Arita Y, Yasui T, Masaki T, et al. Interleukin-6/interleukin-21 signaling axis is critical in the pathogenesis of pulmonary arterial hypertension. Proc Natl Acad Sci USA. National Acad Sciences; 2015 May 19;112(20):E2677–86.

235. Watson M, Pérez-Alegre M, Baron MD, Delmas C, Dovc P, Duval M, et al. Analysis of a simulated microarray dataset: comparison of methods for data normalisation and detection of differential expression (open access publication). Genet Sel Evol. BioMed Central Ltd; 2007 Nov;39(6):669–83.

236. Lindberg J, af Klint E, Ulfgren A-K, Stark A, Andersson T, Nilsson P, et al. Variability in synovial inflammation in rheumatoid arthritis investigated by microarray technology. Arthritis Res Ther. BioMed Central Ltd; 2006;8(2):R47.

237. Lindberg J, af Klint E, Catrina AI, Nilsson P, Klareskog L, Ulfgren A-K, et al. Effect of infliximab on mRNA expression profiles in synovial tissue of rheumatoid arthritis patients. Arthritis Res Ther. BioMed Central Ltd; 2006;8(6):R179.

238. Pohlers D, Beyer A, Koczan D, Wilhelm T, Thiesen H-J, Kinne RW. Constitutive upregulation of the transforming growth factor-beta pathway in rheumatoid arthritis synovial fibroblasts. Arthritis Res Ther. BioMed Central Ltd; 2007;9(3):R59.

239. Badot V, Galant C, Nzeusseu Toukap A, Theate I, Maudoux A-L, Van den Eynde BJ, et al. Gene expression profiling in the synovium identifies a predictive signature of absence of response to adalimumab therapy in rheumatoid arthritis. Arthritis Res Ther. BioMed Central Ltd; 2009;11(2):R57.

240. Bienkowska JR, Dalgin GS, Batliwalla F, Allaire N, Roubenoff R, Gregersen PK, et al. Convergent random forest predictor: Methodology for predicting drug response from genome-scale data applied to anti-TNF response. Genomics. 2009 Dec;94(6):423–32.

241. Koczan D, Drynda S, Hecker M, Drynda A, Guthke R, Kekow J, et al. Molecular discrimination of responders and nonresponders to anti-TNF alpha therapy in rheumatoid arthritis by etanercept. Arthritis Res Ther. BioMed Central Ltd; 2008;10(3):R50.

242. Yarilina A, Park-Min K-H, Antoniv T, Hu X, Ivashkiv LB. TNF activates an IRF1-dependent autocrine loop leading to sustained expression of chemokines and STAT1-dependent type I interferon–response genes. Nat Immunol. 2008 Mar 16;9(4):378–87.

243. You S, Yoo S-A, Choi S, Kim J-Y, Park S-J, Ji J, et al. Identification of key regulators for the migration and invasion of rheumatoid synoviocytes through a systems approach. Proceedings of the National Academy of Sciences. 2014;111(1):550–5.

244. O'Shea JJ, Plenge R. JAK and STAT signaling molecules in immunoregulation and immune-mediated disease. Immunity. 2012 Apr 20;36(4):542–50.

245. Westhovens R, Alten R. Filgotinib (GLPG0634), an oral JAK1 selective inhibitor is effective in combination with methotrexate in patients with active rheumatoid arthritis: results from …. ARTHRITIS & …; 2015.

246. Burger D, Dayer J-M. Assays of T-Cell Contact Dependent Monocyte-Macrophage Functions. 2007. 10 p.

247. Burger D. Cell contact-mediated signaling of monocytes by stimulated T cells: a major pathway for cytokine induction. Eur Cytokine Netw. 2000 Sep;11(3):346–53.

248. Brennan F, Foey A. Cytokine regulation in RA synovial tissue: role of T cell/macrophage contact-dependent interactions. Arthritis Res. BioMed Central; 2002;4 Suppl 3(Suppl 3):S177–82.

249. Beech JT, Andreakos E, Ciesielski CJ, Green P, Foxwell BMJ, Brennan FM. T-cell contact-dependent regulation of CC and CXC chemokine production in monocytes through differential involvement of NFkappaB: implications for rheumatoid arthritis. Arthritis Res Ther. BioMed Central Ltd; 2006;8(6):R168.

250. Mathew EC, Shaw JM, Bonilla FA, Law SK, Wright DA. A novel point mutation in CD18 causing the expression of dysfunctional CD11/CD18 leucocyte integrins in a patient with leucocyte adhesion deficiency (LAD). Clinical & Experimental Immunology. Wiley-Blackwell; 2000 Jul;121(1):133–8.

251. Shelef MA, Bennin DA, Mosher DF, Huttenlocher A. Citrullination of fibronectin modulates synovial fibroblast behavior. Arthritis Res Ther. BioMed Central Ltd; 2012;14(6):R240.

252. Bryant J, Ahern DJ, Brennan FM. CXCR4 and vascular cell adhesion molecule 1 are key chemokine/adhesion receptors in the migration of cytokine-activated T cells. Arthritis Rheum. 2012 Jun 26;64(7):2137–46.

253. Verstovsek S, Mesa RA, Gotlib J, Levy RS, Gupta V, DiPersio JF, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. N Engl J Med. 2012 Mar 1;366(9):799–807.

254. McInnes I, Leung B, Sturrock R, Field M, Liew F. Interleukin-15 mediates T cell-dependent regulation of tumor necrosis factor-α production in rheumatoid arthritis. Nat Med. 1997;3(2):189–95.

255. Sebbag M, Parry SL, Brennan FM, Feldmann M. Cytokine stimulation of T lymphocytes regulates their capacity to induce monocyte production of tumor necrosis factor-alpha, but not interleukin-10: possible relevance to pathophysiology of rheumatoid arthritis. Eur J Immunol. WILEY-VCH Verlag GmbH; 1997 Mar;27(3):624–32.

256. Cohen S, Zwillich SH, Chow V, Labadie RR, Wilkinson B. Co-administration of the JAK inhibitor CP-690,550 and methotrexate is well tolerated in patients with rheumatoid arthritis without need for dose adjustment. Br J Clin

Pharmacol. Blackwell Publishing Ltd; 2010 Feb;69(2):143–51.

257. Lettesjö H, Nordström E, Ström H, Nilsson B, Glinghammar B, Dahlstedt L, et al. Synovial fluid cytokines in patients with rheumatoid arthritis or other arthritic lesions. Scandinavian journal of immunology. 1998 Sep;48(3):286–92.

258. Mesa RA, Cortes J. Optimizing management of ruxolitinib in patients with myelofibrosis: the need for individualized dosing. J Hematol Oncol. BioMed Central Ltd; 2013;6(1):79.

259. Sanchez-Pernaute O, Filkova M, Gabucio A, Klein M, Maciejewska-Rodrigues H, Ospelt C, et al. Citrullination enhances the pro-inflammatory response to fibrin in rheumatoid arthritis synovial fibroblasts. Ann Rheum Dis. 2013 Jul 4;72(8):1400–6.

260. Pattison M, MacKenzie K, Arthur J. Inhibition of JAKs in Macrophages Increases Lipopolysaccharide-Induced Cytokine Production by Blocking IL-10–Mediated Feedback. The Journal of Immunology. 2012;189(6):2784–92.

261. Brennan FM, McInnes IB. Evidence that cytokines play a role in rheumatoid arthritis. J Clin Invest. American Society for Clinical Investigation; 2008 Nov;118(11):3537–45.

262. De Benedetti F, Brunner HI, Ruperto N, Kenwright A, Wright S, Calvo I, et al. Randomized trial of tocilizumab in systemic juvenile idiopathic arthritis. N Engl J Med. 2012 Dec 20;367(25):2385–95.

263. Brunner HI, Ruperto N, Zuber Z, Keane C, Harari O, Kenwright A, et al. Efficacy and safety of tocilizumab in patients with polyarticular-course juvenile idiopathic arthritis: results from a phase 3, randomised, double-blind withdrawal trial. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2015 Jun;74(6):1110–7.

264. Unizony S, Arias-Urdaneta L, Miloslavsky E, Arvikar S, Khosroshahi A, Keroack B, et al. Tocilizumab for the treatment of large-vessel vasculitis (giant cell arteritis, Takayasu arteritis) and polymyalgia rheumatica. Arthritis Care Res (Hoboken). John Wiley & Sons, Inc; 2012 Nov;64(11):1720–9.

265. Karonitsch T, Dalwigk von K, Steiner CW, Blüml S, Steiner G, Kiener HP, et al. Interferon signals and monocytic sensitization of the interferon-γ signaling pathway in the peripheral blood of patients with rheumatoid arthritis. Arthritis Rheum. 2012 Feb;64(2):400–8.

266. Kubo S, Yamaoka K, Kondo M, Yamagata K, Zhao J, Iwata S, et al. The JAK inhibitor, tofacitinib, reduces the T cell stimulatory capacity of human monocyte-derived dendritic cells. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2014 Dec;73(12):2192–8.

267. Yarilina A, Xu K, Chan C, Ivashkiv LB. Regulation of inflammatory responses in tumor necrosis factor-activated and rheumatoid arthritis synovial macrophages by JAK inhibitors. Arthritis Rheum. 2012 Dec;64(12):3856–66.

268. Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, et al. Comparison of treatment effects between animal experiments and clinical

trials: systematic review. BMJ. British Medical Journal Publishing Group; 2007 Jan 27;334(7586):197–7.

269. Burger D, Dayer J-M. The role of human T-lymphocyte-monocyte contact in inflammation and tissue destruction. Arthritis Res. 2002;4(Suppl 3):S169–76.

270. Ferrari-Lacraz S, Sebbag M, Chicheportiche R, Foulquier C, Serre G, Dayer J-M. Contact with stimulated T cells up-regulates expression of peptidylarginine deiminase 2 and 4 by human monocytes. Eur Cytokine Netw. 2012 Jun;23(2):36–44.

271. Unutmaz D, Pileri P, Abrignani S. Antigen-independent activation of naive and memory resting T cells by a cytokine combination. J Exp Med. 1994;180(3):1159–64.

272. Beech J, Andreakos E, Ciesielski C, Green P, Foxwell B, Brennan F. T-cell contact-dependent regulation of CC and CXC chemokine production in monocytes through differential involvement of NFκB: implications for rheumatoid arthritis. Arthritis Res Ther. 2006;8(6):R168.

273. McInnes IB, al-Mughales J, Field M, Leung BP, Huang FP, Dixon R, et al. The role of interleukin-15 in T-cell migration and activation in rheumatoid arthritis. Nat Med. 1996 Feb;2(2):175–82.

274. Wenink MH, Santegoets KCM, Platt AM, van den Berg WB, van Riel PLCM, Garside P, et al. Abatacept modulates proinflammatory macrophage responses upon cytokine-activated T cell and Toll-like receptor ligand stimulation. Ann Rheum Dis. 2012 Jan;71(1):80–3.

275. Yarilina A, Xu K, Chen J, Ivashkiv LB. TNF activates calcium-nuclear factor of activated T cells (NFAT)c1 signaling pathways in human macrophages. Proc Natl Acad Sci USA. 2011 Jan 25;108(4):1573–8.

276. Maeshima K, Yamaoka K, Kubo S, Nakano K, Iwata S, Saito K, et al. The JAK inhibitor tofacitinib regulates synovitis through inhibition of interferon-γ and interleukin-17 production by human CD4+ T cells. Arthritis Rheum. 2012 May 25;64(6):1790–8.

277. Boyle DL, Soma K, Hodge J, Kavanaugh A, Mandel D, Mease P, et al. The JAK inhibitor tofacitinib suppresses synovial JAK1-STAT signalling in rheumatoid arthritis. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2015 Jun;74(6):1311–6.

278. Yellin M, Paliienko I, Balanescu A, Ter-Vartanian S, Tseluyko V, Xu L-A, et al. A phase II, randomized, double-blind, placebo-controlled study evaluating the efficacy and safety of MDX-1100, a fully human anti-CXCL10 monoclonal antibody, in combination with methotrexate in patients with rheumatoid arthritis. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2012 Jun;64(6):1730–9.

279. Gómez-Puerta JA, Celis R, Hernández MV, Ruiz-Esquide V, Ramírez J, Haro I, et al. Differences in synovial fluid cytokine levels but not in synovial tissue cell infiltrate between anti-citrullinated peptide/protein antibody-positive and -negative rheumatoid arthritis patients. Arthritis Res Ther. BioMed Central Ltd; 2013;15(6):R182.

280. Niemand C, Nimmesgern A, Haan S, Fischer P, Schaper F, Rossaint R, et al. Activation of STAT3 by IL-6 and IL-10 in primary human macrophages is differentially modulated by suppressor of cytokine signaling 3. J Immunol. 2003 Mar 15;170(6):3263–72.

281. Wang Y, van Boxel-Dezaire AHH, Cheon H, Yang J, Stark GR. STAT3 activation in response to IL-6 is prolonged by the binding of IL-6 receptor to EGF receptor. Proc Natl Acad Sci USA. National Acad Sciences; 2013 Oct 15;110(42):16975–80.

282. Pratt AG, Swan DC, Richardson S, Wilson G, Hilkens CMU, Young DA, et al. A CD4 T cell gene signature for early rheumatoid arthritis implicates interleukin 6-mediated STAT3 signalling, particularly in anti-citrullinated peptide antibody-negative disease. Ann Rheum Dis. 2012 Aug;71(8):1374–81.

283. Anderson AE, Pratt AG, Sedhom MAK, Doran JP, Routledge C, Hargreaves B, et al. IL-6-driven STAT signalling in circulating CD4+ lymphocytes is a marker for early anticitrullinated peptide antibody-negative rheumatoid arthritis. Ann Rheum Dis. 2015 Feb 3;:1–10.

284. van der Helm-van Mil AHM, le Cessie S, van Dongen H, Breedveld FC, Toes REM, Huizinga TWJ. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. Arthritis Rheum. Wiley Subscription Services, Inc., A Wiley Company; 2007 Feb;56(2):433–40.

285. Kuriya B, Cheng CK, Chen HM, Bykerk VP. Validation of a prediction rule for development of rheumatoid arthritis in patients with early undifferentiated arthritis. Ann Rheum Dis. BMJ Publishing Group Ltd and European League Against Rheumatism; 2009 Sep;68(9):1482–5.

286. Hrdlickova B, Kumar V, Kanduri K, Zhernakova DV, Tripathi S, Karjalainen J, et al. Expression profiles of long non-coding RNAs located in autoimmune disease-associated regions reveal immune cell-type specificity. Genome Med. BioMed Central Ltd; 2014;6(10):88.

287. Rigoutsos I, Huynh T, Miranda K, Tsirigos A, McHardy A, Platt D. Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes. Proc Natl Acad Sci USA. 2006 Apr 25;103(17):6605–10.

288. Consortium TE. Standards, Guidelines and Best Practices for RNA-Seq. 2011 Jun pp. 1–7.

289. Siegel AM, Heimall J, Freeman AF, Hsu AP, Brittain E, Brenchley JM, et al. A Critical Role for STAT3 Transcription Factor Signaling in the Development and Maintenance of Human T Cell Memory. Immunity. Elsevier Inc; 2011 Nov 23;35(5):806–18.

290. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. Oxford University Press; 2009 May 1;25(9):1105–11.

291. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions

and gene fusions. Genome Biol. BioMed Central Ltd; 2013;14(4):R36.

292. Frazee AC, Pertea G, Jaffe AE, Ben Langmead, Salzberg SL, Leek JT. Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat Biotechnol. Nature Publishing Group; 2015 Mar 1;33(3):243–6.

293. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012 Mar 1;7(3):562–78.

294. Pinto A, Aldinucci D, Gloghini A, Zagonel V, Degan M, Improta S, et al. Human eosinophils express functional CD30 ligand and stimulate proliferation of a Hodgkin's disease cell line. Blood. 1996 Nov 1;88(9):3299–305.

295. Song HY, Rothe M, Goeddel DV. The tumor necrosis factor-inducible zinc finger protein A20 interacts with TRAF1/TRAF2 and inhibits NF-kappaB activation. Proc Natl Acad Sci USA. National Academy of Sciences; 1996 Jun 25;93(13):6721–5.

296. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015 Jan 15;43(D1):D447–52.

297. Brohée S, van Helden J. Evaluation of clustering algorithms for protein-protein interaction networks. BMC Bioinformatics. BioMed Central; 2006;7(1):488.

298. Brohée S, Faust K, Lima-Mendez G, Vanderstocken G, van Helden J. Network Analysis Tools: from biological networks to clusters and pathways. Nat Protoc. Nature Publishing Group; 2008;3(10):1616–29.

299. Li P, Spolski R, Liao W, Leonard WJ. Complex interactions of transcription factors in mediating cytokine biology in T cells. Immunol Rev. 2014 Sep;261(1):141–56.

300. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. Oxford University Press; 2013 Jan 1;29(1):15–21.

301. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W, et al. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. Nat Protoc. 2013 Aug 22;8(9):1765–86.

302. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015 Mar 9;12(4):357–60.

303. Bray N, Pimentel H, Melsted P, Pachter L. Near-optimal RNA-Seq quantification. arXiv.org. 2015.

304. Peeters JGC, Vervoort SJ, Tan SC, Mijnheer G, de Roock S, Vastert SJ, et al. Inhibition of Super-Enhancer Activity in Autoinflammatory Site-Derived T Cells Reduces Disease-Associated Gene Expression. Cell Rep. 2015 Sep 29;12(12):1986–96.

305. Vahedi G, Kanno Y, Furumoto Y, Jiang K, Parker SCJ, Erdos MR, et al.

Super-enhancers delineate disease-associated regulatory nodes in T cells. Nature. 2015 Feb 16;520(7548):558–62.

306. Jiang K, Zhu L, Buck MJ, Chen Y, Carrier B, Liu T, et al. Disease-Associated Single-Nucleotide Polymorphisms From Noncoding Regions in Juvenile Idiopathic Arthritis Are Located Within or Adjacent to Functional Genomic Elements of Human Neutrophils and CD4+ T Cells. Arthritis & Rheumatology. 2015 Jun 26;67(7):1966–77.

307. Rochman Y, Yukawa M, Kartashov AV, Barski A. Functional Characterization of Human T Cell Hyporesponsiveness Induced by CTLA4-Ig. Turner SJ, editor. PLoS ONE. 2015 Apr 10;10(4):e0122198–18.

308. Donlin LT, Jayatilleke A, Giannopoulou EG, Kalliolias GD, Ivashkiv LB. Modulation of TNF-Induced Macrophage Polarization by Synovial Fibroblasts. The Journal of Immunology. 2014 Aug 15;193(5):2373–83.

309. Kasmi El KC, Smith AM, Williams L, Neale G, Panopoulos AD, Panopolous A, et al. Cutting edge: A transcriptional repressor and corepressor induced by the STAT3-regulated anti-inflammatory signaling pathway. J Immunol. 2007 Dec 1;179(11):7215–9.

310. Maruyama K, Uematsu S, Kondo T, Takeuchi O, Martino MM, Kawasaki T, et al. Strawberry notch homologue 2 regulates osteoclast fusion by enhancing the expression of DC-STAMP. J Exp Med. 2013 Sep 23;210(10):1947–60.

311. Grill M, Syme TE, Noçon AL, Lu AZX, Hancock D, Rose-John S, et al. Strawberry notch homolog 2 is a novel inflammatory response factor predominantly but not exclusively expressed by astrocytes in the central nervous system. Glia. 2015 Apr 22;63(10):1738–52.

312. Duroux-Richard I, Presumey J, Courties G, Gay S, Gordeladze J, Jorgensen C, et al. MicroRNAs as new player in rheumatoid arthritis. Joint Bone Spine. 2011;78(1):17–22.

313. Stanczyk J, Ospelt C, Karouzakis E, Filer A, Raza K, Kolling C, et al. Altered expression of microRNA-203 in rheumatoid arthritis synovial fibroblasts and its role in fibroblast activation. Arthritis Rheum. 2011 Feb;63(2):373–81.

314. Baxter D, McInnes I, Kurowska-Stolarska M. Novel regulatory mechanisms in inflammatory arthritis: a role for microRNA. Immunol Cell Biol. 2012;90(3):288–92.

315. Shalek AK, Satija R, Shuga J, Trombetta JJ, Gennert D, Lu D, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature. Nature Publishing Group; 2015 Apr 9;510(7505):263–9.

316. Hanchate NK, Kondoh K, Lu Z, Kuang D, Ye X, Qiu X, et al. Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. Science. American Association for the Advancement of Science; 2015 Nov 5;350(6265):1251–5.

317. Peeters JGC, Vervoort SJ, Mijnheer G, de Roock S, Vastert SJ, Nieuwenhuis EES, et al. Autoimmune disease-associated gene expression is reduced by BET-inhibition. Genom Data. 2016 Mar;7:14–7.

Appendix

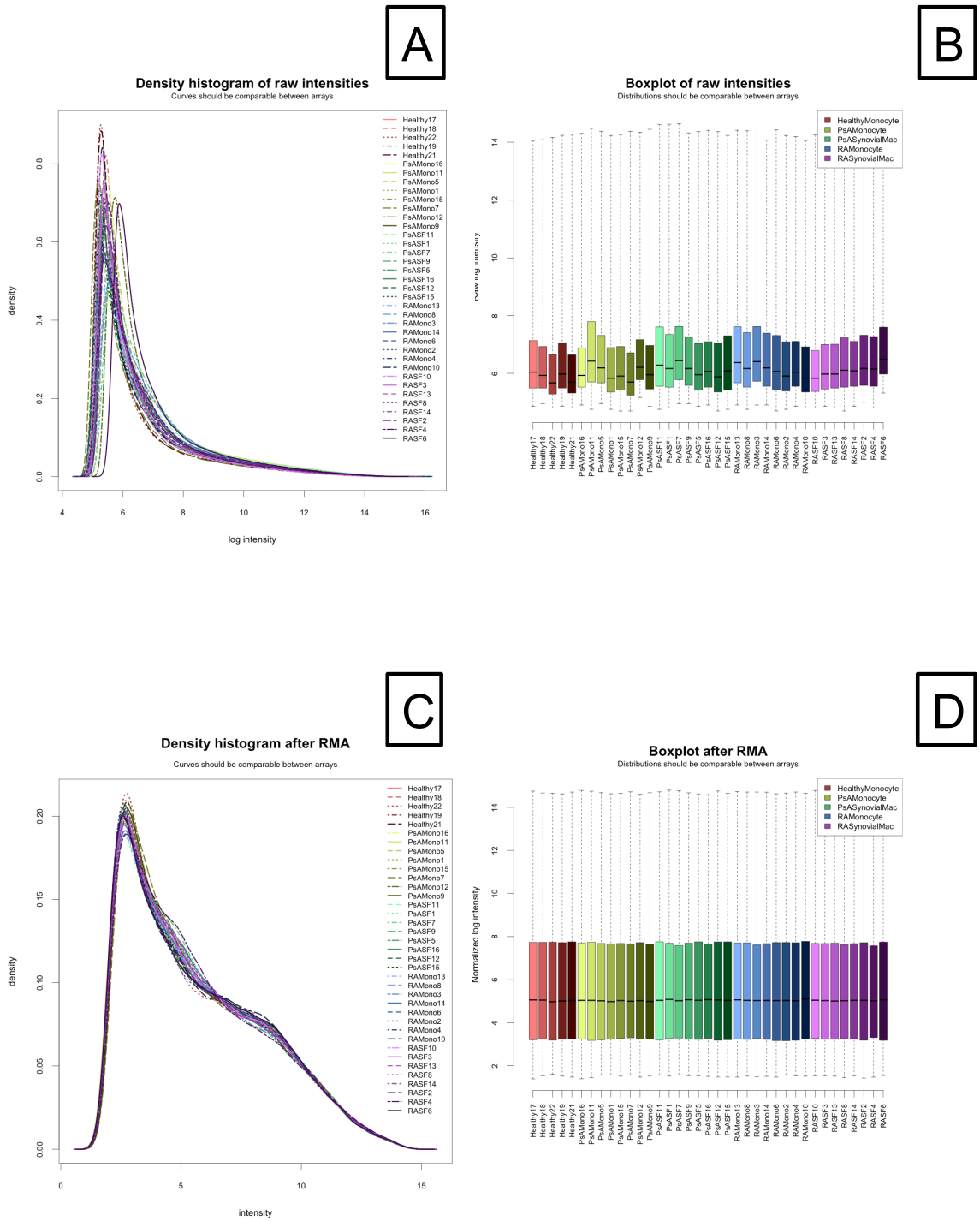


Figure 0-1 Signal intensity histograms and boxplots of primary cells only before and after RMA normalisation demonstrate good quality data. Plots were generated by reading primary cell array files into R and processed using affyAnalysisQC. A + B – Raw intensity density histogram shows similar peak intensity and spread of data. Raw intensity boxplots show a similar pattern. C and D - RMA normalised data shows a distribution similar to normal although there is still a positive skew with a small number of genes expressed at high values although this is consistent amongst samples showing that further analysis of these samples is appropriate

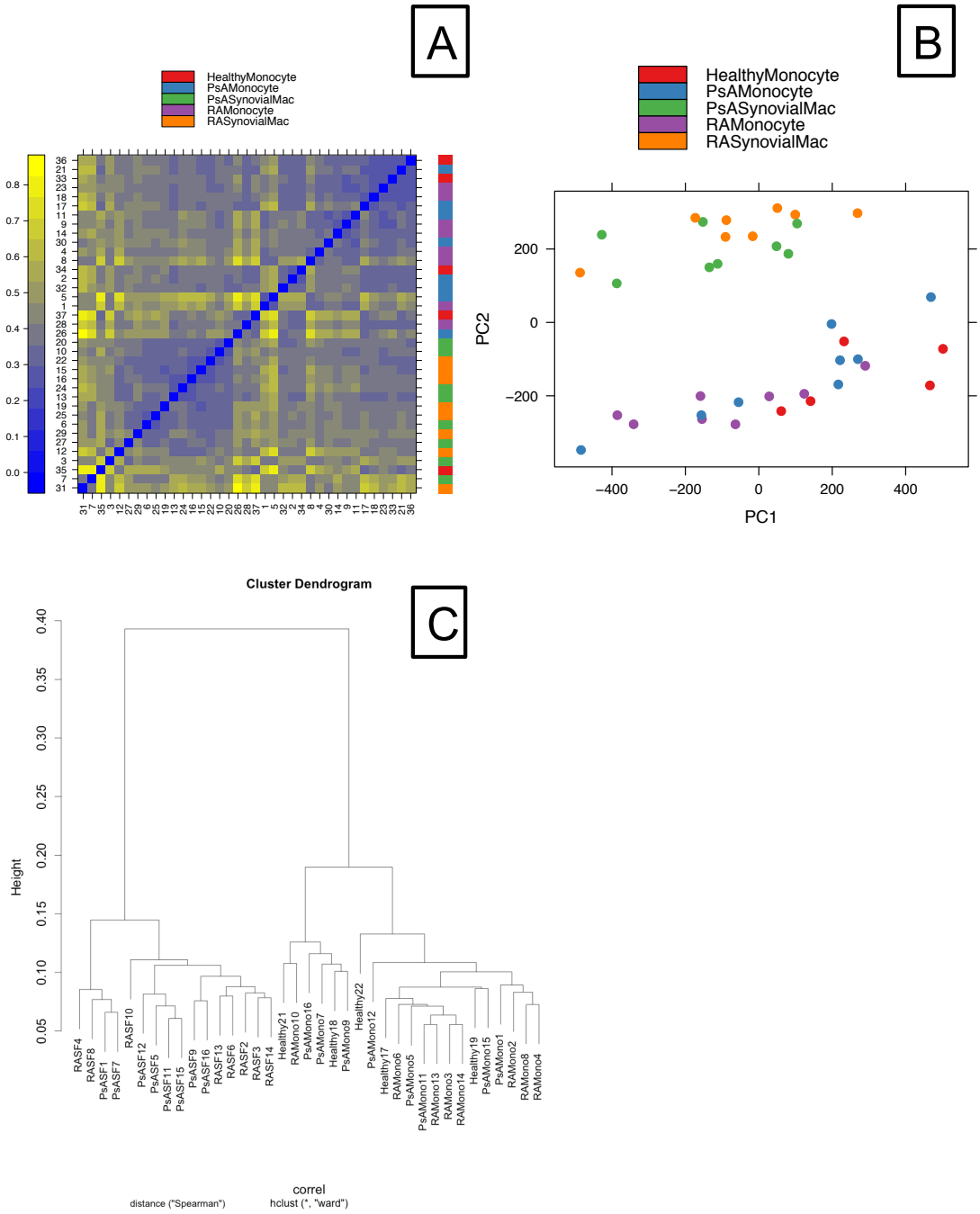


Figure 0-2 Unsupervised correlation methods of raw data demonstrate underlying structure with peripheral blood monocytes and synovial macrophages in distinct groups despite disease status. Plots were generated by reading primary cell array files into R and processed using arrayQualityMetrics. A-C: The correlation, PCA and hierarchical clustering demonstrate that subgroups are present with peripheral blood monocytes and synovial macrophages in separate groups. Disease status or absence of disease does not influence grouping at this level

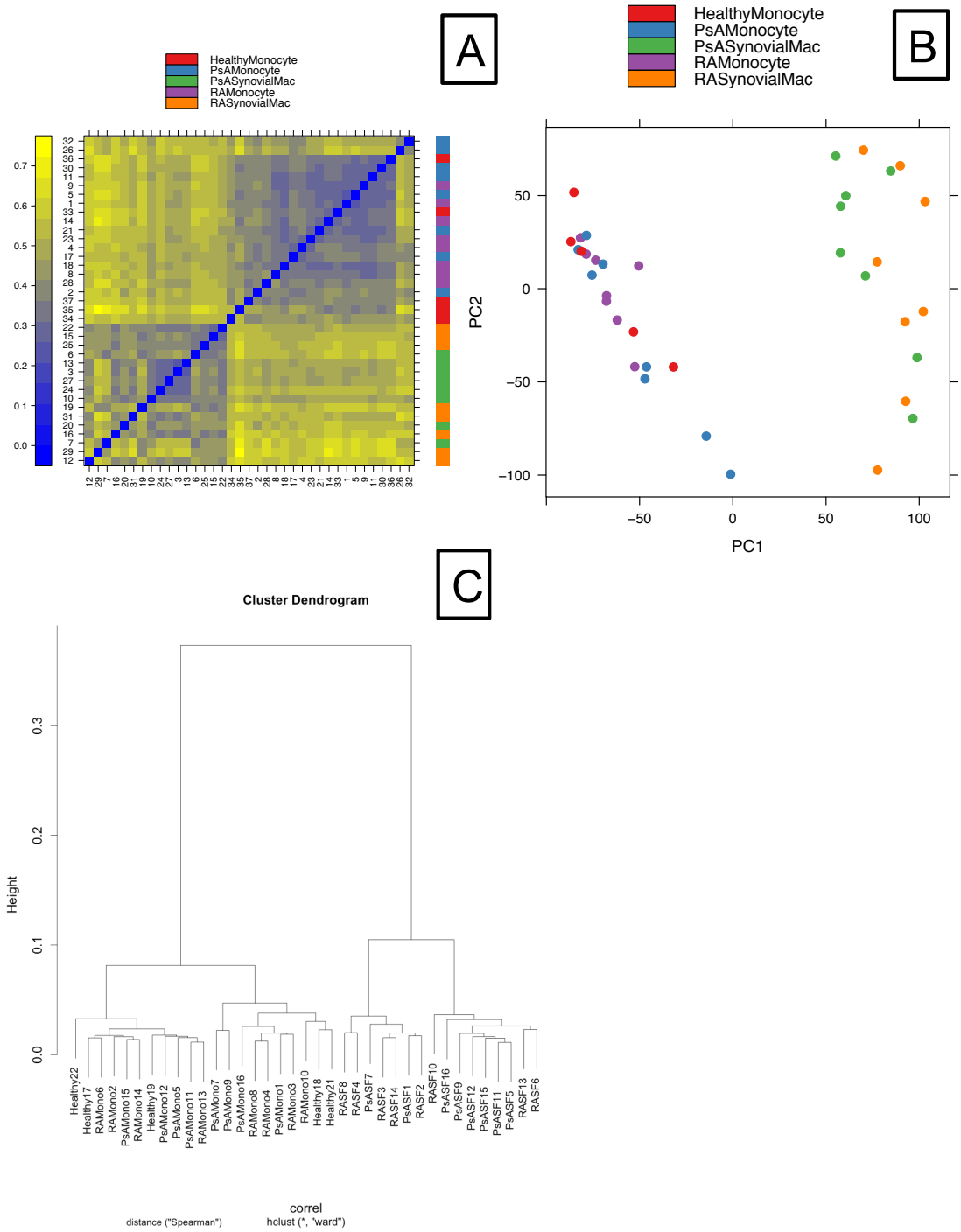


Figure 0-3 Unsupervised correlation methods of RMA normalisaed data demonstrate underlying structure more clearly with peripheral blood monocytes and synovial macrophages in distinct groups despite disease status. Plots were generated by reading primary cell array files into R and processed using arrayQualityMetrics. A-C: The correlation, PCA and hierarchical clustering demonstrate that subgroups are present with peripheral blood monocytes and synovial macrophages in separate groups. Disease status or absence of disease does not influence grouping at this level although separation of monocyte and macrophage groups is more apparent after normalisation.

Name	FileName	Target	Patient	Disease	Gender	MTX	Cell
RAMono13	GP2341_R03.CEL	RAMonocyte	13	RA	M	Yes	Monocyte
PsAMono16	GP2341_R05.CEL	PsAMonocyte	16	PsA	F	NA	Monocyte
PsASF11	GP2341_R06.CEL	PsASynovialMac	11	PsA	M	NA	Synovial Mac
RAMono8	GP2341_R07.CEL	RAMonocyte	8	RA	M	No	Monocyte
PsAMono11	GP2341_R10.CEL	PsAMonocyte	11	PsA	M	NA	Monocyte
PsASF1	GP2341_R12.CEL	PsASynovialMac	1	PsA	F	Yes	Synovial Mac
PsASF7	GP2341_R14.CEL	PsASynovialMac	7	PsA	F	Yes	Synovial Mac
RAMono3	GP2341_R15.CEL	RAMonocyte	3	RA	F	NA	Monocyte
RAMono14	GP2341_R18.CEL	RAMonocyte	14	RA	F	Yes	Monocyte
PsASF9	GP2341_R19.CEL	PsASynovialMac	9	PsA	M	No	Synovial Mac
PsAMono5	GP2341_R20.CEL	PsAMonocyte	5	PsA	M	No	Monocyte
RASF10	GP2341_R21.CEL	RASynovialMac	10	RA	F	Yes	Synovial Mac
PsASF5	GP2341_R24.CEL	PsASynovialMac	5	PsA	M	No	Synovial Mac
RAMono6	GP2341_R26.CEL	RAMonocyte	6	RA	F	Yes	Monocyte
RASF3	GP2341_R30.CEL	RASynovialMac	3	RA	F	NA	Synovial Mac
RASF13	GP2341_R31.CEL	RASynovialMac	13	RA	M	Yes	Synovial Mac
PsAMono1	GP2341_R33.CEL	PsAMonocyte	1	PsA	F	Yes	Monocyte
RAMono2	GP2341_R35.CEL	RAMonocyte	2	RA	F	NA	Monocyte
RASF8	GP2341_R36.CEL	RASynovialMac	8	RA	M	No	Synovial Mac
PsASF16	GP2341_R38.CEL	PsASynovialMac	16	PsA	F	NA	Synovial Mac
PsAMono15	GP2341_R40.CEL	PsAMonocyte	15	PsA	F	No	Monocyte
RASF14	GP2341_R41.CEL	RASynovialMac	14	RA	F	Yes	Synovial Mac
RAMono4	GP2341_R44.CEL	RAMonocyte	4	RA	M	No	Monocyte
PsASF12	GP2341_R45.CEL	PsASynovialMac	12	PsA	M	Yes	Synovial Mac
RASF2	GP2341_R47.CEL	RASynovialMac	2	RA	F	NA	Synovial Mac
PsAMono7	GP2341_R49.CEL	PsAMonocyte	7	PsA	F	Yes	Monocyte
PsASF15	GP2341_R51.CEL	PsASynovialMac	15	PsA	F	No	Synovial Mac
RAMono10	GP2341_R53.CEL	RAMonocyte	10	RA	F	Yes	Monocyte
RASF4	GP2341_R55.CEL	RASynovialMac	4	RA	M	No	Synovial Mac
PsAMono12	GP2341_R57.CEL	PsAMonocyte	12	PsA	M	Yes	Monocyte
RASF6	GP2341_R59.CEL	RASynovialMac	6	RA	F	Yes	Synovial

							Mac
PsAMono 9	GP2341_R61.CEL	PsAMonocyte	9	PsA	M	No	Monocyte
Healthy17	Healthy_blood_1_R1 3.CEL	HealthyMonocyte	17	None	F	No	Monocyte
Healthy18	Healthy_blood_2_R2 2.CEL	HealthyMonocyte	18	None	M	No	Monocyte
Healthy22	Healthy_blood_3_R2 9.CEL	HealthyMonocyte	22	None	M	No	Monocyte
Healthy19	Healthy_blood_4_R4 3.CEL	HealthyMonocyte	19	None	M	No	Monocyte
Healthy21	Healthy_blood_5_R5 0.CEL	HealthyMonocyte	21	None	M	No	Monocyte

Table 0-1 Microarray file metadata for arrays used in Chapter 3. Patient disease status and also cell type was used to assess quality of arrays and also determine batch effect. Some data was available on methotrexate use but this was incomplete. This table was used to create the phenoData file which was read into the R ExpressionSet.

```
library(affy)

### Read in CEL files and add phenodata #######

celfiles <- list.celfiles(full.names=TRUE)

raw <- ReadAffy(filenames=celfiles)
pd <- read.csv("phenodata.csv", row.names="FileName", header=TRUE)

pData(raw) <- pd


### generate QC reports on raw data

#library(affyQCReport)
#library(arrayQualityMetrics)
#QCReport(raw,file="RawaffyQCReport.pdf")


### Remove array 3

#raw <- raw[ ,-3]
#pData(raw) <- droplevels(pData(raw))

### Rerun QC report

#QCReport(raw,file="RawaffyQCReportminus3.pdf")

### Normalise with RMA ##############

rma <- rma(raw)

### QC reports on raw and rma data

#arrayQualityMetrics(expressionset=raw, do.logtransform=TRUE, force=FALSE, outdir="QARaw",
intgroup=c("Target"))

#arrayQualityMetrics(expressionset=rma, do.logtransform=FALSE, force=FALSE, outdir="QARMA",
intgroup = c("Target"))

### Normalise with gcrma and rma

library(gcrma)
library(frma)

gcrma <- gcrma(raw)
frma <- frma(raw)
mas <- mas5(raw)

### Create Correlation plots

e <- exprs(rma)
f <- exprs(frma)
g <-exprs(gcrma)
m <- exprs(mas)

library(maCorrPlot)

corrrma <- CorrSample(e, np=1000, seed=213)
corrfrma <- CorrSample(f, np=1000, seed=213)
corrgcrma <- CorrSample(g, np=1000, seed=213)
corrmas5 <- CorrSample(m, np=1000, seed=213)
```

```
pdf(file="correlationplot.pdf")
plot(corrrma, corrfrma, corrgcrma, corrmas5, cond=c("RMA", "FRMA", "GCRMA", "MAS5"), scatter=T,
curve=T)
dev.off()

#### perform a correlation analysis using CCA package on expression matrices from different
methods of normalisation

library(CCA)


### Perform limma with all methods of normalisation

library(limma)

rna <- factor(pData(rma)[,"Target"])
design <- model.matrix(~0+rna)
colnames(design) <- levels(rna)
aw <- arrayWeights(exprs(rma), design)
fit <- lmFit(exprs(rma), design, weights=aw)

### Contrasts of Synovial Mac and venn diagram

#contrastsa <- makeContrasts(RASynovialMac-PsASynovialMac, PsASynovialMac-
PsAMonocyte,RASynovialMac-RAMonocyte, levels=design)
#contr.fita <- eBayes(contrasts.fit(fit, contrastsa))
#resultsa <- decideTests(contr.fita)
#pdf(file="Venn Diagram of limma with rma normalisation -a.pdf", height=10, width=15)
#vennDiagram(resultsa)
#dev.off()

### Contrasts of Monocytes and venn diagram

contrastsb <- makeContrasts(RAMonocyte-PsAMonocyte, RAMonocyte-HealthyMonocyte,
PsAMonocyte-HealthyMonocyte, levels=design)
contr.fitb <- eBayes(contrasts.fit(fit, contrastsb))
resultsb <- decideTests(contr.fitb)
pdf(file="Venn Diagram of limma with rma normalisation no fdr -b.pdf", height=10, width=15)
vennDiagram(resultsb)
dev.off()


### Write out tables

### RMA

write.table(topTable(contr.fita, coef=1,number=nrow(fit)), file="ProbesinRAvsPsAMac_RMA.txt",
sep="\t", row.names=TRUE, quote=FALSE, col.names=NA )

write.table(topTable(contr.fita, coef=2,number=nrow(fit)),
file="ProbesinPsAMacvsMonocyte_RMA.txt", sep="\t", row.names=TRUE, quote=FALSE, col.names=NA
)

write.table(topTable(contr.fita, coef=3,number=nrow(fit)),
file="ProbesinRAMacvsMonocyte_RMA.txt", sep="\t", row.names=TRUE, quote=FALSE, col.names=NA
)

write.table(topTable(contr.fitb, coef=1,number=nrow(fit)), file="ProbesinRAvsPsAMonocyte_RMA.txt",
sep="\t", row.names=TRUE, quote=FALSE, col.names=NA )

write.table(topTable(contr.fitb, coef=2,number=nrow(fit)), file="ProbesinRAvs
HealthyMonocyte_RMA.txt", sep="\t", row.names=TRUE, quote=FALSE, col.names=NA )
```

```
write.table(topTable(contr.fitb, coef=3,number=nrow(fit)),
file="ProbesinPsAvsHealthyMonocyte_RMA.txt", sep="\t", row.names=TRUE, quote=FALSE ,
col.names=NA)

### gcrma normalisation

rna1 <- factor(pData(gcrma)[,"Target"])
design1 <- model.matrix(~0+rna1)
colnames(design1) <- levels(rna1)
aw1 <- arrayWeights(exprs(gcrma), design1)
fit1 <- lmFit(exprs(gcrma), design1, weights=aw1)

### Contrasts of Synovial Mac and venn diagram

#contrasts1a <- makeContrasts(RASynovialMac-PsASynovialMac, PsASynovialMac-
PsAMonocyte,RASynovialMac-RAMonocyte, levels=design1)
#contr.fit1a <- eBayes(contrasts.fit(fit1, contrasts1a))
#results1a <- decideTests(contr.fit1a)
#pdf(file="Venn Diagram of limma with gcrma normalisation -a.pdf", height=10, width=15)
#vennDiagram(results1a)
#dev.off()

### Contrasts of Monocytes and venn diagram

contrasts1b <- makeContrasts(RAMonocyte-PsAMonocyte, RAMonocyte-HealthyMonocyte,
PsAMonocyte-HealthyMonocyte, levels=design1)
contr.fit1b <- eBayes(contrasts.fit(fit1, contrasts1b))
results1b <- decideTests(contr.fit1b, adjust.method="none")
pdf(file="Venn Diagram of limma with gcrma normalisation no fdr-b.pdf", height=10, width=15)
vennDiagram(results1b)
dev.off()

### Write out tables

### GCRMA

write.table(topTable(contr.fit1a, coef=1,number=nrow(fit1)), file="ProbesinRAvsPsAMac_GCRMA.txt",
sep="\t", row.names=TRUE, quote=FALSE, col.names=NA )

write.table(topTable(contr.fit1a, coef=2,number=nrow(fit1)),
file="ProbesinPsAMacvsMonocyte_GCRMA.txt", sep="\t", row.names=TRUE, quote=FALSE ,
col.names=NA)

write.table(topTable(contr.fit1a, coef=3,number=nrow(fit1)),
file="ProbesinRAMacvsMonocyte_GCRMA.txt", sep="\t", row.names=TRUE, quote=FALSE,
col.names=NA )

write.table(topTable(contr.fit1b, coef=1,number=nrow(fit1)),
file="ProbesinRAvsPsAMonocyte_GCRMA.txt", sep="\t", row.names=TRUE, quote=FALSE ,
col.names=NA)

write.table(topTable(contr.fit1b, coef=2,number=nrow(fit1)), file="ProbesinRAvs
HealthyMonocyte_GCRMA.txt", sep="\t", row.names=TRUE, quote=FALSE, col.names=NA )

write.table(topTable(contr.fit1b, coef=3,number=nrow(fit1)),
file="ProbesinPsAvsHealthyMonocyte_GCRMA.txt", sep="\t", row.names=TRUE, quote=FALSE,
col.names=NA )

### frma normalisation

rna2 <- factor(pData(frma)[,"Target"])
design2 <- model.matrix(~0+rna2)
colnames(design2) <- levels(rna2)
```

```
aw2 <- arrayWeights(exprs(frma), design2)
fit2 <- lmFit(exprs(frma), design2, weights=aw2)

### Contrasts of Synovial Mac and venn diagram

#contrasts2a <- makeContrasts(RASynovialMac-PsASynovialMac, PsASynovialMac-
PsAMonocyte,RASynovialMac-RAMonocyte, levels=design2)
#contr.fit2a <- eBayes(contrasts.fit(fit2, contrasts2a))
#results2a <- decideTests(contr.fit2a)
#pdf(file="Venn Diagram of limma with frma normalisation-a.pdf", height=10, width=15)
#vennDiagram(results2a)
#dev.off()

### Contrasts of Monocytes and venn diagram

contrasts2b <- makeContrasts(RAMonocyte-PsAMonocyte, RAMonocyte-HealthyMonocyte,
PsAMonocyte-HealthyMonocyte,  levels=design2)
contr.fit2b <- eBayes(contrasts.fit(fit2, contrasts2b))
results2b <- decideTests(contr.fit2b, adjust.method="none")
pdf(file="Venn Diagram of limma with frma normalisation no fdr-b.pdf", height=10, width=15)
vennDiagram(results2b)
dev.off()



### Write out tables

### FRMA

write.table(topTable(contr.fit2a, coef=1,number=nrow(fit2)), file="ProbesinRAvsPsAMac_FRMA.txt",
sep="\t", row.names=TRUE, quote=FALSE , col.names=NA)

write.table(topTable(contr.fit2a, coef=2,number=nrow(fit2)),
file="ProbesinPsAMacvsMonocyte_FRMA.txt", sep="\t", row.names=TRUE, quote=FALSE,
col.names=NA )

write.table(topTable(contr.fit2a, coef=3,number=nrow(fit2)),
file="ProbesinRAMacvsMonocyte_FRMA.txt", sep="\t", row.names=TRUE, quote=FALSE, col.names=NA
)

write.table(topTable(contr.fit2b, coef=1,number=nrow(fit2)),
file="ProbesinRAvsPsAMonocyte_FRMA.txt", sep="\t", row.names=TRUE, quote=FALSE, col.names=NA
)

write.table(topTable(contr.fit2b, coef=2,number=nrow(fit2)), file="ProbesinRAvs
HealthyMonocyte_FRMA.txt", sep="\t", row.names=TRUE, quote=FALSE, col.names=NA )

write.table(topTable(contr.fit2b, coef=3,number=nrow(fit2)),
file="ProbesinPsAvsHealthyMonocyte_FRMA.txt", sep="\t", row.names=TRUE, quote=FALSE ,
col.names=NA)

### Compare Diff expr genes in RASFMac vs PsASFMAC

### define which contr.fit to use by a number

j=3

rma_deg_result <- topTable(contr.fitb, coef=j, number=nrow(fit))
rma_deg_result <- rma_deg_result[rma_deg_result$adj.P.Val<=0.05,]
gcrma_deg_result <- topTable(contr.fit1b, coef=j, number=nrow(fit1))
gcrma_deg_result <- gcrma_deg_result[gcrma_deg_result$adj.P.Val<=0.05,]
frma_deg_result <- topTable(contr.fit2b, coef=j, number=nrow(fit2))
frma_deg_result <- frma_deg_result[frma_deg_result$adj.P.Val<=0.05,]
```

```
### VD between methods of normalisation

library(VennDiagram)

venn.diagram(list("Probes significant from RMA"= as.factor(row.names(rma_deg_result)), "Probes
significant from GCRMA"=as.factor(row.names(gcrma_deg_result)), "Probes signficant from
FRMA"=as.factor(row.names(frma_deg_result))), "Normalisation method overlap PsAMono vs
HealthyMono.tiff",
scaled=FALSE,
euler.d=FALSE,
height=800,
width=800,
resolution=50,
units="px",
main="Normalisation method overlap PsAMono vs HealthyMono",
sub="",
#main.pos=c(2, 3),
main.fontfamily="sans",
main.col="black",
main.cex="3",
sub.fontfamily="sans",
sub.col="black",
sub.cex="3",
fill=c("lightblue", "cornflowerblue", "skyblue"),
fontfamily="sans",
cat.fontfamily="sans",
label.col="black",
cex=2,
cat.dist=0.035,
cat.cex=1.9,
cat.dafaultpos="text")


### Generate Heatmaps ###

require(Heatplus)
corrdist = function(x) as.dist(1-cor(t(x)))
hclust.avl = function(x) hclust(x, method="average")

RASFMACPsASFMac <- rma[ , rma$Target %in% c("RASynovialMac", "PsASynovialMac")]
pData(RASFMACPsASFMac) <- droplevels(pData(RASFMACPsASFMac))
esetrma <- exprs(RASFMACPsASFMac)[row.names(rma_deg_result),]
ann <- annHeatmap(esetrma, ann=pData(RASFMACPsASFMac)$Target,legend=3,dendrogram
=list(clustfun=hclust.avl, distfun=corrdist, Col=list(status="hide")))
plot(ann)

### try attract package for JAK

eset1 <- rma[ , rma$Target %in% c("RASynovialMac", "PsASynovialMac", "HealthyMonocyte",
"RAMonocyte", "PsAMonocyte")]
pData(eset1) <- droplevels(pData(eset1))

library(attract)

attractor.states <- findAttractors(eset1, "Target", nperm=10, annotation="hgu133plus2.db")

remove.these.genes <- removeFlatGenes(eset1, "Target", contrasts=NULL, limma.cutoff=0.05)

top10.syn <- findSynexprs(attractor.states@rankedPathways[1:10,1], attractor.states,
removeGenes=remove.these.genes)

pretty.col <- rainbow(15)
```

```
for(i in 1:10)

{plotsynexprs(get(ls(top10.syn)[4], top10.syn), vertLines=c(5,10,16,24), tickLabels=c("CellContact",
"Normal", "MCSF", "PsA", "RA"), tickMarks=c(2.5,7.5,13,20, 28), index=i,
main=paste("SynexpressionGroup ", i, sep=""), col=pretty.col[i])

}

### JAK syn exprs ###

jak.syn <- findSynexprs("04630", attractor.states, remove.these.genes)

pretty.col <- rainbow(15)

for(i in 1:20)

{plotsynexprs(jak.syn, vertLines=c(5,10,16,24), tickLabels=c("CellContact", "Normal", "MCSF", "PsA",
"RA"), tickMarks=c(2.5,7.5,13,20, 28), index=i, main=paste("SynexpressionGroup JAK-STAT", i, sep=""),
col=pretty.col[i])

}

jak.cor <- findCorrPartners(jak.syn, eset1, remove.these.genes)
jak.func <- calcFuncSynexprs(jak.syn, attractor.states, "CC",
annotation="hgu133plus2.db")
```

List of genes differentially expressed in RNA Sequencing experiment of synovial fluid treated versus control CD4 T-cells from patients with RA

Test ID	Gene	Locus	log2(control FPKM)	log2(SF FPKM)	log2(Ratio)	q Value
ENSG00000256424.1	RP11-495K9.7	chr12:132131657-132132666	-10	1.19	11.19	0.00350769
ENSG00000161544.5	CYGB	chr17:74523437-74553765	-10	0.37	10.37	0.0127552
ENSG00000173369.11	C1QB	chr1:22979254-22988031	-10	0.21	10.21	0.00350769
ENSG00000223561.2	AC003090.1	chr7:25632970-25790614	-10	0.2	10.2	0.00350769
ENSG00000109906.9	ZBTB16	chr11:113930314-114227293	0.39	4.98	4.59	0.00350769
ENSG00000258919.1	RP11-1029J19.4	chr14:102095320-102098867	-0.37	3.46	3.84	0.0127552
ENSG00000182585.5	EPGN	chr4:75174189-75181024	-1.94	1.85	3.79	0.00350769
ENSG00000258512.1	RP11-796G6.2	chr14:102100790-102198859	1.31	4.67	3.36	0.00350769
ENSG00000182489.7	XKRX	chrX:100168430-100184422	-1.13	2.21	3.34	0.00350769
ENSG00000269926.1	RP11-442H21.2	chr10:74033677-74035794	3.67	6.91	3.24	0.00350769
ENSG00000168209.4	DDIT4	chr10:74033677-74035794	2.48	5.6	3.12	0.00350769
ENSG00000096060.10	FKBP5	chr6:35541361-35696360	4.11	7.17	3.06	0.00350769
ENSG00000164674.11	SYTL3	chr6:159071045-159185908	4.16	6.96	2.8	0.00350769
ENSG00000106952.3	TNFSF8	chr9:117656002-117692697	4.2	6.96	2.76	0.00350769
ENSG00000117289.7	TXNIP	chr1:145438468-145442635	7.85	10.61	2.76	0.00350769
ENSG00000049089.9	COL9A2	chr1:40766158-40782966	-0.9	1.85	2.75	0.00350769
ENSG00000182021.5	RP11-381O7.3	chr9:67017399-67032072	-2.16	0.55	2.71	0.00350769
ENSG00000157514.12	TSC22D3	chrX:106956450-107020572	6.83	9.53	2.69	0.00350769

ENSG00000170522.5	ELOVL6	chr4:110967001-111120355	-1.9	0.75	2.65	0.0164748
ENSG00000171451.13	DSEL	chr18:65173818-65566856	-1.78	0.86	2.64	0.00350769
ENSG00000263424.1	CTD-2541J13.2	chr18:65173818-65566856	-0.7	1.94	2.64	0.00350769
ENSG00000136235.11	GPNMB	chr7:23275585-23314727	-1.67	0.93	2.6	0.00350769
ENSG00000057657.10	PRDM1	chr6:106534194-106557814	1.75	4.23	2.48	0.00350769
ENSG00000184557.3	SOCS3	chr17:76352863-76356158	2.54	4.95	2.41	0.00350769
ENSG00000235842.1	RP11-356I2.2	chr6:138144809-138147031	1.53	3.92	2.39	0.00350769
ENSG00000090376.4	IRAK3	chr12:66582658-66651214	-2.07	0.21	2.28	0.00868571
ENSG00000227145.1	IL21-AS1	chr4:123533782-123610311	-2.1	0.08	2.18	0.0180706
ENSG00000237513.1	RP11-325F22.2	chr7:104581509-104653491	1.37	3.52	2.15	0.00631688
ENSG00000099250.12	NRP1	chr10:33466419-33625190	-1.84	0.27	2.11	0.014592
ENSG00000164120.9	HPGD	chr4:175411327-175444305	0.58	2.68	2.09	0.00350769
ENSG00000271321.1	CTAGE6	chr7:143452181-143454789	-1.97	0.08	2.04	0.0283523
ENSG00000228484.1	RP11-106M7.1	chr10:116524546-116539662	-1.83	0.19	2.02	0.0378566
ENSG00000080546.9	SESN1	chr6:109307639-109416022	4.15	6.14	1.98	0.00350769
ENSG00000084070.7	SMAP2	chr1:40810521-40888998	6.03	8	1.97	0.00350769
ENSG00000137801.9	THBS1	chr15:39873279-39891667	0.59	2.53	1.94	0.00350769
ENSG00000115604.6	IL18R1	chr2:102927961-103015218	2.55	4.4	1.85	0.00350769
ENSG00000270348.1	RP11-436H22.1	chr2:70883915-70885457	-1.1	0.74	1.84	0.0496749
ENSG00000110848.4	CD69	chr12:9905081-9913497	4.13	5.9	1.77	0.00350769

ENSG00000101187.11	SLCO4A1	chr20:61272070-61317137	0.97	2.72	1.76	0.00350769
ENSG00000104381.8	GDAP1	chr8:75146934-75401107	0.38	2.13	1.75	0.02688
ENSG00000113269.8	RNF130	chr5:179338650-179499118	3.72	5.48	1.75	0.00350769
ENSG00000270127.1	RP11-526I2.5	chr15:101087971-101090358	0.61	2.31	1.7	0.00350769
ENSG00000097096.8	SYDE2	chr1:85622555-85666729	1.17	2.81	1.64	0.00350769
ENSG00000130340.10	SNX9	chr6:158244295-158366109	4.03	5.64	1.62	0.00350769
ENSG00000163629.8	PTPN13	chr4:87515467-87736324	1.18	2.8	1.62	0.0109714
ENSG00000141384.7	TAF4B	chr18:23805899-23971649	2.82	4.43	1.61	0.00350769
ENSG00000198478.6	SH3BGRL2	chr6:80340999-80413372	-0.1	1.46	1.56	0.00350769
ENSG00000113916.13	BCL6	chr3:187416046-187463515	2.37	3.91	1.54	0.00350769
ENSG00000173597.3	SULT1B1	chr4:70592565-70653679	2.55	4.1	1.54	0.00350769
ENSG00000090104.7	RGS1	chr1:192544856-192549161	4.54	6.05	1.51	0.00350769
ENSG00000172985.8	SH3RF3	chr2:109745803-110262207	1.39	2.87	1.48	0.00350769
ENSG00000105982.12	RNF32	chr7:156264889-156469824	2.35	3.82	1.47	0.0164748
ENSG00000150347.10	ARID5B	chr10:63661058-63856703	3.8	5.23	1.43	0.00350769
ENSG00000119138.3	KLF9	chr9:72999502-73029540	3.2	4.62	1.42	0.00350769
ENSG00000136153.14	LMO7	chr13:76123618-76434004	4.09	5.49	1.4	0.00350769
ENSG00000141469.12	SLC14A1	chr18:42792959-43332485	0.78	2.17	1.39	0.0109714
ENSG00000113249.8	HAVCR1	chr5:156456423-156486130	3.08	4.44	1.37	0.0180706
ENSG00000111863.7	ADTRP	chr6:11713887-	5.38	6.74	1.35	0.00350769

		11807279				
ENSG00000087589.12	CASS4	chr20:54987167-55034396	1.99	3.33	1.34	0.00350769
ENSG00000188211.4	NCR3LG1	chr11:17373272-17398888	-1.21	0.13	1.34	0.0180706
ENSG00000080986.7	NDC80	chr18:2571509-2616634	0.91	2.24	1.33	0.0127552
ENSG00000156510.11	HKDC1	chr10:70975088-71027904	0.85	2.17	1.33	0.0211478
ENSG00000259863.1	SH3RF3-AS1	chr2:109743782-109745386	0.22	1.55	1.33	0.0496749
ENSG00000114423.14	CBLB	chr3:105374304-105588396	5.17	6.5	1.32	0.00350769
ENSG00000145990.6	GFOD1	chr6:13358061-13487894	-0.77	0.55	1.32	0.02688
ENSG00000260923.1	AC137934.1	chr16:90252404-90289086	2.22	3.54	1.32	0.0390251
ENSG00000140848.12	CPNE2	chr16:57126448-57181878	0.67	1.98	1.31	0.0352464
ENSG00000174946.5	GPR171	chr3:150803483-151176497	6.24	7.53	1.3	0.00350769
ENSG00000198879.7	SFMBT2	chr10:7200585-7453450	3.66	4.95	1.29	0.00350769
ENSG00000120129.5	DUSP1	chr5:172189982-172204777	2.2	3.45	1.26	0.00350769
ENSG00000260910.1	LINC00565	chr13:114629486-114631817	1.3	2.56	1.26	0.00631688
ENSG00000074966.6	TXK	chr4:48068409-48136273	5.6	6.83	1.23	0.00350769
ENSG00000155324.5	GRAMD3	chr5:125695823-125832186	2.01	3.23	1.22	0.0309307
ENSG00000173262.7	SLC2A14	chr12:7965107-8043744	-0.14	1.08	1.22	0.0324267
ENSG00000138646.4	HERC5	chr4:89378267-89427314	2.94	4.12	1.18	0.00350769
ENSG00000118515.7	SGK1	chr6:134490383-134639250	1.02	2.19	1.17	0.00868571
ENSG00000182118.5	FAM89A	chr1:231154703-231175992	1.86	3.02	1.16	0.0211478
ENSG00000107890.12	ANKRD26	chr10:27280842-27389421	3.12	4.25	1.14	0.00868571
ENSG0000	LINS	chr15:10110	4.63	5.76	1.14	0.008685

0140471.12		7118-101191910				71
ENSG00000072310.12	SREBF1	chr17:17715068-17740325	1.65	2.78	1.13	0.00631688
ENSG00000112167.5	SAYSD1	chr6:39071839-39082965	4.3	5.43	1.13	0.00350769
ENSG00000135842.12	FAM129A	chr1:184759857-184943682	4.34	5.47	1.13	0.00350769
ENSG00000154589.2	LY96	chr8:74903586-74941322	3.82	4.95	1.13	0.0198381
ENSG00000121749.11	TBC1D15	chr12:72233486-72320629	4.59	5.72	1.12	0.00350769
ENSG00000118503.10	TNFAIP3	chr6:138178422-138204449	6.26	7.37	1.11	0.00350769
ENSG00000104312.6	RIPK2	chr8:90769974-90803291	1.54	2.63	1.09	0.024119
ENSG00000253424.1	CTC-436K13.3	chr5:157683144-157685444	3.75	4.84	1.09	0.024119
ENSG00000135048.9	TMEM2	chr9:74298281-74431606	5.31	6.38	1.08	0.00350769
ENSG00000144802.7	NFKBIZ	chr3:101498045-101579866	4.76	5.84	1.08	0.0127552
ENSG00000159399.5	HK2	chr2:75061107-75120486	0.91	1.97	1.07	0.00350769
ENSG00000105983.14	LMBR1	chr7:156473570-156685924	5.02	6.07	1.05	0.00350769
ENSG00000159674.7	SPON2	chr4:1160719-1202750	2.6	3.65	1.05	0.0451965
ENSG00000112245.6	PTP4A1	chr6:64231665-64293492	3.18	4.22	1.04	0.0451965
ENSG00000162946.16	DISC1	chr1:231762560-232177018	2.71	3.75	1.04	0.00350769
ENSG00000064932.11	SBNO2	chr19:1107635-1174282	4.22	5.25	1.03	0.00350769
ENSG00000174839.8	DENND6A	chr3:57611183-57678816	3.9	4.93	1.03	0.0390251
ENSG00000145416.9	MARCH1	chr4:164445449-165305202	1.7	2.72	1.02	0.00350769
ENSG00000145850.4	TIMD4	chr5:156346292-156390266	2.51	3.53	1.02	0.0180706
ENSG0000	UGP2	chr2:640680	5.15	6.16	1.01	0.003507

0169764.10		73-64118696				69
ENSG00000215417.5	MIR17HG	chr13:92000073-92006833	3.37	4.37	1.01	0.0127552
ENSG00000121966.6	CXCR4	chr2:136871918-136875735	8.53	9.53	1	0.00350769
ENSG00000122862.4	SRGN	chr10:70847861-70864567	7.06	8.05	1	0.00350769
ENSG00000155749.8	ALS2CR12	chr2:202152993-202222121	2.95	3.95	1	0.0464
ENSG00000152518.5	ZFP36L2	chr2:43393799-43823185	5.93	6.91	0.98	0.00350769
ENSG00000113263.8	ITK	chr5:156512842-156682201	7	7.96	0.96	0.00350769
ENSG00000136997.10	MYC	chr8:128747679-128753674	5.45	6.41	0.96	0.00350769
ENSG00000145703.11	IQGAP2	chr5:75699073-76031606	5.17	6.1	0.93	0.00350769
ENSG00000151150.15	ANK3	chr10:61788158-62493248	5.47	6.39	0.92	0.0337285
ENSG00000271204.1	RP11-138A9.1	chr7:130614967-130616965	3.32	4.24	0.92	0.00631688
ENSG00000174885.8	NLRP6	chr11:278364-285359	1.56	2.46	0.91	0.0309307
ENSG00000105639.14	JAK3	chr19:17935594-17958841	4.64	5.53	0.89	0.00350769
ENSG00000173281.4	PPP1R3B	chr8:8993764-9009084	2.31	3.2	0.89	0.00350769
ENSG00000112297.10	AIM1	chr6:106959729-107018326	4.92	5.8	0.88	0.00868571
ENSG00000134333.9	LDHA	chr11:18415934-18429972	5.57	6.45	0.88	0.00350769
ENSG00000145391.8	SETD7	chr4:140417094-140477928	2.85	3.73	0.88	0.00631688
ENSG00000163565.14	IFI16	chr1:158969757-159024945	6.04	6.9	0.86	0.00350769
ENSG00000110002.11	VWA5A	chr11:123986068-124018428	2.7	3.55	0.85	0.00868571
ENSG00000198604.6	BAZ1A	chr14:35221936-35345665	4.57	5.42	0.85	0.00631688
ENSG00000064012.1	CASP8	chr2:202098165-	7.04	7.88	0.84	0.0211478

7		202152434				
ENSG00000112576.8	CCND3	chr6:41902670-42018095	6.99	7.82	0.83	0.00350769
ENSG00000155926.9	SLA	chr8:133879204-134147147	6.15	6.98	0.83	0.00350769
ENSG00000156804.3	FBXO32	chr8:124510128-124553446	4.01	4.84	0.83	0.0309307
ENSG00000048740.12	CELF2	chr10:11047258-11378666	5.85	6.67	0.82	0.0464
ENSG00000101347.7	SAMHD1	chr20:35504523-35580246	5.93	6.76	0.82	0.00350769
ENSG00000164512.13	ANKRD55	chr5:55395506-55529186	4.3	5.11	0.82	0.0127552
ENSG00000007944.10	MYLIP	chr6:16129355-16148479	4.1	4.91	0.81	0.00350769
ENSG00000125630.11	POLR1B	chr2:113299491-113334635	3.19	4	0.81	0.00631688
ENSG00000170222.10	ADPRM	chr17:10583653-10718481	5.68	6.49	0.81	0.0324267
ENSG00000091409.10	ITGA6	chr2:173292081-173489823	4.89	5.68	0.79	0.00868571
ENSG00000131507.8	NDFIP1	chr5:141488069-141534008	5.5	6.28	0.78	0.00350769
ENSG00000163599.10	CTLA4	chr2:204732508-204738683	3.55	4.34	0.78	0.0211478
ENSG00000173762.3	CD7	chr17:80272743-80275478	5.2	5.97	0.77	0.00350769
ENSG00000145715.10	RASA1	chr5:86563704-86687748	4.12	4.88	0.76	0.00350769
ENSG00000177613.7	CSTF2T	chr10:52750944-54073888	4.45	5.22	0.76	0.00350769
ENSG00000115758.8	ODC1	chr2:10580093-10588630	3.76	4.51	0.75	0.0109714
ENSG00000162739.9	SLAMF6	chr1:160454819-160493052	5	5.74	0.74	0.00350769
ENSG00000259820.1	AC083843.1	chr8:135804262-135810515	1.64	2.37	0.74	0.0254968
ENSG00000185885.11	IFITM1	chr11:307630-315272	6.54	7.27	0.73	0.014592
ENSG0000	CD53	chr1:111415	7.93	8.64	0.71	0.012755

0143119.8		774-111442550				2
ENSG00000198898.7	CAPZA2	chr7:116451123-116559315	5.59	6.3	0.71	0.014592
ENSG00000088930.6	XRN2	chr20:21283941-21370463	5.29	5.99	0.7	0.00350769
ENSG00000128050.4	PAICS	chr4:57301906-57327534	3.42	4.12	0.7	0.024119
ENSG00000170379.15	FAM115C	chr7:143318042-143427502	4.82	5.52	0.7	0.0109714
ENSG00000067066.12	SP100	chr2:231280656-231444721	6.62	7.3	0.68	0.0127552
ENSG00000100906.6	NFKBIA	chr14:35870716-35873952	5.78	6.46	0.68	0.00350769
ENSG00000143167.7	GPA33	chr1:167021787-167059868	3.71	4.39	0.68	0.00868571
ENSG00000147894.10	C9orf72	chr9:27546543-27573864	4.51	5.19	0.68	0.0211478
ENSG00000183696.9	UPP1	chr7:48128224-48148330	5.06	5.74	0.68	0.0254968
ENSG00000172292.10	CERS6	chr2:169312371-169642939	3.36	4.03	0.67	0.0366545
ENSG00000168610.10	STAT3	chr17:40465341-40540586	5.57	6.22	0.65	0.0198381
ENSG00000233251.3	AC007743.1	chr2:56400668-56613308	4.31	4.96	0.65	0.014592
ENSG00000168906.8	MAT2A	chr2:85766287-85772403	5.15	5.76	0.62	0.0378566
ENSG00000120800.4	UTP20	chr12:101673886-101780394	2.58	3.18	0.6	0.024119
ENSG00000095002.8	MSH2	chr2:47630107-47798078	3.86	4.43	0.57	0.0496749
ENSG00000124575.5	HIST1H1D	chr6:26234439-26235216	8.47	9.03	0.56	0.0309307
ENSG00000165732.8	DDX21	chr10:70715883-70744829	4.71	5.27	0.56	0.0254968
ENSG00000124203.5	ZNF831	chr20:57766074-57834168	3.75	4.3	0.54	0.0254968
ENSG00000179456.9	ZBTB18	chr1:244214584-244220778	5.31	4.8	-0.52	0.0390251

ENSG00000186517.9	ARHGAP30	chr1:161016737-161039760	5.92	5.4	-0.52	0.0402538
ENSG00000197448.9	GSTK1	chr7:142941185-142985141	7.47	6.95	-0.52	0.0464
ENSG00000126882.8	FAM78A	chr9:134133462-134151934	5.34	4.8	-0.54	0.0390251
ENSG00000104904.7	OAZ1	chr19:2252251-2273487	8.61	8.05	-0.56	0.0431127
ENSG00000142784.11	WDTC1	chr1:27561006-27635110	4.1	3.54	-0.56	0.0451965
ENSG00000197976.6	AKAP17A	chrX:1710485-1721407	5.9	5.34	-0.56	0.0378566
ENSG00000076924.7	XAB2	chr19:7684410-7694451	5.12	4.55	-0.57	0.0309307
ENSG00000092820.13	EZR	chr6:159186772-159241625	6.1	5.51	-0.58	0.0228
ENSG00000101265.11	RASSF2	chr20:4760668-4804291	4.28	3.7	-0.58	0.0211478
ENSG00000115085.9	ZAP70	chr2:98330022-98356325	7.07	6.49	-0.58	0.0211478
ENSG00000173208.3	ABCD2	chr12:39943834-40013553	3.66	3.08	-0.58	0.0283523
ENSG00000188452.9	CERKL	chr2:182401402-182545603	4.02	3.45	-0.58	0.024119
ENSG00000149930.13	TAOK2	chr16:29984961-30003582	3.93	3.33	-0.6	0.0127552
ENSG00000103502.9	CDIPT	chr16:29869677-29879371	5.3	4.68	-0.61	0.0352464
ENSG00000110711.4	AIP	chr11:67250511-67258574	6.47	5.86	-0.61	0.0297796
ENSG00000123143.8	PKN1	chr19:14543864-14582678	5.47	4.86	-0.61	0.0198381
ENSG00000159753.9	RLTPR	chr16:67678821-67694713	4.43	3.82	-0.61	0.0297796
ENSG00000184678.8	HIST2H2BE	chr1:149856009-149858232	5.1	4.49	-0.61	0.0164748
ENSG00000063244.8	U2AF2	chr19:56165511-56186081	5.59	4.97	-0.62	0.0211478
ENSG00000100239.11	PPP6R2	chr22:50781732-50883514	4.87	4.26	-0.62	0.0228
ENSG00000117984.8	CTSD	chr11:1753639-1785222	5.58	4.96	-0.62	0.0254968

ENSG00000153179.7	RASSF3	chr12:65004292-65091347	5.91	5.29	-0.62	0.0496749
ENSG00000162413.12	KLHL21	chr1:6650783-6684093	4.91	4.29	-0.62	0.0297796
ENSG00000178980.10	SEPW1	chr19:48281828-48287941	6.33	5.71	-0.62	0.0254968
ENSG00000184260.4	HIST2H2AC	chr1:149858524-149858961	7.85	7.22	-0.62	0.024119
ENSG00000185905.3	C16orf54	chr16:29753783-29759620	6.2	5.57	-0.62	0.00868571
ENSG00000205220.7	PSMB10	chr16:67968404-67970990	5.87	5.26	-0.62	0.0283523
ENSG00000242802.2	AP5Z1	chr7:4815252-4831399	4.46	3.84	-0.62	0.0416438
ENSG00000087074.7	PPP1R15A	chr19:49375648-49379314	5.1	4.47	-0.63	0.0283523
ENSG00000105669.8	COPE	chr19:19010322-19030206	6.17	5.54	-0.63	0.0283523
ENSG00000182162.5	P2RY8	chrX:1581464-1656000	5.85	5.22	-0.63	0.0366545
ENSG00000262319.1	CTC-457L16.2	chr17:19030781-19062489	5.1	4.47	-0.63	0.0378566
ENSG00000101216.6	GMEB2	chr20:62218954-62258394	3.76	3.11	-0.64	0.0198381
ENSG00000111252.6	SH2B3	chr12:111843751-111889427	3.53	2.89	-0.64	0.0496749
ENSG00000141959.12	PFKL	chr21:45719933-45747259	5.29	4.65	-0.64	0.0211478
ENSG00000157303.6	SUSD3	chr9:95820988-95847420	5.98	5.33	-0.64	0.0402538
ENSG00000160570.9	DEDD2	chr19:42702749-42724292	4.95	4.31	-0.64	0.0451965
ENSG00000211450.5	C11orf31	chr11:57508824-57510510	5.92	5.28	-0.64	0.02688
ENSG00000154016.9	GRAP	chr17:18853657-18950950	5.72	5.07	-0.65	0.0366545
ENSG00000164896.15	FASTK	chr7:150773710-150777949	4.96	4.31	-0.65	0.0211478
ENSG00000183688.4	FAM101B	chr17:289768-295730	4.24	3.59	-0.65	0.0180706
ENSG00000157873.1	TNFRSF14	chr1:2481358-2496821	6.3	5.64	-0.66	0.0254968

3						
ENSG00000072958.4	AP1M1	chr19:16308388-16346160	5.14	4.47	-0.67	0.0211478
ENSG00000111679.11	PTPN6	chr12:7055630-7070479	5.85	5.17	-0.67	0.0337285
ENSG00000109062.5	SLC9A3R1	chr17:72666716-72765492	5.73	5.05	-0.68	0.0283523
ENSG00000119669.3	IRF2BPL	chr14:77490887-77495034	3.76	3.09	-0.68	0.0127552
ENSG00000126464.9	PRR12	chr19:50094899-50129696	2.28	1.6	-0.68	0.0337285
ENSG00000135046.9	ANXA1	chr9:75766672-75785309	7.6	6.92	-0.68	0.0127552
ENSG00000146112.7	PPP1R18	chr6:30644165-30655672	5.61	4.93	-0.68	0.0496749
ENSG00000160255.12	ITGB2	chr21:46305867-46351904	6.1	5.41	-0.68	0.0337285
ENSG00000163191.5	S100A11	chr1:151967006-152020383	6.81	6.14	-0.68	0.014592
ENSG00000172354.5	GNB2	chr7:100271153-100276797	5.4	4.72	-0.68	0.0180706
ENSG00000013306.11	SLC25A39	chr17:42396992-42402238	5.13	4.43	-0.69	0.0309307
ENSG00000084207.11	GSTP1	chr11:67351065-67354131	6.28	5.59	-0.69	0.024119
ENSG00000118640.6	VAMP8	chr2:85788684-85809154	5.54	4.85	-0.69	0.0431127
ENSG00000164054.11	SHISA5	chr3:48509196-48542259	6.83	6.14	-0.69	0.00631688
ENSG00000173846.8	PLK3	chr1:45265896-45272957	5.95	5.26	-0.69	0.0443676
ENSG00000198055.6	GRK6	chr5:176829140-176883283	6.07	5.39	-0.69	0.0390251
ENSG00000067225.13	PKM	chr15:72491369-72524164	6.71	6.01	-0.7	0.00350769
ENSG00000115232.9	ITGA4	chr2:182321933-182400914	6.33	5.64	-0.7	0.00868571
ENSG00000159069.9	FBXW5	chr9:139834886-139839148	6.11	5.41	-0.7	0.0198381
ENSG0000	CTDSP2	chr12:58213	6.34	5.64	-0.7	0.040253

0175215.5		709-58240522				8
ENSG00000109736.10	MFSD10	chr4:2932287-2936586	4.16	3.45	-0.71	0.0416438
ENSG00000131634.9	TMEM204	chr16:1543363-1662111	5.95	5.24	-0.71	0.0109714
ENSG00000140406.2	MESDC1	chr15:81293294-81296342	3.66	2.96	-0.71	0.0211478
ENSG00000145901.10	TNIP1	chr5:150409505-150473138	6.13	5.43	-0.71	0.0390251
ENSG00000160185.9	UBASH3A	chr21:43824007-43867791	5.07	4.36	-0.71	0.014592
ENSG00000162032.11	SPSB3	chr16:1826712-1844972	6.16	5.45	-0.71	0.0198381
ENSG00000164733.16	CTSB	chr8:11700032-11726957	7	6.29	-0.71	0.0297796
ENSG00000182179.6	UBA7	chr3:49842639-49851379	5.41	4.71	-0.71	0.00868571
ENSG00000269858.1	EGLN2	chr19:41284120-41314338	6.42	5.71	-0.71	0.0254968
ENSG00000125753.9	VASP	chr19:46009836-46030241	4.81	4.09	-0.72	0.0464
ENSG00000125817.7	CENPB	chr20:3764497-3767337	4.62	3.9	-0.72	0.00350769
ENSG00000213402.2	PTPRCAP	chr11:67202980-67211292	7.24	6.52	-0.72	0.00868571
ENSG00000023902.8	PLEKHO1	chr1:150121372-150136916	6.91	6.18	-0.73	0.0127552
ENSG00000100258.13	LMF2	chr22:50941375-50946135	5.09	4.36	-0.73	0.00350769
ENSG00000108622.6	ICAM2	chr17:62073430-62097994	6.23	5.5	-0.73	0.00350769
ENSG00000142444.6	C19orf52	chr19:11042743-11044211	4.17	3.44	-0.73	0.0451965
ENSG00000142634.8	EFHD2	chr1:15736390-15756839	4.32	3.59	-0.73	0.0127552
ENSG00000238227.3	C9orf69	chr9:139006426-139010731	4.39	3.65	-0.73	0.0180706
ENSG00000009790.10	TRAF3IP3	chr1:209929376-209957904	8.01	7.27	-0.74	0.014592
ENSG00000051523.6	CYBA	chr16:88709690-	7.11	6.36	-0.74	0.00631688

		88717560				
ENSG00000099308.5	MAST3	chr19:18169804-18262498	3.02	2.29	-0.74	0.0496749
ENSG00000110660.10	SLC35F2	chr11:107661716-107799019	3.47	2.73	-0.74	0.0476863
ENSG00000116871.11	MAP7D1	chr1:36621179-36646450	5.51	4.77	-0.74	0.0228
ENSG00000173020.6	ADRBK1	chr11:67033880-67054027	7.3	6.56	-0.74	0.0402538
ENSG00000178996.8	SNX18	chr5:53813588-53842415	3.72	2.98	-0.74	0.0109714
ENSG00000180644.6	PRF1	chr10:72357103-72362531	3.89	3.15	-0.74	0.049062
ENSG00000188229.5	TUBB4B	chr9:140135664-140142222	4.63	3.9	-0.74	0.0228
ENSG00000204525.10	HLA-C	chr6:31236525-31239863	9.72	8.98	-0.74	0.0180706
ENSG00000104964.10	AES	chr19:3052907-3063105	8.51	7.76	-0.75	0.0476863
ENSG00000105404.6	RABAC1	chr19:42460832-42463542	5.69	4.93	-0.75	0.02688
ENSG00000123146.14	CD97	chr19:14491312-14519537	4.99	4.24	-0.75	0.0309307
ENSG00000144579.3	CTDSP1	chr2:219262978-219270664	6.01	5.26	-0.75	0.0164748
ENSG00000147138.1	GPR174	chrX:78426468-78427726	5.22	4.48	-0.75	0.0109714
ENSG00000149091.10	DGKZ	chr11:46354454-46402104	6.11	5.36	-0.75	0.0476863
ENSG00000156860.11	FBRS	chr16:30669751-30682135	3.85	3.1	-0.75	0.014592
ENSG00000116584.11	ARHGEF2	chr1:155916644-155976861	5.92	5.17	-0.76	0.014592
ENSG00000167716.14	WDR81	chr17:1614804-1641893	3.77	3.01	-0.76	0.0297796
ENSG00000196154.7	S100A4	chr1:153516088-153522612	6.96	6.2	-0.76	0.00350769
ENSG00000160271.10	RALGDS	chr9:135973106-136039332	6.13	5.35	-0.77	0.0309307
ENSG0000	ARHGAP4	chrX:153172	5.95	5.17	-0.78	0.003507

0089820.11		820-153193873				69
ENSG00000106351.8	AGFG2	chr7:100136847-100165842	4.28	3.5	-0.78	0.0324267
ENSG00000122224.12	LY9	chr1:160765927-160798045	6.14	5.35	-0.78	0.00350769
ENSG00000159388.5	BTG2	chr1:203274663-203278730	6.39	5.61	-0.78	0.00350769
ENSG00000160410.10	SHKBP1	chr19:41082756-41097305	4.43	3.65	-0.78	0.00350769
ENSG00000196544.6	C17orf59	chr17:8091651-8093564	3.93	3.14	-0.78	0.0228
ENSG00000197956.5	S100A6	chr1:153506078-153508720	8.03	7.25	-0.78	0.00350769
ENSG00000028137.12	TNFRSF1B	chr1:12227059-12269285	4.77	3.98	-0.79	0.00350769
ENSG00000101236.11	RNF24	chr20:3912067-3996229	2.89	2.09	-0.79	0.0443676
ENSG00000122122.9	SASH3	chrX:128913954-128930939	6.53	5.74	-0.79	0.0164748
ENSG00000168067.7	MAP4K2	chr11:64556608-64570713	5.67	4.88	-0.79	0.0109714
ENSG00000172819.12	RARG	chr12:53604353-53626764	3.65	2.87	-0.79	0.0211478
ENSG00000175467.10	SART1	chr11:65729159-65747299	4.2	3.4	-0.79	0.00631688
ENSG00000180448.6	HMHA1	chr19:1065921-1095598	7.88	7.09	-0.79	0.0109714
ENSG00000183741.7	CBX6	chr22:39257454-39268319	4.44	3.66	-0.79	0.00350769
ENSG00000099860.4	GADD45B	chr19:2476124-2478257	4.89	4.1	-0.8	0.02688
ENSG00000108518.7	PFN1	chr17:4848946-4860426	7.84	7.04	-0.8	0.00350769
ENSG00000115306.11	SPTBN1	chr2:54683421-54896812	5.58	4.78	-0.8	0.00350769
ENSG00000149781.8	FERMT3	chr11:63974149-64001824	5.38	4.58	-0.8	0.0451965
ENSG00000160888.6	IER2	chr19:13261228-13265722	3.55	2.74	-0.8	0.0496749
ENSG00000197043.8	ANXA6	chr5:150480272-150537443	6.9	6.09	-0.8	0.00350769
ENSG0000	APOBEC3G	chr22:39436	4.87	4.07	-0.8	0.010971

0239713.3		608-39483748				4
ENSG00000076928.13	ARHGEF1	chr19:42387227-42434302	7.34	6.53	-0.81	0.0127552
ENSG00000118046.10	STK11	chr19:1189405-1238026	4.78	3.97	-0.81	0.0127552
ENSG00000127084.13	FGD3	chr9:95709732-95798518	6.85	6.04	-0.81	0.00350769
ENSG00000188603.12	CLN3	chr16:28467692-28510291	4.66	3.84	-0.82	0.0180706
ENSG00000197540.2	GZMM	chr19:544033-549919	4.67	3.85	-0.82	0.0451965
ENSG00000074370.13	ATP2A3	chr17:3827168-3867736	4.77	3.94	-0.83	0.00350769
ENSG00000101224.13	CDC25B	chr20:3767577-3786762	6.4	5.57	-0.83	0.0337285
ENSG00000115687.9	PASK	chr2:242045513-242123067	6.8	5.97	-0.83	0.0402538
ENSG00000129968.11	ABHD17A	chr19:1876808-1885546	6.13	5.31	-0.83	0.00350769
ENSG00000130522.4	JUND	chr19:18390562-18392432	5.59	4.76	-0.83	0.02688
ENSG00000135926.7	TMBIM1	chr2:219135114-219232822	6.13	5.3	-0.83	0.00350769
ENSG00000172932.10	ANKRD13D	chr11:67056017-67069956	5.38	4.55	-0.83	0.0164748
ENSG00000184271.11	POU6F1	chr12:51580718-51611477	5.27	4.44	-0.83	0.0337285
ENSG00000099849.10	RASSF7	chr11:537526-564021	4.92	4.08	-0.84	0.024119
ENSG00000147813.11	NAPRT1	chr8:144655659-144660783	3.97	3.13	-0.84	0.0180706
ENSG00000165175.11	MID1IP1	chrX:38660684-38665790	5.1	4.26	-0.84	0.00350769
ENSG00000179218.8	CALR	chr19:13049420-13055303	7.66	6.82	-0.84	0.00350769
ENSG00000064666.9	CNN2	chr19:1026211-1039068	6.8	5.95	-0.85	0.00350769
ENSG00000101665.4	SMAD7	chr18:46446222-46477081	5.28	4.42	-0.85	0.014592
ENSG00000104998.2	IL27RA	chr19:14142559-	5.09	4.24	-0.85	0.00350769

		14163743				
ENSG00000122694.11	GLIPR2	chr9:36136731-36163910	3.4	2.55	-0.85	0.0443676
ENSG00000137309.15	HMGA1	chr6:34204649-34214008	5.17	4.32	-0.85	0.00350769
ENSG00000170604.3	IRF2BP1	chr19:46386865-46389376	4.12	3.27	-0.85	0.0109714
ENSG00000077150.13	NFKB2	chr10:104153866-104162281	4.18	3.31	-0.86	0.014592
ENSG00000099785.6	Mar-02	chr19:8478153-8503901	4.29	3.43	-0.86	0.0164748
ENSG00000107404.13	DVL1	chr1:1266693-1284730	3.94	3.08	-0.86	0.00350769
ENSG00000119403.9	PHF19	chr9:123617976-123639606	5.75	4.88	-0.86	0.024119
ENSG00000140368.8	PSTPIP1	chr15:77285699-77329673	5.55	4.69	-0.86	0.00350769
ENSG00000173457.6	PPP1R14B	chr11:64011955-64016966	5.59	4.73	-0.86	0.00350769
ENSG00000198286.5	CARD11	chr7:2945774-3083579	5.35	4.49	-0.86	0.0443676
ENSG00000213654.5	GPSM3	chr6:32158542-32191844	7.1	6.24	-0.86	0.00350769
ENSG00000071655.12	MBD3	chr19:1576638-1592882	4.41	3.53	-0.87	0.00631688
ENSG00000128340.10	RAC2	chr22:37621300-37640488	6.82	5.95	-0.87	0.00350769
ENSG00000139641.8	ESYT1	chr12:56511942-56538455	6.18	5.31	-0.87	0.00350769
ENSG00000163545.7	NUAK2	chr1:205271186-205290883	3.82	2.95	-0.87	0.00350769
ENSG00000175550.3	DRAP1	chr11:65686727-65689032	5.74	4.86	-0.87	0.00350769
ENSG00000182500.7	ORAI1	chr12:122064454-122080583	4.28	3.41	-0.87	0.00868571
ENSG00000113088.5	GZMK	chr5:54273691-54330398	4.02	3.13	-0.88	0.0180706
ENSG00000117643.10	MAN1C1	chr1:25943958-26112698	5.35	4.47	-0.88	0.00350769
ENSG00000165915.9	SLC39A13	chr11:47404698-47438047	4.58	3.7	-0.88	0.0127552

ENSG00000011132.7	APBA3	chr19:3750877-3761697	3.88	2.99	-0.89	0.00868571
ENSG00000026025.9	VIM	chr10:17256237-17279592	8.93	8.03	-0.89	0.00350769
ENSG00000077238.9	IL4R	chr16:27324988-27376099	5.59	4.7	-0.89	0.014592
ENSG00000110719.5	TCIRG1	chr11:67806482-67888911	5.96	5.07	-0.89	0.00350769
ENSG00000167468.12	GPX4	chr19:1103935-1106787	7.12	6.23	-0.89	0.00350769
ENSG00000170638.5	TRABD	chr22:50624343-50638027	5.81	4.92	-0.89	0.00350769
ENSG00000095370.15	SH2D3C	chr9:130500595-130541020	3.85	2.95	-0.9	0.00350769
ENSG00000101298.9	SNPH	chr20:1206699-1289972	3.88	2.98	-0.9	0.00350769
ENSG00000142765.13	SYTL1	chr1:27668512-27680421	5.38	4.48	-0.9	0.00350769
ENSG00000166925.4	TSC22D4	chr7:100054237-100076902	4.8	3.9	-0.9	0.00350769
ENSG00000182866.12	LCK	chr1:32716839-32751766	8.2	7.3	-0.9	0.00350769
ENSG00000197471.6	SPN	chr16:29674299-29710020	5.73	4.83	-0.9	0.00631688
ENSG00000197530.8	MIB2	chr1:1550794-1565990	3.95	3.05	-0.91	0.0378566
ENSG00000100599.11	RIN3	chr14:92980117-93155339	3.38	2.46	-0.92	0.00350769
ENSG00000102879.11	CORO1A	chr16:30194147-30200397	8.18	7.25	-0.92	0.00350769
ENSG00000126432.9	PRDX5	chr11:64085559-64089283	5.26	4.34	-0.92	0.00350769
ENSG00000136490.4	LIMD2	chr17:61699774-61778532	8.3	7.38	-0.92	0.0109714
ENSG00000213145.5	CRIP1	chr14:105952653-105955284	6.64	5.71	-0.93	0.00350769
ENSG00000019582.10	CD74	chr5:149781199-149792492	6.25	5.31	-0.94	0.00350769
ENSG00000183889.8	PKD1P1	chr16:16411300-16444447	5.87	4.93	-0.94	0.00350769
ENSG00000206503.7	HLA-A	chr6:29909036-	10.73	9.79	-0.94	0.00868571

		29913661				
ENSG00000223865.6	HLA-DPB1	chr6:33032345-33054978	4.88	3.94	-0.94	0.0228
ENSG00000225663.3	FAM195B	chr17:79780286-79791178	5.27	4.33	-0.94	0.00350769
ENSG00000008256.11	CYTH3	chr7:6201406-6312275	2.17	1.22	-0.95	0.0416438
ENSG00000064687.8	ABCA7	chr19:1040101-1065568	5.48	4.53	-0.95	0.024119
ENSG00000108840.11	HDAC5	chr17:42154113-42201070	4.52	3.57	-0.95	0.0109714
ENSG00000161638.6	ITGA5	chr12:54747444-54891472	6.1	5.15	-0.95	0.014592
ENSG00000090238.7	YPEL3	chr16:30103634-30115437	7.01	6.05	-0.96	0.00350769
ENSG00000100241.15	SBF1	chr22:50885183-50913454	5.93	4.97	-0.96	0.0254968
ENSG00000171222.6	SCAND1	chr20:34541538-34547394	5.05	4.09	-0.96	0.00631688
ENSG00000104783.7	KCNN4	chr19:44270684-44285409	3.68	2.71	-0.97	0.00868571
ENSG00000178199.9	ZC3H12D	chr6:149768793-149806197	6.46	5.48	-0.97	0.00350769
ENSG00000005844.13	ITGAL	chr16:30483978-30534506	6.69	5.71	-0.98	0.00350769
ENSG00000073350.9	LLGL2	chr17:73521160-73571289	3.04	2.05	-0.98	0.0180706
ENSG00000112242.10	E2F3	chr6:20402397-20493941	3.65	2.67	-0.98	0.00350769
ENSG00000125898.7	FAM110A	chr20:814357-838106	2.93	1.95	-0.98	0.0496749
ENSG00000128185.5	DGCR6L	chr22:20301798-20307603	5.01	4.02	-0.98	0.00350769
ENSG00000178038.12	ALS2CL	chr3:46710678-46735191	2.97	1.99	-0.98	0.00631688
ENSG00000178951.4	ZBTB7A	chr19:4044361-4066943	3.28	2.31	-0.98	0.00631688
ENSG00000061273.13	HDAC7	chr12:48099867-48231681	5.3	4.31	-0.99	0.00631688
ENSG00000177548.8	RABEP2	chr16:28889725-28950667	3.29	2.3	-0.99	0.0451965
ENSG0000	FKBP8	chr19:18642	6.76	5.76	-1	0.003507

0105701.9		227-18654863				69
ENSG00000125910.4	S1PR4	chr19:3172343-3180329	7.58	6.57	-1	0.00350769
ENSG00000130787.9	HIP1R	chr12:123319972-123347507	5.24	4.24	-1	0.00350769
ENSG00000160408.10	ST6GALNAC6	chr9:130628758-130667687	6.43	5.43	-1	0.00350769
ENSG00000162783.8	IER5	chr1:181057637-181059977	3.4	2.4	-1	0.00631688
ENSG00000172164.9	SNTB1	chr8:121547984-121825513	3.6	2.6	-1	0.00631688
ENSG00000147443.8	DOK2	chr8:21766383-21771371	4.21	3.18	-1.03	0.00631688
ENSG00000162676.7	GFI1	chr1:92940318-92952433	3.93	2.89	-1.04	0.0127552
ENSG00000175602.2	CCDC85B	chr11:65657874-65659105	3.6	2.57	-1.04	0.0198381
ENSG00000184640.12	Sep-09	chr17:75253818-75496678	8.5	7.46	-1.04	0.00350769
ENSG00000136286.10	MYO1G	chr7:45002264-45018697	3.93	2.87	-1.06	0.00350769
ENSG00000175130.6	MARCKSL1	chr1:32799432-32801980	3.32	2.25	-1.07	0.00868571
ENSG00000090382.2	LYZ	chr12:69742120-69748014	5.21	4.13	-1.08	0.00350769
ENSG00000115756.8	HPCAL1	chr2:10443014-10567743	5.42	4.33	-1.08	0.00350769
ENSG00000128271.15	ADORA2A	chr22:24813846-24924358	4.18	3.1	-1.08	0.00350769
ENSG00000162591.11	MEGF6	chr1:3406483-3528059	4.67	3.59	-1.08	0.00350769
ENSG00000253522.1	hsa-mir-146a	chr5:159895274-159914433	3.03	1.95	-1.08	0.00350769
ENSG00000154165.3	GPR15	chr3:98250742-98251960	6.01	4.92	-1.09	0.00350769
ENSG00000175463.7	TBC1D10C	chr11:67159175-67193078	7.09	6	-1.09	0.00350769
ENSG00000100242.11	SUN2	chr22:39101727-39190203	8.19	7.09	-1.1	0.00350769
ENSG0000	PSD4	chr2:113914	5.66	4.56	-1.1	0.003507

0125637.10		901-113960814				69
ENSG00000139626.10	ITGB7	chr12:53585101-53601091	6.6	5.49	-1.1	0.00350769
ENSG00000160685.9	ZBTB7B	chr1:154975126-154990998	4.64	3.54	-1.11	0.00631688
ENSG00000135127.7	CCDC64	chr12:120427672-120532298	6.25	5.12	-1.13	0.00350769
ENSG00000176973.7	FAM89B	chr11:65337900-65341669	5.24	4.11	-1.13	0.00350769
ENSG00000185736.11	ADARB2	chr10:1228072-1779670	3.62	2.49	-1.13	0.0198381
ENSG00000186350.8	RXRA	chr9:137208943-137332431	1.48	0.35	-1.13	0.0309307
ENSG00000196924.10	FLNA	chrX:153576893-153603006	6.54	5.41	-1.13	0.0127552
ENSG00000100299.12	ARSA	chr22:51063445-51066607	3.75	2.6	-1.15	0.00350769
ENSG00000165272.10	AQP3	chr9:33441151-33447609	6.23	5.08	-1.15	0.00350769
ENSG00000063180.4	CA11	chr19:49141198-49149569	2.43	1.28	-1.16	0.02688
ENSG00000131759.13	RARA	chr17:38465443-38513094	3.9	2.75	-1.16	0.0180706
ENSG00000158717.6	RNF166	chr16:88762902-88772829	6.5	5.35	-1.16	0.00350769
ENSG00000138172.6	CALHM2	chr10:105206542-105222452	4.37	3.2	-1.17	0.00350769
ENSG00000162496.4	DHRS3	chr1:12627938-12677737	5.47	4.3	-1.17	0.00350769
ENSG00000105122.7	RASAL3	chr19:15562437-15575377	6.29	5.1	-1.18	0.00350769
ENSG00000163508.8	EOMES	chr3:27757439-27764206	3.75	2.57	-1.18	0.00350769
ENSG00000167797.3	CDK2AP2	chr11:67273967-67276102	4.75	3.58	-1.18	0.00350769
ENSG00000135047.10	CTSL1	chr9:90340433-90346308	6.5	5.32	-1.19	0.00350769
ENSG00000136280.11	CCM2	chr7:45039073-45116068	6.7	5.51	-1.19	0.00350769

ENSG00000164088.13	PPM1M	chr3:52279840-52284613	4.84	3.65	-1.19	0.00350769
ENSG00000188322.4	SBK1	chr16:28303839-28335170	1.01	-0.19	-1.19	0.0164748
ENSG00000215788.5	TNFRSF25	chr1:6484847-6580121	7.71	6.52	-1.19	0.00350769
ENSG00000142227.6	EMP3	chr19:48799713-48833810	6.99	5.79	-1.2	0.00350769
ENSG00000135362.9	PRR5L	chr11:36317837-36486754	4.68	3.46	-1.21	0.00350769
ENSG00000169442.4	CD52	chr1:26605666-26647014	9.79	8.58	-1.21	0.00350769
ENSG00000088899.10	PROSAPIP1	chr20:3143262-3154192	3.2	1.98	-1.22	0.00350769
ENSG00000102760.12	RGCC	chr13:42031694-42045018	7.27	6.05	-1.22	0.00350769
ENSG00000112667.8	DNPH1	chr6:43193366-43197222	5.47	4.25	-1.22	0.00350769
ENSG00000056558.6	TRAF1	chr9:123664670-123691451	6.33	5.1	-1.23	0.00350769
ENSG00000110944.4	IL23A	chr12:56732662-56734193	4.81	3.58	-1.23	0.00350769
ENSG00000129250.7	KIF1C	chr17:4901242-4931696	3.62	2.39	-1.23	0.0127552
ENSG00000185347.13	C14orf80	chr14:105956191-105965912	2.48	1.25	-1.23	0.0297796
ENSG00000106003.8	LFNG	chr7:2552162-2568811	5.87	4.62	-1.25	0.00350769
ENSG00000151651.11	ADAM8	chr10:135075906-135090372	3.85	2.61	-1.25	0.00631688
ENSG00000141556.15	TBCD	chr17:80709939-81009686	6.7	5.44	-1.26	0.00350769
ENSG00000135916.11	ITM2C	chr2:231729353-231743963	4.13	2.86	-1.27	0.00350769
ENSG00000170004.12	CHD3	chr17:7760002-7816078	6.01	4.74	-1.27	0.00350769
ENSG00000141858.7	SAMD1	chr19:14198651-14201848	4.13	2.85	-1.28	0.00350769
ENSG00000168056.10	LTBP3	chr11:65306275-65326401	4.56	3.28	-1.28	0.00350769
ENSG00000267519.1	CTD-3252C9.4	chr19:13945329-	2.27	0.99	-1.28	0.0164748

		13947103				
ENSG00000142303.9	ADAMTS10	chr19:8645125-8675585	3.84	2.53	-1.31	0.00868571
ENSG00000106123.7	EPHB6	chr7:142552791-142568847	3.75	2.42	-1.33	0.00868571
ENSG00000161570.4	CCL5	chr17:34195970-34212867	7.06	5.73	-1.33	0.00350769
ENSG00000158856.13	EPB49	chr8:21906505-21940038	0.67	-0.68	-1.34	0.0324267
ENSG00000130592.9	LSP1	chr11:1874199-1913497	7.65	6.3	-1.35	0.00350769
ENSG00000124641.10	MED20	chr6:41873091-41888877	2.58	1.22	-1.36	0.0127552
ENSG00000006704.6	GTF2IRD1	chr7:73868119-74016931	1.48	0.11	-1.37	0.0228
ENSG00000124762.8	CDKN1A	chr6:36644304-36655116	1.67	0.29	-1.38	0.014592
ENSG00000197077.8	KIAA1671	chr22:25348696-25593415	2.55	1.15	-1.39	0.00350769
ENSG00000241657.1	TRBV11-2	chr7:142197569-142198069	5.35	3.96	-1.39	0.00350769
ENSG00000251060.1	U66061.31	chr7:142413073-142426272	2.16	0.77	-1.39	0.00350769
ENSG00000029534.15	ANK1	chr8:41510738-41754280	2.08	0.68	-1.41	0.00631688
ENSG00000181588.15	MEX3D	chr19:1554667-1568057	1	-0.41	-1.41	0.049062
ENSG00000213420.3	GPC2	chr7:99767228-99774992	4.07	2.66	-1.41	0.02688
ENSG00000232810.3	TNF	chr6:31543343-31546113	2.37	0.96	-1.41	0.00631688
ENSG00000135114.8	OASL	chr12:121458094-121477045	3.47	2.04	-1.43	0.00350769
ENSG00000184613.6	NELL2	chr12:44902057-45315631	7.1	5.67	-1.43	0.00350769
ENSG00000123159.11	GIPC1	chr19:14588571-14606944	4.41	2.96	-1.44	0.00350769
ENSG00000183473.5	SSTR3	chr22:37600277-37608362	2.4	0.96	-1.44	0.00350769
ENSG00000246526.2	RP11-539L10.2	chr4:6689174-6692246	2.65	1.19	-1.46	0.00350769
ENSG0000	FCER1A	chr1:159259	2.31	0.82	-1.49	0.024119

0179639.6		503-159278014				
ENSG00000260231.1	RP4-659J6.2	chr7:139877060-139879440	3.39	1.89	-1.5	0.00350769
ENSG00000104081.9	BMF	chr15:40380090-40401093	2.95	1.44	-1.51	0.0109714
ENSG00000145911.5	N4BP3	chr5:177540443-177553088	0.09	-1.42	-1.51	0.00868571
ENSG00000213658.5	LAT	chr16:28996146-29002104	7.08	5.56	-1.52	0.00350769
ENSG00000130748.6	TMEM160	chr19:47549164-47551888	4.34	2.81	-1.53	0.00350769
ENSG00000163235.11	TGFA	chr2:70674411-70781325	2.22	0.68	-1.55	0.0211478
ENSG00000104856.9	RELB	chr19:45504687-45541452	2.51	0.95	-1.56	0.00350769
ENSG00000168758.6	SEMA4C	chr2:97525452-97536494	5.61	4.05	-1.56	0.00350769
ENSG00000197093.6	GAL3ST4	chr7:99756866-99766373	3.54	1.89	-1.65	0.00350769
ENSG00000215475.3	SIAH3	chr13:46354404-46425871	2.54	0.88	-1.66	0.00350769
ENSG00000100097.7	LGALS1	chr22:38071614-38075813	4.58	2.89	-1.69	0.00350769
ENSG00000101412.9	E2F1	chr20:32263488-32274210	1.21	-0.49	-1.7	0.00350769
ENSG00000197696.5	NMB	chr15:85198359-85201794	1.7	0	-1.71	0.0366545
ENSG00000173114.8	LRRN3	chr7:110303109-111202573	5.03	3.32	-1.72	0.00350769
ENSG00000148737.11	TCF7L2	chr10:114710008-114927437	0.94	-0.81	-1.75	0.0352464
ENSG00000137331.11	IER3	chr6:30710975-30712331	2.85	0.99	-1.86	0.00350769
ENSG00000183691.4	NOG	chr17:54671059-54672951	5.82	3.96	-1.86	0.00350769
ENSG00000186810.7	CXCR3	chrX:70835765-70838367	3.47	1.59	-1.88	0.00350769
ENSG00000085117.7	CD82	chr11:44585976-44641913	4.3	2.41	-1.89	0.00350769

ENSG00000157551.12	KCNJ15	chr21:39493544-39673748	0.18	-1.72	-1.89	0.00350769
ENSG00000152268.8	SPON1	chr11:13983913-14295237	5.18	3.22	-1.97	0.00350769
ENSG00000227507.2	LTB	chr6:31548301-31550299	7.78	5.75	-2.03	0.00350769
ENSG00000178773.10	CPNE7	chr16:89642175-89663654	2.91	0.87	-2.05	0.00868571
ENSG00000117318.8	ID3	chr1:23884408-23886285	5.94	3.82	-2.12	0.00350769
ENSG00000100628.7	ASB2	chr14:94400498-94443137	1.6	-0.54	-2.14	0.0164748
ENSG00000182379.9	NXPH4	chr12:57610577-57620232	0.72	-1.75	-2.47	0.0378566

Over 100 genes are differentially expressed in between CD4 T-cells stimulated with synovial fluid and those stimulated with synovial fluid and tofacitinib

Test ID	Gene	Locus	log2(SF FPKM)	log2(SF + Tofa FPKM)	log2(Ratio)	q Value
ENSG00000248309.1	CTC-454M9.1	chr5:87803362-88762215	0.24	-4.54	-4.78	0.0137537
ENSG00000184557.3	SOCS3	chr17:76352863-76356158	4.98	2.45	-2.53	0.0137537
ENSG00000113916.13	BCL6	chr3:187416046-187463515	3.94	1.61	-2.33	0.0137537
ENSG00000111863.7	ADTRP	chr6:11713887-11807279	6.76	5.11	-1.65	0.0137537
ENSG00000181827.10	RFX7	chr15:56379665-56535483	3.6	2.14	-1.46	0.0137537
ENSG00000150347.10	ARID5B	chr10:63661058-63856703	5.26	4.05	-1.21	0.0137537
ENSG00000198879.7	SFMBT2	chr10:7200585-7453450	4.97	3.79	-1.19	0.0345442
ENSG00000176542.5	KIAA2018	chr3:113367231-113415493	3.29	2.19	-1.1	0.0399301
ENSG00000163564.10	PYHIN1	chr1:158900585-158946844	6.05	5.1	-0.96	0.0408826
ENSG00000169446.4	MMGT1	chrX:135044228-135056222	4.54	3.59	-0.95	0.0225061
ENSG00000100258.13	LMF2	chr22:50941375-50946135	4.39	5.25	0.86	0.044435
ENSG00000186517.9	ARHGAP30	chr1:161016737-161039760	5.43	6.29	0.86	0.0399301
ENSG00000173762.3	CD7	chr17:80272743-80275478	5.99	6.86	0.87	0.0137537
ENSG00000177105.9	RHOG	chr11:3848207-3862213	4.75	5.64	0.89	0.0408826
ENSG00000214753.2	HNRNPUL2	chr11:62480101-62494821	3.6	4.49	0.89	0.0408826
ENSG00000149930.13	TAOK2	chr16:29984961-30003582	3.36	4.28	0.93	0.0399301
ENSG00000159753.9	RLTPR	chr16:67678821-67694713	3.85	4.79	0.94	0.0408826
ENSG00000170638.5	TRABD	chr22:50624343-50638027	4.94	5.89	0.95	0.0137537
ENSG00000177764.6	ZCCHC3	chr20:277736-280965	3.14	4.1	0.96	0.0408826
ENSG00000162496.4	DHRS3	chr1:12627938-12677737	4.33	5.31	0.98	0.0345442
ENSG00000254470.2	AP5B1	chr11:65543363-65548273	2.54	3.53	0.99	0.0408826
ENSG00000089820.11	ARHGAP4	chrX:153172820-153193873	5.2	6.2	1	0.044435
ENSG00000158106.8	RHPN1	chr8:144451056-144466390	2.64	3.64	1	0.044435
ENSG00000070047.7	PHRF1	chr11:576485-612222	3.59	4.61	1.01	0.0137537
ENSG00000068831.14	RASGRP2	chr11:64494382-64512928	7.17	8.19	1.02	0.0499294

ENSG00000170604.3	IRF2BP1	chr19:46386865-46389376	3.3	4.31	1.02	0.0293171
ENSG00000101298.9	SNPH	chr20:1206699-1289972	3	4.03	1.03	0.0137537
ENSG00000176438.8	SYNE3	chr14:95883830-95942173	3.14	4.17	1.03	0.0345442
ENSG00000215440.7	NPEPL1	chr20:57264186-57290900	3.33	4.36	1.03	0.0399301
ENSG00000125910.4	S1PR4	chr19:3172343-3180329	6.6	7.64	1.04	0.0137537
ENSG00000197976.6	AKAP17A	chrX:1710485-1721407	5.37	6.41	1.04	0.0137537
ENSG00000215788.5	TNFRSF25	chr1:6484847-6580121	6.55	7.59	1.04	0.0408826
ENSG00000118046.10	STK11	chr19:1189405-1238026	4	5.05	1.05	0.0408826
ENSG00000158545.11	ZC3H18	chr16:88636788-88698374	3.12	4.16	1.05	0.0293171
ENSG00000110046.8	ATG2A	chr11:64662006-64684722	2.28	3.34	1.06	0.0293171
ENSG00000185950.7	IRS2	chr13:110406183-110438915	1.83	2.91	1.08	0.0293171
ENSG00000119669.3	IRF2BPL	chr14:77490887-77495034	3.11	4.2	1.09	0.0137537
ENSG00000148296.5	SURF6	chr9:136197551-136203235	3.48	4.57	1.1	0.044435
ENSG00000149527.12	PLCH2	chr1:2357418-2436969	2.35	3.45	1.1	0.0345442
ENSG00000137818.7	RPLP1	chr15:69745122-69748255	9.54	10.66	1.11	0.0408826
ENSG00000169871.8	TRIM56	chr7:100728719-100735017	3.74	4.84	1.11	0.0399301
ENSG00000171206.8	TRIM8	chr10:104404252-104418075	4.39	5.51	1.13	0.0137537
ENSG00000105122.7	RASAL3	chr19:15562437-15575377	5.13	6.29	1.15	0.0345442
ENSG00000127528.5	KLF2	chr19:16435627-16438685	5.83	6.97	1.15	0.0408826
ENSG00000088899.10	PROSAPIP1	chr20:3143262-3154192	2.01	3.17	1.16	0.0293171
ENSG00000100599.11	RIN3	chr14:92980117-93155339	2.49	3.65	1.16	0.0408826
ENSG00000166925.4	TSC22D4	chr7:100054237-100076902	3.92	5.09	1.17	0.0293171
ENSG00000139718.6	SETD1B	chr12:122242085-122270562	2.56	3.74	1.18	0.0293171
ENSG00000142765.13	SYTL1	chr1:27668512-27680421	4.5	5.7	1.19	0.0137537
ENSG00000166341.6	DCHS1	chr11:6642555-6677085	0.73	1.92	1.19	0.0225061
ENSG00000167716.14	WDR81	chr17:1614804-1641893	3.04	4.25	1.21	0.0293171
ENSG00000105287.8	PRKD2	chr19:47150868-47220384	3.05	4.27	1.22	0.0408826
ENSG00000106003.8	LFNG	chr7:2552162-2568811	4.65	5.88	1.23	0.0137537
ENSG00000175467.10	SART1	chr11:65729159-65747299	3.43	4.67	1.24	0.0137537
ENSG000001	ALS2CL	chr3:46710678-	2.02	3.28	1.26	0.034

78038.12		46735191				5442
ENSG00000161847.9	RAVER1	chr19:10426887-10444316	1.8	3.08	1.27	0.0225061
ENSG00000203485.8	INF2	chr14:105155942-105185942	2.74	4.01	1.27	0.044435
ENSG00000140368.8	PSTPIP1	chr15:77285699-77329673	4.71	5.99	1.28	0.0137537
ENSG00000074370.13	ATP2A3	chr17:3827168-3867736	3.96	5.25	1.29	0.0137537
ENSG00000227507.2	LTB	chr6:31548301-31550299	5.76	7.08	1.31	0.0137537
ENSG00000061273.13	HDAC7	chr12:48099867-48231681	4.33	5.65	1.32	0.044435
ENSG00000255026.1	RP11-326C3.2	chr11:287304-288987	5.7	7.02	1.32	0.0225061
ENSG00000157637.8	SLC38A10	chr17:79218799-79269347	3.32	4.66	1.34	0.0499294
ENSG00000147813.11	NAPRT1	chr8:144655659-144660783	3.14	4.49	1.35	0.0225061
ENSG00000123143.8	PKN1	chr19:14543864-14582678	4.88	6.25	1.36	0.0137537
ENSG00000176248.7	ANAPC2	chr9:140069235-140082989	4.1	5.46	1.36	0.0225061
ENSG00000089639.6	GMIP	chr19:19740284-19754476	4.46	5.88	1.42	0.0137537
ENSG00000182979.13	MTA1	chr14:105886158-105937066	3.69	5.14	1.44	0.0225061
ENSG00000182154.7	MRPL41	chr9:140445650-140447007	3.99	5.44	1.45	0.0137537
ENSG00000114626.13	ABTB1	chr3:127391780-127399768	3.88	5.36	1.48	0.0137537
ENSG00000198467.8	TPM2	chr9:35681988-35691017	3.53	5.02	1.48	0.0137537
ENSG00000197530.8	MIB2	chr1:1550794-1565990	3.08	4.57	1.49	0.0225061
ENSG00000103326.6	SOLH	chr16:577716-604636	2.67	4.17	1.5	0.0137537
ENSG00000130522.4	JUND	chr19:18390562-18392432	4.79	6.29	1.5	0.0137537
ENSG00000177548.8	RABEP2	chr16:28889725-28950667	2.32	3.81	1.5	0.0225061
ENSG00000220008.2	LINGO3	chr19:2289773-2308156	2.65	4.15	1.5	0.0137537
ENSG00000059122.12	FLYWCH1	chr16:2961937-3004277	4.01	5.56	1.55	0.0293171
ENSG00000205336.6	GPR56	chr16:57644563-57698944	0.62	2.17	1.55	0.0408826
ENSG00000077454.11	LRCH4	chr7:100171633-100205798	5.53	7.09	1.56	0.0137537
ENSG00000076928.13	ARHGEF1	chr19:42387227-42434302	6.55	8.14	1.59	0.0399301
ENSG00000131584.14	ACAP3	chr1:1227755-1243398	1.69	3.28	1.59	0.0345442
ENSG00000168071.17	CCDC88B	chr11:64107694-64125006	4.21	5.81	1.59	0.0137537
ENSG00000155034.14	FBXL18	chr7:5520188-5553429	0.93	2.54	1.61	0.0225061
ENSG00000111676.9	ATN1	chr12:7033625-7051484	0.97	2.62	1.65	0.0137537

ENSG00000175602.2	CCDC85B	chr11:65657874-65659105	2.59	4.25	1.66	0.0293171
ENSG00000183889.8	PKD1P1	chr16:16411300-16444447	4.96	6.62	1.67	0.0137537
ENSG00000205307.6	SAP25	chr7:100169854-100171270	4.04	5.7	1.67	0.0137537
ENSG00000228434.1	AC004951.6	chr7:44040487-44049721	3.43	5.1	1.67	0.0399301
ENSG00000178951.4	ZBTB7A	chr19:4044361-4066943	2.33	4.01	1.68	0.0137537
ENSG00000162591.11	MEGF6	chr1:3406483-3528059	3.61	5.32	1.71	0.0137537
ENSG00000254144.1	RP11-661A12.4	chr8:144624142-144631899	7.18	8.89	1.71	0.0137537
ENSG00000142235.4	LMTK3	chr19:48988527-49016446	0.43	2.16	1.73	0.0225061
ENSG00000061938.12	TNK2	chr3:195590234-195638816	3.99	5.74	1.75	0.0137537
ENSG00000079313.7	REXO1	chr19:1815247-1848452	2.43	4.19	1.76	0.0137537
ENSG00000136213.7	CHST12	chr7:2443222-2474242	1.66	3.42	1.76	0.0137537
ENSG00000063245.10	EPN1	chr19:56186591-56249768	2.49	4.26	1.77	0.0137537
ENSG00000113504.15	SLC12A7	chr5:1050498-1112150	1.95	3.76	1.81	0.0137537
ENSG00000126464.9	PRR12	chr19:50094899-50129696	1.63	3.51	1.88	0.0137537
ENSG00000153443.8	UBALD1	chr16:4658883-4665028	1.6	3.52	1.92	0.0345442
ENSG00000105193.4	RPS16	chr19:39923846-39926588	7.25	9.21	1.95	0.0137537
ENSG00000267874.1	CTD-2527I21.9	chr19:35521587-35531352	3.93	5.88	1.95	0.0137537
ENSG00000162458.8	FBLIM1	chr1:16083101-16113089	-0.1	1.86	1.97	0.0137537
ENSG00000171604.7	CXXC5	chr5:139026883-139063467	0.39	2.38	1.99	0.0408826
ENSG00000267598.1	CTC-250I14.6	chr19:13261228-13265722	3.75	5.75	2	0.044435
ENSG00000128011.4	LRFN1	chr19:39797207-39805976	0.67	2.69	2.02	0.0137537
ENSG00000227184.3	EPPK1	chr8:144939496-144952632	1.82	3.85	2.03	0.0137537
ENSG00000261221.1	ZNF865	chr19:56116770-56128635	0.47	2.53	2.06	0.0137537
ENSG00000133250.9	ZNF414	chr19:8575461-8579044	1.09	3.19	2.1	0.0408826
ENSG00000168056.10	LTBP3	chr11:65306275-65326401	3.31	5.47	2.16	0.0137537
ENSG00000130653.10	PNPLA7	chr9:140354403-140444986	0.52	2.68	2.17	0.0137537
ENSG00000111319.8	SCNN1A	chr12:6456008-6500729	0.37	2.57	2.2	0.0408826
ENSG00000167685.10	ZNF444	chr19:56643967-56672262	2.39	4.67	2.28	0.0137537
ENSG00000257900.2	RP11-454K7.1	chr14:45846470-45858489	2.06	4.38	2.32	0.0345442
ENSG000002	RP11-	chr3:186525480	0.56	2.92	2.36	0.022

32233.1	573D15.2	-186543310				5061
ENSG00000269468.1	AC004824.2	chr1:17634689-17690499	3.6	6.09	2.49	0.0137537
ENSG00000124194.11	GDAP1L1	chr20:42875886-42909013	-1.41	1.33	2.74	0.0345442
ENSG00000186081.7	KRT5	chr12:52908358-52914471	-1.4	2.37	3.76	0.0137537
ENSG00000249612.1	AC063980.1	chr5:135712035-135732730	-10	0.81	10.81	0.044435
ENSG00000165685.4	TMEM52B	chr12:10310901-10344400	-10	1.68	11.68	0.0137537

R script for analysis of qPCR data using HTqPCR package

```
library("HTqPCR")

raw <- readCtData(files=c("B15590010_1_4.sdm-Amplification Data.txt", "B15590008_5_8.sdm-
Amplification Data.txt", "B15590011_9_12.sdm-Amplification Data.txt", "B15590009_27_30.sdm-
Amplification Data.txt"), head=T, na.value=40, column.info=c(Ct=6, position=1, feature=4, type=5),
n.features=96, n.data=4)

p <- read.csv("pdata1.csv", row.names="Sample")

pData(raw) <-p

### Re-read featurenames - May need to alter in excel and then read in a data frame with gene names
alone.

z<- read.table(file="test.txt")

featureNames(raw) <- as.character(z[ ,])



### Plot QC files
pdf(file="Samples 27-30.pdf")

plotCtCard(raw, col.range = c(10, 35), well.size = 2.6, card=c(13,14,15,16), main="Samples 27-30")

dev.off()

### Plot Ct overview of Housekeeping genes

featureNames(raw)[c(11)] -> b

plotCtOverview(raw, genes=b, xlim=c(0,50), conf.int=TRUE,
ylim=c(0,55))

plotCtOverview(raw, genes=b, xlim=c(0,50), groups=pData(raw)$Patient, conf.int=TRUE,
ylim=c(0,55))

### Normalise files

q.norm <- normalizeCtData(raw, norm="quantile")
d18s.norm <- normalizeCtData(raw, norm="deltaCt", deltaCt.genes=c("X18S"))
dGAPDH.norm<- normalizeCtData(raw, norm="deltaCt", deltaCt.genes=c("GAPDH"))
dGPI.norm<- normalizeCtData(raw, norm="deltaCt", deltaCt.genes=c("GPI"))
dTBP.norm<- normalizeCtData(raw, norm="deltaCt", deltaCt.genes=c("TBP"))

dall.norm <- normalizeCtData(raw, norm="deltaCt", deltaCt.genes=c("18S", "GAPDH", "GPI", "TBP"))

d <- exprs(raw)
e <- t(d)
f <- cbind(e, pData(raw))
f$SampleA <- factor(f$SampleA, levels=f$SampleA[order(f$Order)])


### Plot Correlation Heatmaps of raw and normalised data

pdf(file="Ct correlation of raw data.pdf")
plotCtCor(raw, main="Ct correlation of raw data")
dev.off()
```

```
pdf(file="Ct correlation of 18s normalised data.pdf")
plotCtCor(d18s.norm, main="Ct correlation of 18s norm data")
dev.off()

pdf(file="Ct correlation of GAPDH normalised data.pdf")
plotCtCor(dGAPDH.norm, main="Ct correlation of GAPDH norm data")
dev.off()

pdf(file="Ct correlation of GPI normalised data.pdf")
plotCtCor(dGPI.norm, main="Ct correlation of GPI norm data")
dev.off()

pdf(file="Ct correlation of TBP normalised data.pdf")
plotCtCor(dTBP.norm, main="Ct correlation of TBP norm data")
dev.off()

pdf(file="Ct correlation of quantile normalised data.pdf")
plotCtCor(q.norm, main="Ct correlation of quantile norm data")
dev.off()

pdf(file="Ct correlation of 4 gene normalised data.pdf")
plotCtCor(dall.norm, main="Ct correlation of 4 gene norm data")
dev.off()

### Plot Ct Densitites of raw and normalised data

pdf(file="Ct densities of raw data.pdf")
plotCtDensity(raw, main="Ct densities of raw data", lty=1:16)
dev.off()

pdf(file="Ct densities of 18s normalised data.pdf")
plotCtDensity(d18s.norm, main="Ct densities of 18s norm data", lty=1:16)
dev.off()

pdf(file="Ct densities of GAPDH normalised data.pdf")
plotCtDensity(dGAPDH.norm, main="Ct densities of GAPDH norm data", lty=1:16)
dev.off()

pdf(file="Ct densities of GPI normalised data.pdf")
plotCtDensity(dGPI.norm, main="Ct densities of GPI norm data", lty=1:16)
dev.off()

pdf(file="Ct densities of TBP normalised data.pdf")
plotCtDensity(dTBP.norm, main="Ct densities of TBP norm data", lty=1:16)
dev.off()

pdf(file="Ct densities of quantile normalised data.pdf")
plotCtDensity(q.norm, main="Ct densities of quantile norm data", lty=1:16)
dev.off()

pdf(file="Ct densities of 4 gene normalised data.pdf")
plotCtDensity(dall.norm, main="Ct densities of 4 gene norm data", lty=1:16)
dev.off()

### PlotCt Pairs

pdf(file="Ct pair-wise correlation of raw data.pdf", height=10, width=10)
plotCtPairs(raw, col="type", diag=TRUE, main="Ct pair-wise correlation of Raw data")
dev.off()

pdf(file="Ct pair-wise correlation of 18s normalised data.pdf", height=10, width=10)
plotCtPairs(d18s.norm, col="type", diag=TRUE, main="Ct pair-wise correlation of 18s Normalised
data")
```

```
dev.off()

pdf(file="Ct pair-wise correlation of GAPDH normalised data.pdf", height=10, width=10)
plotCtPairs(dGAPDH.norm, col="type", diag=TRUE, main="Ct pair-wise correlation of GAPDH
Normalised data")
dev.off()

pdf(file="Ct pair-wise correlation of GPI normalised data.pdf", height=10, width=10)
plotCtPairs(dGPI.norm, col="type", diag=TRUE, main="Ct pair-wise correlation of GPI Normalised
data")
dev.off()

pdf(file="Ct pair-wise correlation of TBP normalised data.pdf", height=10, width=10)
plotCtPairs(dTBP.norm, col="type", diag=TRUE, main="Ct pair-wise correlation of TBP Normalised
data")
dev.off()

pdf(file="Ct pair-wise correlation of quantile normalised data.pdf", height=10, width=10)
plotCtPairs(q.norm, col="type", diag=TRUE, main="Ct pair-wise correlation of Quantile Normalised
data")
dev.off()

pdf(file="Ct pair-wise correlation of 4 gene normalised data.pdf", height=10, width=10)
plotCtPairs(dall.norm, col="type", diag=TRUE, main="Ct pair-wise correlation of 4 gene Normalised
data")
dev.off()

### Plot Ct Heatmaps

pdf(file="Ct Heatmap of raw data.pdf")
plotCtHeatmap(raw, gene.names="", dist="euclidean", main="Ct Heatmap of raw data")
dev.off()

pdf(file="Ct Heatmap of 18s normalised data.pdf")
plotCtHeatmap(d18s.norm, gene.names="", dist="euclidean", main="Ct Heatmap of 18s Normalised
data")
dev.off()

pdf(file="Ct Heatmap of GAPDH normalised data.pdf")
plotCtHeatmap(dGAPDH.norm, gene.names="", dist="euclidean", main="Ct Heatmap of GAPDH
Normalised data")
dev.off()

pdf(file="Ct Heatmap of GPI normalised data.pdf")
plotCtHeatmap(dGPI.norm, gene.names="", dist="euclidean", main="Ct Heatmap of GPI Normalised
data")
dev.off()

pdf(file="Ct Heatmap of TBP normalised data.pdf")
plotCtHeatmap(dTBP.norm, gene.names="", dist="euclidean", main="Ct Heatmap of TBP Normalised
data")
dev.off()

pdf(file="Ct Heatmap of quantile normalised data.pdf")
plotCtHeatmap(q.norm, gene.names="", dist="euclidean", main="Ct Heatmap of Quantile Normalised
data")
dev.off()

pdf(file="Ct Heatmap of 4 gene normalised data.pdf")
plotCtHeatmap(dall.norm, gene.names="", dist="euclidean", main="Ct Heatmap of 4 gene Normalised
data")
dev.off()
```

```
### Plot PCAs

pdf(file="PCA of raw data.pdf")
plotCtPCA(raw, features=FALSE)
dev.off()

pdf(file="PCA of 18s normalised data.pdf")
plotCtPCA(d18s.norm, features=FALSE)
dev.off()

pdf(file="PCA of GAPDH normalised data.pdf")
plotCtPCA(dGAPDH.norm, features=FALSE)
dev.off()

pdf(file="PCA of GPI normalised data.pdf")
plotCtPCA(dGPI.norm, features=FALSE)
dev.off()

pdf(file="PCA of TBP normalised data.pdf")
plotCtPCA(dTBP.norm, features=FALSE)
dev.off()

pdf(file="PCA of quantile normalised data.pdf")
plotCtPCA(q.norm, features=FALSE)
dev.off()

pdf(file="PCA of quantile 4 gene data.pdf")
plotCtPCA(dall.norm, features=FALSE)
dev.off()


### Limma for different raw and normalisation methods.  Tables and plots will need to be appended
with normalisation method


library(limma)

# Preparing experiment design raw

xnorm <- dall.norm

design <- model.matrix(~0+xnorm$Treatment)
colnames(design) <- c("DMSO", "SF", "SFTofa", "Tofa")
print(design)
contrasts <- makeContrasts(SFTofa-SF, SF-DMSO, Tofa-DMSO, levels=design)
colnames(contrasts) <- c("SFTofa-SF", "SF-DMSO", "Tofa-DMSO")
print(contrasts)
# Reorder data to get the genes in consecutive rows
xnorm2    <- xnorm[order(featureNames(xnorm)),]
qDE.limma.xnorm       <- limmaCtData(xnorm2, design=design, contrasts=contrasts,
ndups=1, adjust.method="none")

write.table(qDE.limma.xnorm[["SFTofa-SF"]], file="SFTofa-SF dall.norm.txt", sep="\t", col.names=NA)
write.table(qDE.limma.xnorm[["SF-DMSO"]], file="SF-DMSO dall.norm.txt", sep="\t", col.names=NA)
#write.table(qDE.limma.xnorm[["Tofa-DMSO"]], file="Tofa-DMSO dall.norm.txt", sep="\t",
col.names=NA)

pdf(file="SFTofa vs SF significant genes - dall.norm.pdf")
plotCtRQ(qDE.limma.xnorm, comparison=1, p.val=0.05, transform="log2", col="#9E0142", legend=F)
dev.off()

pdf(file="SF vs DMSO significant genes - dall.norm.pdf")
```

```
plotCtRQ(qDE.limma.xnorm, comparison=2, p.val=0.05, transform="log2", col="#9E0142", legend=F)
dev.off()


#plotCtRQ(qDE.limma.xnorm, comparison=3, p.val=0.05, transform="log2", col="#9E0142")


pdf(file="SFTofa vs SF significant genes with Ct - dall.norm.pdf")
plotCtSignificance(qDE.limma.xnorm, q=xnorm, comparison=1,
groups=xnorm$Treatment, target="SFTofa",
calibrator="SF", jitter=0.1, p.val=0.05)
dev.off()

pdf(file="SF vs DMSO significant genes with Ct - dall.norm.pdf")
plotCtSignificance(qDE.limma.xnorm, q=xnorm, comparison=2,
groups=xnorm$Treatment, target="SF",
calibrator="DMSO", jitter=0.1, p.val=0.05)
dev.off()
```

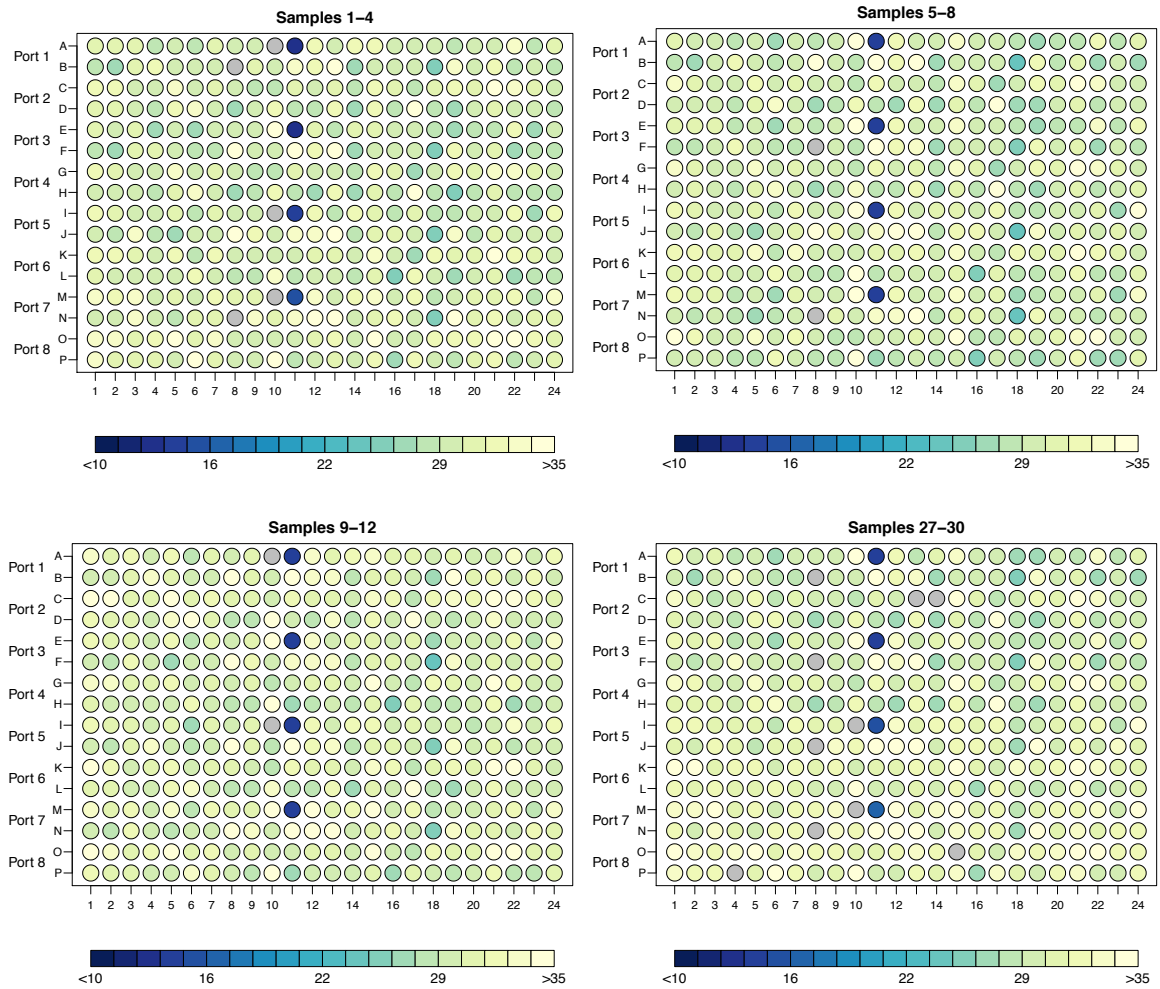


Figure 0-4 Visual inspection of rawCt values does not reveal any obvious spatial batch effects. Plots were generated using the HTqPCR package following qRT-PCR of samples from CD4 T-cells treated with synovial fluid and tofacitinib. Raw Ct values are colour coded by scale and blue dots correspond to 18S transcript showing that in each sample, 18S transcript is expressed at a similar level.

Sample	Treatment	Patient
Sample1	DMSO	D1
Sample2	Tofa	D1
Sample3	SF	D1
Sample4	SFTofa	D1
Sample5	DMSO	D2
Sample6	Tofa	D2
Sample7	SF	D2
Sample8	SFTofa	D2
Sample9	DMSO	D3
Sample10	SF	D3
Sample11	Tofa	D3
Sample12	SFTofa	D3
Sample13	DMSO	D4
Sample14	Tofa	D4
Sample15	SF	D4
Sample16	SFTofa	D4

Figure 0-5 Sample information for TLDA qPCR validation experiment.