Surveying the Epigenetic Landscape of Murine Synovitis to Explore the Heterogeneity of Rheumatoid Arthritis

A thesis submitted to Cardiff University in accordance with the requirements for the degree of Doctor of Philosophy in the School of Medicine

By

Stuart Taliesin Owain Hughes BSc (Hons) MSc

December 2023

Acknowledgements

While undertaking the research project of my master's degree I swore to myself I would never do a PhD. I am immensely glad that I changed my mind. As I come to the very final stages of compiling this thesis, I look back over the last four years and just how monumental they have been I wish to thank those whose support has been invaluable.

The mentorship and patience of Professor Simon Jones has made this work possible. I am exceptionally grateful for the opportunity and guidance given to me.

I would like to thank Professors Stephen Turner and Tony Tiganis of Monash University for their support from afar. Additionally, I would like to thank Sir Stanley Thomas, the Systems Immunity Research Institute, the Monash Biomedical Discovery Institute and Versus Arthritis for their generous support of the project.

Likewise, I am grateful for teaching and assistance of doctors Anna Cardus Figueras, Robert Jenkins, James Burston, Robert Andrews, and Barbara Szomolay. When I first began work in the lab, I felt lost and had very little idea what I was supposed to be doing. But Anna held my hand through the first ChIP and ATAC experiments and somehow at the end of three years I felt vaguely confident. When Anna, Rob Jenkins and James departed from the lab they took with them their wisdom and expertise. That I was able to complete experiments after they had left, I believe stands as a testament to knowledge and skills they had imparted and their willingness to explain. In the year or so since I have left the lab the thing I have missed above all else is the people.

My Mother's own doctoral journey partly served as a lesson in what not to do. Doing a PhD while she continued to work, looked after the family, and everything else was certainly not the easiest. However, her commitment showed that everything is possible as long as one never gives in and keeps smiling. My Father has been my most steadfast supporter, and when things were at their most trying reminded me of what will be to come when, not if, the dam thing is finished. I thank you both.

I would also like to thank my friends of Swansea Medical School and Swansea Medicals RFC for putting up with my incessant complaining and keeping me sane.

Finally, Daniela, by far and away the most important part of this PhD. Shared animal models certainly have an awful lot to answer for. Thank you for all you help and unwavering support over the last few years. I love you.

Summary

Rheumatoid arthritis, an immune-mediated inflammatory disease, attacks joints, causing cartilage, bone erosion and chronic pain. Despite significant advances treatment options, many patients do not respond to commonly prescribed targeted therapies, necessitating improvements in their stratification. Ultrasound-directed biopsies of inflamed joints show the clinical presentation of synovitis is highly heterogeneous, with histological features identifying fibroblast-rich, myeloid-rich, and lymphoid-rich synovitis. The focus of my thesis is to understand the mechanisms driving these alternate patterns of disease. To explore how these forms of synovitis evolve, I applied analytical methods to survey the epigenetic landscape of inflamed joint tissues from mice with antigen-induced arthritis (AIA). Experiments were conducted in wild-type (Wt) mice and mice deficient in the α-receptor subunit for either interleukin-6 (*Il6ra*$^{-/-}$) or interleukin-27 (*Il27ra*$^{-/-}$), which develop hallmarks of synovitis resembling myeloid-rich, fibroblast-rich, and lymphoid-rich synovitis respectively. Next-generation sequencing methods assessing chromatin accessibility (ATAC-seq) and transcription factor (ChIP-seq) involvement in synovitis were generated and mapped against synovial RNA-seq datasets previously generated from these mice. Considering the biology of IL-6 and IL-27, my analysis focussed on the role of the Janus-activated kinase-Signal Transducer and Activator of Transcription (Jak-STAT) pathway and the STAT1 and STAT3 transcription factors. Analysis of synovial tissues from mice with AIA identified elements of gene regulation common to all three strains with AIA and others unique to Wt, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice. In this regard, ATAC-seq revealed subtle differences in the epigenetic control of gene expression specific to each synovial pathotype. Building on these findings, I next evaluated the role of STAT1 and STAT3, which share a complex regulatory interplay affecting alternate patterns of gene regulation. My results confirmed the importance of this mechanism in determining synovitis in mice with AIA, with STAT1 and STAT3 effecting processes including leukocyte recruitment and activation, pannus formation and joint damage. Here, molecular pathway analysis identified genomic signatures linked with disease heterogeneity in synovitis, highlighting the role of chromatin accessibility and transcription factor activity in shaping the course of disease. My results open opportunities to consider the pathways driving arthritis progression and clinical responses to biological medicines commonly used in treating rheumatoid arthritis.

List of publications and presentations

Jenkins RH, **Hughes STO**, Figueras AC, Jones SA. Unravelling the broader complexity of IL-6 involvement in health and disease. Cytokine. 2021 Dec;148:155684. doi: 10.1016/j.cyto.2021.155684. Epub 2021 Aug 16. PMID: 34411990.

Millrine D, Jenkins RH, **Hughes STO**, Jones SA. Making sense of IL-6 signalling cues in pathophysiology. FEBS Lett. 2022 Mar;596(5):567-588. doi: 10.1002/1873-3468.14201. Epub 2021 Oct 15. PMID: 34618359; PMCID: PMC9673051.

Hughes STO, Cardus Figueras A, Andrews R, Jenkins RH, Hill D, Costa D, Burston J, Jones G, Turner S, Tiganis T, Jones S. *Next Generation Sequencing Methods Open New Insights into Disease Heterogeneity in Synovitis*. 9th annual meeting of the International Cytokine and Interferon Society. Cardiff. 2021 (POSTER)

Abbreviations

ACPAs	Anti-citrullinated protein antibodies
ACR	American College of Rheumatology
ADCC	Antigen-dependent cellular cytotoxicity
AIA	Antigen Induced Arthritis
APCs	Antigen-presenting cells
ATAC-seq	Assay for Transposase Accessible Chromatin using sequencing
BAM	Binary alignment map
BEDPE	Browser extensible data paired-end
ChIP-seq	Chromatin Immunoprecipitation sequencing
CIA	Collagen Induced Arthritis
CFA	Complete Freund's adjuvant
CLB	Cell Lysis Buffer
CRP	C-reactive protein
CVD	Cardiovascular disease
DAS28	Disease Activity Score
DHFR	Dihydrofolate reductase
DMARDs	Disease-modifying antirheumatic drugs
DNA	Deoxyribonucleic acid
DNase-seq	DNase I hypersensitive site sequencing
EGFR	Epidermal growth factor receptor
ELMO1	Engulfment and Cell Motility 1
ELS	ectopic lymphoid-like structure
ERP	End Repair Mix
ESR	Erythrocyte sedimentation rate
EULAR	European League Against Rheumatism
FAIRE-seq	Formaldehyde-Assisted Isolation of Regulatory Elements and sequencing
FAPα^+	Fibroblast activator protein-α
GAP	Gene associated peak
GM-CSF	Granulocyte-macrophage colony stimulating factor
gp130	Glycoprotein 130
GWAS	Genome wide association studies
HAT	Histone acetyltransferase
HLA	HLA-antigen
HLA-DR	HLA-antigen D related
IA	Intra-articular
IgG	Immunoglobulin G
IL	Interleukin
IL-27R	IL-27 receptor

IL-6R	IL-6 receptor
IMID	immune-mediated inflammatory disease
INF	Interferon
IP	Intraperitoneal
IPA	Ingenuity Pathway Analysis
IPDB	Immunoprecipitation dilution buffer
IPWB	Immunoprecipitation wash buffer
Irf1	Interferon Regulatory Factor 1
JAK	Janus Kinase
mAbs	Monoclonal antibodies
MACS	Magnetic-activated cell sorting
MACS2	Model-based Analysis of ChIP-seq algorithm
MAPK	Mitogen activated protein kinase
mBSA	methylated bovine serum albumin
MCP	Metacarpophalangeal
mBSA	Methylated bovine serum albumin
MHC	Major histocompatibility complex
MMPs	Matrix metalloproteinases
MNase-seq	micrococcal nuclease digestion with sequencing
MTP	Metatarsophalangeal
NA	Non-associated
NF-κB	Nuclear factor κ-light-chain-enhancer of activated B cells
NGS	Next generation sequencing
NICE	National Institute of Heath and Care Excellence
NSAIDs	Non-steroidal anti-inflammatory drugs
OPG	Osteoprotegerin
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PIP	Proximal interphalangeal
PMM	PCR Master Mix
PPC	PCR Primer Cocktail
pY-STAT	Tyrosine phosphorylated STAT transcription factor
QCed	Quality controlled
RA	Rheumatoid Arthritis
RANKL	Receptor Activator of Nuclear factor κ-B Ligand
RF	Rheumatoid factor
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RSB	Resuspension Buffer
SC	Subcutaneous
SPF	Specific-pathogen-free

SOCS	Suppressor of cytokine signaling
STAT	Signal Transducer and Activator of Transcription
STL	Stop Ligation Buffer
TB	Tuberculosis
TE	Tris-EDTA
Th	CD4 positive Helper T-cells
TNFα	Tumour necrosis factor-α
tsDMARDs	Target synthetic DMARDs
TSS	Transcription start sites
Un	Unassigned
VTE	Venous thromboembolism
WHO	World Health organisation
Wt	Wild type C57/Bl6 mice

Table of Contents

List of Figures

Chapter 1: Introduction

Chapter 2: Methods and Materials

Chapter 3: Tracking chromatin accessibility as a response to synovitis in murine antigen-induced arthritis.

Chapter 4: Relating chromatin accessibility to the transcriptional control of genes involved in synovial pathology

List of Tables

Chapter 1:

Introduction

1.1 Rheumatoid Arthritis

Rheumatoid arthritis (RA) is an immune-mediated inflammatory disease (IMID) that drives a severe pathology of synovial joints, causing swelling and pain. Here, inflammation of the synovium, known as synovitis, contributes to the deterioration of bone and cartilage in the affected joints. Although initial symptoms are relatively mild, this debilitating autoimmune disease rapidly leads to irreparable disability (Smolen et al. 2018). It is a very common autoimmune condition, affecting ~1% of the global adult population, with an increased prevalence in certain ethnicities and geographical areas (Tobón et al. 2010). In particular, RA shows a two-threefold increase in disease incidence in women over men (Crowson et al. 2011) and factors contributing to disease risk include genetic susceptibilities and environmental factors including cigarette smoking.

RA is a highly complex disease and consistent with the clinical presentation of other IMIDs, patients with RA display various comorbidities that impact their treatment management and quality of life. These include increases in cardiovascular risk, alterations in iron, glucose, and lipid metabolism, and processes affecting fatigue, sleep, and mental health (McInnes and Schett 2011). Thus, RA is a systemic chronic inflammatory disease which affects various tissues and organs beyond joint pathology.

1.1.1 Inflammation, Autoimmunity, and the Transition to Chronicity

Inflammation is a complex, tightly regulated immune process essential for the maintenance of normal physiology and anti-microbial host defence. Activated as a response to infection, trauma or injury, inflammation coordinates the activation of cell types that facilitate communication between stromal tissue cells and immune cells involved in shaping innate and adaptive immunity. An appropriate regulation of the inflammatory process ensures protection from pathogens (e.g., bacteria, viruses, parasites) and promotes tissue repair and a return to physiology. However, a recurrent or persistent activation of the inflammatory cascade can trigger septic shock-type syndromes, inflammation-induced tissue injury (e.g., fibrosis) and the development of autoimmune reactions linked with IMIDs. These range from immediate life-threatening conditions to chronic debilitating diseases that increase the risk of morbidity and mortality (Furman et al. 2019).

It is important to distinguish between acute and chronic inflammation. In response to infection or injury, an acute inflammatory episode provides essential host defence and ensures restoration of tissue function and integrity. Furthermore, acute inflammation is characterised by an initial engagement of innate immune cells, followed by a transition towards cell types supporting adaptive immunity, which are then cleared from the tissue once the inflammatory challenge has resolved (Kotas and Medzhitov 2015). Any disruption to the control of these steps leads to the retention of activated immune cells within the tissue, which contributes to a loss of tissue architecture and the development of pathophysiology (Kotas and Medzhitov 2015; Fullerton and Gilroy 2016; Furman et al. 2019). In RA, this occurs after the concurrence of multiple risk factors reach a threshold at which the disease is triggered (Smolen et al. 2018). However, the exact mechanism of this change is not fully understood. Localisation of inflammation within the synovium could be caused by a result of local complement activation, microvascular damage or joint microtrauma. Alternately joint damage could be initiated by circulating autoantibodies recognising immune complexes causing activation of periarticular osteoclasts and the release of inflammatory cytokines (McInnes and Schett 2017).

1.1.2 Onset and Disease Course

The progression of RA can be divided into preclinical, early, and established stages (Raza and Gerlag 2014; Smolen et al. 2018). During the preclinical phase, which can extend over years, there are no outward signs or symptoms of the disease. Occasionally, however, an abrupt onset may occur due to immune disturbances (Masi 1983). Preclinical RA is characterized by asymptomatic autoimmunity, with the identification of autoantibodies targeting citrullinated proteins (termed anti-citrullinated protein antibodies; ACPAs), and the presence of rheumatoid factor (RF), which recognizes the Fc region of Immunoglobulin G (IgG) (Holers 2013).

These autoantibodies are critical drivers of IMIDs, including RA. However, not all patients possess autoantibodies and RA patients are commonly classified by the presence or absence of autoreactive antibodies (Malmström et al. 2016). These patient groups often show differences in genetic risk factors, the rate of disease progression and treatment response (Martin-Mola et al. 2016). A lack of RF or ACPAs in patient blood serum can hinder

diagnosis and treatment (Coffey et al. 2019). However, ‘seronegative’ patents may become reclassified as further autoantibodies are identified as determinates of RA (Sokolova et al. 2021).

The presence of autoantibodies during the preclinical stages of RA suggests that some immune challenge is required to trigger their pathogenic properties.
While the drivers of this process are complex and not entirely understood, several mechanisms have been proposed. For example, in the presence of asymptomatic autoimmunity, disease onset is driven by a switch in autoantibody activity caused by a pathway comprising interleukin (IL)-23 and a subset of helper T-cells which produce IL-17 (Th17) (Pfeifle et al. 2016).

When the level of autoimmunity breaches tolerance, becoming clinically evident, antigen-presenting cells (APCs) stimulate the production of the aforementioned antibodies, activating other immune cells. Early RA involves synovitis with activated macrophages, infiltrating effector $CD4^+$ T-cells, and synovial tissue cells, leading to inflammation through pro-inflammatory cytokines, bioactive lipids, and chemokines. This cascade recruits and sustains infiltrating leukocytes within the inflamed joint. Synovial fibroblasts, producing matrix metalloproteinases (MMPs), degrade the extracellular matrix composed of collagen, aggrecan, elastin, fibronectin, gelatine, and laminin. This degradation arises from altered gene transcription due to changes in chromatin accessibility and transcription factor binding that disrupt tissue homeostasis (Araki and Mimura 2017). IL-23 and its associated mechanism is one pathway by which this can occur. As a proinflammatory cytokine, and member of the IL-12 family, secreted by macrophages and dendritic cells, IL-23 promotes and maintains the differentiation of T cells into Th17 cells. These cells then secrete other inflammatory cytokines including IL-17A which contribute to synovial inflammation and the regulation of processes affecting osteoclastogenesis (Najm and McInnes 2021).

As the disease advances, autoantibody profiles expand due to increased production and antigen specificity. Untreated or unresolved inflammation leads to established RA. Here, persistent inflammatory cytokines and adaptive immune responses sustain the disease, causing cartilage and bone damage, resulting in sensory and neuropathic pain, and classic deformities such as the 'swan-neck deformity' in fingers. These changes significantly decrease patients' quality of life (Smolen et al. 2018).

1.1.3 The composition of the synovium under homeostasis and the inflammatory state

Though a systemic disease, with a range of immunological events across the body, joint synovitis is a core component of RA. In a healthy joint, the synovium has two principal roles. First, to produce lubricants for the cartilage, creating a low-friction environment for ease of movement. Second, it provides essential nutrients for the maintenance of a healthy cartilage, which lacks its own blood supply. The intima of the synovium lacks a basement membrane and does not behave like a barrier seen in other tissues and organs (e.g., endothelial, or epithelial surfaces). Instead, it is comprised of macrophage-like and fibroblast-like synoviocytes with an additional sub-lining made up of blood vessels, adipose tissue, fibroblasts, and a variety of other immune cells. This loose configuration of cells is leaky and allows for relatively free movement of cells and proteins in and out of the synovium and synovial fluid (Castor 1960).

In RA there are two principal pathogenic alterations which occur in the synovium. The first; the macrophage-like and fibroblast-like synoviocytes increase in number and become activated, leading to an expansion of the intimal lining. Synoviocytes are also major producers of inflammatory cytokines and degradative proteases, which shape the development of synovitis and associated joint damage (McInnes and Schett 2011). Subsets of fibrocytes, fibroblast activator protein-α (FAPα^+) thymocyte differentiation antigen THY1$^+$ and FAPα^+THY1$^-$, within the synovium can, however, perform different functions and thus contribute to different disease outcomes(Croft et al. 2019)

The FAPα^+ THY1$^+$ cell population contribute to a persistent, high-inflammatory disease course, though they have little effect on bone or cartilage damage. The FAPα^+THY1$^-$ population behaves in a converse manner. The disease presents with a lower inflammatory state with, however, more severe structural damage. The differential production by these cell subtypes of matrix metalloproteases (MMPs) or inflammatory cytokines produces these different outcomes in patients favouring each subtype (Croft et al. 2019).

The second pathogenic alteration involves the infiltration by leukocytes of the synovial sub-lining to drive adaptive immunity. In the majority of patients, it is CD4$^+$ T-cells that infiltrate the synovium. These cells can harbour multiple functions and fall into a variety

of cell subtypes contributing greatly to the variety of phenotypic variation in RA. Regardless of their presence in the synovium, the cytokines they release and the other cells they recruit drive the inflammation of the joint (Chemin et al. 2019).

In a minority of patients (20-40%), these cells form ectopic lymphoid-like structure (ELS), comprising simply T-cell and B-cell aggregates or highly organised lymphocyte networks containing follicular dendritic cells and supporting stromal cells resembling germinal centres in lymph nodes. It is thought that their formation is predetermined, by a combination of genetic and environmental factors, at the very early stages of disease rather than evolving over time (Pitzalis et al. 2014). These structures drive mature B cell proliferation, and differentiation into autoreactive antibody-generating populations that exacerbate local synovitis and joint pathology (Randen et al. 1992; Humby et al. 2009).

Synovitis contributes to joint damage through the disruption of regulatory pathways responsible for the turnover of cartilage and bone damage. Matrix-degrading enzymes and cytokines are released by macrophages, neutrophils, and mast cells. Fibroblast-like synovocytes also produce matrix-degrading enzymes (e.g., stromelysins and collagenases). These enzymes are the principal causes of damage to the joint's cartilage (Kiener et al. 2009). Bone erosion is caused by the overactivation of osteoclastogenesis resulting in osteoclast maturation and activation (Redlich and Smolen 2012). Osteoblasts also play a role in joint damage via their dysregulation which results from an unresolved inflammatory state (Berardi et al. 2021).

Cytokines (e.g., IL-1β, IL-17, tumour necrosis factor; TNFα) that activate the nuclear factor κ-B pathway are major determinants of bone and cartilage destruction (Jung et al. 2014). They are involved with the regulation of metalloproteases by chondrocytes and signals relayed by the Receptor Activator of Nuclear factor κ-B Ligand (RANKL) and the Wnt pathway. (Ostrowska et al. 2018). The RANKL pathway is also involved in RA's joint pathology, particularly bone destruction. As the principal regulator of osteoclast activity. In RA there is an imbalance between RANKL and its inhibitor osteoprotegerin (OPG), thus RANKL is responsible for the hallmark bone reabsorption of RA (Papadaki et al. 2019).

Cytokine signalling networks which can influence the course of inflammation as well as the cellular activities which result in the tissue damage described are a key part of RA. As alluded to previously, cytokine signalling is also heterogeneous as there is a wide range of signalling cascades which impact the disease course (Smolen et al. 2018).

1.2 Risk factors for RA

Various genetic and environmental risk factors affect clinical susceptibility to RA. However, it is thought that it is a combination of cumulative factors which leads to the initiation of autoimmunity and associated systemic inflammation, these are described below.

1.2.1 Sex

RA is more prevalent in women, with the cumulative lifetime being 3.6% risk in women, while in men is only 1.7% (Crowson et al. 2011). It is proposed that this increase in disease susceptibility is, in part, due to the effects of oestrogen and other sex hormones, such as androgens and progesterone, on the immune system. For instance, oestrogen's capacity to increase self-recognising T and B-cell survival (Alpízar-Rodríguez et al. 2017). However, as is the case with much of the biology surrounding RA, this aspect of the disease is uncertain, complicated, and not well understood. There are two principal theories as to this gender imbalance. The first is there are sex-linked genetic components which contribute to RA. The second is that female sex hormones (oestrogen/ lower testosterone levels) play a key role in autoimmune diseases (Gerosa et al. 2008). Interestingly, although RA onset typically takes place between the ages of 40-50, after most women would have had children, some women with RA experience an improvement, even full remission, in their RA symptoms during or following pregnancy (Krause and Makol 2016).

1.2.2 Environmental Factors

One of the foremost environmental risks for RA is smoking. However, it should be noted that it is believed that it is not nicotine or tobacco itself but rather the other components of cigarette smoke which increase the risk of RA given that an increase in RA incidence is not noted in those who use tobacco-based snuff products (Jiang et al. 2014). A meta-analysis of cigarette smoking in RA shows that the risk of disease increases with the number of cigarettes smoked a day and the length of time an individual has continued smoking (Sugiyama et al. 2010). There exists a shared epitope; an amino acid motif which is often encoded by some alleles of the human leukocyte antigen D-related (HLA -DR) locus which shows an association with RA (Holoshitz 2010). The link between RA and smoking is

strongest in patients with at least one copy of this epitope. The risk of developing RA can increase 20-fold in smokers with the epitope compared to non-smokers who do not possess it (Källberg et al. 2011). This relationship highlights the importance of a combination between environmental and genetic factors in the onset of RA. Similar to smoking, dust inhalation, especially silica dust, has also been associated with RA (Stolt et al. 2005).

The microbiota has also been implicated in the development of RA. Patients with RA commonly show less diversity in gut microbiota than the general population and often display reduced Actinobacteria species (Chen et al. 2016). The post-translational modification of proteins described in **Section 1.1.1** typically takes place in the mucus membranes of the mouth, lungs, and gut. It is to the cells in these locations where cigarette smoke, dust and microbiota are exposed and where their affects begin to manifest (Smolen et al. 2018).

The most common cause of premature death in those with RA is cardiovascular disease (CVD), and there is a high prevalence of CVD risk factors in RA patients. These include obesity, diabetes mellitus, hypertension, and hyperlipidaemia (Radner et al. 2017). Given this association, it is somewhat unsurprising that there is an association between a diet which includes a high consumption of sugar and fat yields a higher incidence of RA (Skoczyńska and Swierkot 2018).

1.2.3 Genetic Factors

Genetics plays a large part in RA's onset and disease is frequently seen in different generations from the same family (Frisell et al. 2016). Genome-wide association studies (GWAS) have revealed around 100 loci across the genome with links to RA susceptibility. While several show relatively weak associations with disease onset, the combination of genetic predisposition and environmental factors represent major drivers of RA (Smolen et al. 2018). Furthermore, these genes and the proteins for which they encode provide avenues of exploration for possible therapeutic intervention and drug discovery (Okada et al. 2014)

Perhaps the most prominent genetic factor in RA is the shared epitope mentioned in **section 1.2.2**. Loci which encode for the class II human leukocyte antigen (*HLA*), which is also known as the major histocompatibility complex (MHC), molecules possibly contain the

shared epitope. The most common alleles of the aforementioned HLA locus which are most prominently linked to the development of RA are *HLA-DRB1*01* and *HLA-DRB1*04* (Gregersen et al. 1987). A not altogether uncommon problem with GWAS and similar exploration of the genetic component of RA is pleiotropy. Many genes identified as playing a key role in the pathology of RA are not unique to the disease but also play a role in other similar autoimmune and/or inflammatory conditions. Thus, when researching a particular disease of this nature it becomes necessary to remove these common genes from the analysis with the aim of identifying genes unique to the disease of interest (Matthews et al. 2009).

Although the majority of RA genetic studies seek to uncover variants which mark susceptibility it is also important to elucidate how genetic polymorphisms impact biological processes affecting arthritis severity, rate of disease progression, or response to therapy. The genes *DKK1* (dickkopf-1; involved in adult bone regulation), *GRZB* (granzyme-B; which processes cytokines and is involved in chronic inflammation), *HLA-DRB1* (a key player in human immune function), *IL2RA* (also involved in the immune response), *MMP9* (matrix metalloproteinase-9; a type-IV collagenase) and *SPAG16* (which, though principally involved with sperm motility has been implicated in increasing RA susceptibility) are examples which have been associated with both susceptibility and severity to and of RA, though a few genes; *FOX03* for instance, has been associated with only severity (Lee et al. 2013; Krabben et al. 2015; Viatte et al. 2016).

As mentioned above, genetic loci associated with RA are frequently associated with other IMIDs (Matthews et al. 2009). Thus, efforts are taken to discover risk loci which are unique to RA. SNPs associated with one such gene is *IL6ST*, which encodes the IL6 signal transducer glycoprotein 130 (gp130)(Stahl et al. 2010). The importance of the IL-6 cytokine to the work of this thesis is discussed further below.

Another gene of interest is *PTPN22*, which encodes a protein tyrosine phosphatase which is a key mediator of immune homeostasis. This is achieved by the inhibition of T-cell receptor signalling and the selective promotion of type I interferon (INF) responses. Furthermore, a SNP, *PTPN22* 1858C>T is a risk factor for various IMIDs including RA (Stanford and Bottini 2014). Additionally, PTPN22 and a similar molecule PTPN2 regulate the phosphorylation of the transcription factor Signal Transducer and Activator of Transcription (STAT) 1 downstream of IL-6 signalling, a key pathway in RA's pathophysiology. PTPN2 has

also been linked with the disease properties ELS (discussed earlier) and synovial infiltrating monocytic cells (Svensson et al. 2019; Twohig et al. 2019).

Risk loci which are shared between conditions can suggest a shared aetiology for those conditions. STAT4 orchestrates immune responses by mediating signals from IL-12 and IL-23. Its actions include promoting the differentiation of proinflammatory Th1 cells, enhancing immune defence against pathogens and influencing the production of autoantibodies. A SNP haplotype in the gene *STAT4* is linked with RA and systemic lupus erythematosus, thus it is probable the two conditions share disease pathways(Remmers et al. 2007).

As yet there are no genes which have been identified which can predict the success of a treatment. However, analysis of genetic loci which are linked to RA risk, in conjunction with drugs which have been approved for use in RA (or other diseases such as cancer) enables specific loci and their biological impact to be linked with an existing drug (Okada et al. 2013). Nevertheless, this lack of evidence and the relatively few genes (compared to susceptibility genes) which have been robustly linked with severity is in part due to the small sample size (RA patients) as compared with the healthy population (Oliver et al. 2015).

1.2.4 Epigenetic Factors

Both histone acetylation and DNA methylation are epigenetic mechanisms which have been implicated in the development of RA. Histone acetylation is a process in which the DNA wrapped around histone proteins becomes loose. Acetyl groups are added to the lysine resides of the histone proteins in a reaction catalysed by histone acetyltransferases (HATs). This reduces the positive charge of lysine reducing its affinity for DNA which is negatively charged. This looseness enables DNA to be more easily accessed for transcription and expression (Gujral et al. 2020). DNA methylation involves the addition of a methyl group to the cytosine base of DNA. A high degree of DNA methylation can make the DNA less accessible and thus harder to be transcribed and expressed. However, this gene silencing can aid in genome stability as well as the suppression of inappropriate genes (Jin et al. 2011). DNA methylation in particular is involved in the degree of risk conferred in various genetic variations as well as a mechanism by which environmental factors influence disease. Within the HLA region there have been identified clusters, the methylation status of which

has been linked with risk for RA (Liu et al. 2013). Both hypermethylation and hypomethylation of different genes have been associated with RA (Zhao et al. 2022). Methylation levels have been found to be greater in smokers who have the *HLA-DRB1* risk allele versus those who do not. However, this difference is not observed in those who do not smoke (Meng et al. 2017). With regards to the severity of the disease, fibroblast-like synoviocytes in different joints, in the same patient, can exhibit a diversity of methylation thus, this is a possible explanation as to why some joints are more severely affected by RA than others (Frank-Bertoncelj et al. 2017).

Aside from protein-coding genes, many genetic/epigenetic risk variants map to enhancer regions involved in epigenetic gene regulation. Genes linked to these promoter regions commonly influence the identity and functional properties of stromal and haematopoietic cells (Trynka et al. 2013). Conversely, a given gene may be regulated by a number of different enhancer regions, thus different enhancer variants in individual patients may illicit different disease progression and severity though they act through the same gene (Martin et al. 2015).

1.3 Diagnosis and Assessment of RA

The early stage of RA, as soon as the disease becomes clinically perceptible, or better yet even earlier when there are detectable levels of RF and ACPAs but there are no symptoms, is often described as the window of opportunity for intervention. Developments in early diagnosis and the availability of biological medicines and small molecule inhibitors have meant that many of the most severe symptoms (e.g., joint deformity) are less common in patients with RA. However, without an accurate diagnosis, this is not possible. The mechanism underlying the window is not fully understood, though there are two primary suggestions. The synovium of early RA is different in terms of the cytokine and cellular environment to late/ established RA. This is a qualitative difference. Alternatively, the cells and cytokines in early and late RA may be the same though there are fewer in the early stage. This is a quantitative difference. Unfortunately, delays in clinical assessments and referrals to a rheumatologist frequently limit the benefits afforded by early intervention (Raza and Filer 2015).

There are key symptoms which may suggest RA over other forms of arthritis, these are positivity for self-reactive antibodies (established through serological assays), articular pain and swelling in the metatarsophalangeal (MTP) and or metacarphoangeal (MCP) joints, joint stiffness in the morning which lasts longer than 30 minutes and disease symmetry (joints being equally affected bilaterally) (Emery et al. 2002).

As important to devising treatment strategies as an initial diagnosis is measuring disease activity and subsequently a patient's response (or lack thereof) to treatment. With the absence of concrete disease biomarkers, a no less valid and reliable tool, which considers symptoms, patient feedback and serological assays is required. Several such scoring methods have been devised. An example is Disease Activity Score (DAS28) which marks disease on a continuous scale using the three aforementioned criteria. The 28 refers to the number of joints assessed. They are the shoulders, elbows, wrists, knees, MCP joints and the proximal interphalangeal (PIP) joints. In some circumstances, the MTP joints can also be included. A numeric value is calculated based on the number of swollen and tender joints (out of 28) in combination with the patient's erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) values and a general wellness score. Thus, designations of remission, low disease activity, moderate disease activity and high disease activity can be assigned (Fransen and van Riel 2009).

Another criterion is the American College of Rheumatology (ACR)/ European League Against Rheumatism (EULAR) classification criteria. This is based on the number of large and small joints involved, levels of RF and ACPA, ESR and CRP number and the duration of the arthritic symptoms. A key purpose of the ACR/EULAR is to emphasise the hallmarks of RA which emerge in the early disease stages, to enable clinicians to diagnose sooner and obtain better treatment outcomes (Kay and Upchurch 2012).

In addition to methods for diagnosing RA, tools exist for assessing disease severity and progression. The Sharp score (and its derivations) is a radiological tool to assess structural damage caused by inflammation. It measures bone erosion, cartilage destruction and joint space narrowing (Sharp et al. 1971).

The ultimate goal in treating RA is drug-free remission or remission of disease that prevents irreversible articular damage and long-term disability. The various treatments used in RA (see **Section 1.4**) are aimed at reducing inflammation. They are effective in many patients, but not all, and patients commonly display a lack of response to certain therapies

(Vital et al. 2010). A major challenge in the treatment of RA is identifying the right drug for the right patient. Thus, there is a need to predict the course of undifferentiated disease onset and the contributing signalling pathways that may be targeted for clinical benefit.

While the diagnostic tools discussed here, and genetic testing described earlier can be useful, they do not tell us which specific inflammatory pathway is driving disease. In essence, clinical criteria alone are not enough to predict disease outcomes and provide limited information on the most suitable treatment regime for an individual patient. This is a key area for development.

1.4 Treatment of Rheumatoid Arthritis

There is no outright cure for RA, though modern biological drugs offer opportunities to clinically control the progression of the disease, and the frequency of patients classified as in remission or drug-free remission is becoming more common (Van Den Broek et al. 2011). Many of these newer drugs operate through the manipulation of complex biological pathways. There are three principal types of biological drugs: cytokine blockers, lymphocyte-targeting agents, and small-molecule inhibitors. Owing to the complex nature and heterogeneity of RA finding an appropriate treatment plan for an individual patient remains challenging and patients often switch therapies to identify the most effective therapy or to reduce adverse outcomes. Indeed, in some cases the precise biological mechanism by which a drug acts is unknown as is why a drug may be effective in one patient and not another (Choy et al. 2013).

1.4.1 Disease-modifying anti-rheumatic drugs

Biological drugs are not the first treatment option offered by rheumatologists to patients. Instead, they are prescribed more conventional, broad-spectrum therapeutic strategies that target an overall reduction in inflammation. The principal goal of these early strategies is to restrict inflammation and reduce the symptoms of pain. These drugs include non-steroidal anti-inflammatory drugs (NSAIDs), corticosteroids, which are not suitable for long-term use, and various disease-modifying antirheumatic drugs (DMARDs) (Choy et al. 2013; Smolen et al. 2018).

Drugs classified as DMARDs often display different modes of action to the aforementioned steroids and NSAIDs. The most commonly prescribed DMARD is methotrexate, which is often combined with a short-term course of glucocorticoids to control the emerging inflammation (Onuora 2014). While methotrexate is highly effective in many patients it is not the only DMARD utilised. Others include leflunomide, hydroxychloroquine and sulfasalazine. A combination of these DMARDs can be used (methotrexate- sulfasalazine- hydroxychloroquine is a frequent 'triple therapy' combination) and is shown to be more effective than monotherapy (Verschueren et al. 2015).

Methotrexate is an immunosuppressant, used not only in the treatment of RA but in a host of other conditions such as cancer, Crohn's disease, and psoriasis. As a testament to the drug's success the World Health Organisation (WHO) consider it an essential medication due to its broad biological action (WHO EML 23rd List (2023). 2023). Developed in the middle of the last century, the therapeutic action of methotrexate is not entirely understood. Originally used in the treatment of cancer, methotrexate inhibits cell proliferation and prevents rapid immune cell division. The enzyme inhibited by methotrexate is dihydrofolate reductase (DHFR). DHFR is required for the conversion of dihydrofolate to the biologically active form of tetrahydrofolate. For this reason, folic acid supplementation is frequently given with methotrexate to offset many of the side effects of the drug's hepatotoxicity and gastrointestinal issues (Liu et al. 2019). Without the latter molecule the building blocks of Deoxyribonucleic acid (DNA) and Ribonucleic acid (RNA) nucleotides; purines and pyrimidines, cannot be synthesised. Without DNA synthesis cells cannot replicate. By inhibiting the proliferation of immune cells, their disease-causing actions, for example, the release of pro-inflammatory cytokines, are limited (Bedoui et al. 2019). Furthermore, methotrexate has been found to inhibit the Janus Kinase (JAK) Signal Transducer and Activator of Transcription (STAT) pathway (which is discussed in greater detail below), a crucial pathway to inflammation and immunity (Alqarni and Zeidler 2020).

While DMARDs, including methotrexate, often control or slow the rate of disease progression, patients frequently show a lack of long-term clinical efficacy and drug compliance due to side effects. These symptoms include gastrointestinal complications (e.g., nausea, vomiting, mucosal ulcers), loss of appetite, hair loss and swelling of limb and facial features (Benjamin et al. 2023). Despite these contraindications, methotrexate and other DMARDs are commonly combined with more specific biological drugs to enhance clinical

efficacy. Decisions to apply combination therapies rely on a longitudinal assessment of autoantibody levels and radiographic changes in joint damage (Smolen et al. 2006; Smolen et al. 2017). For patients unable to tolerate the side effects of methotrexate, recent clinical trials reveal that some biological drugs offer better clinical outcomes when applied as a monotherapy for patients unable to tolerate methotrexate (Gabay et al. 2013).

1.4.2 Biological Agents

Biological drugs (also termed biologics or biological medicines) are now an established part of RA therapy and have revolutionised the standard of care for patients with IMIDs. These therapies are grouped into three classes: cytokine blockers, lymphocyte-targeting agents, and small-molecule inhibitors (Choy et al. 2013). While their mode of action targets inflammatory or immune pathways, these interventions often contribute to improvements in the wider features of the disease, including fatigue, sleep, metabolism, and mental well-being.

As with much of RA, these are not always understood to the full. Arguably, their numerosity can be a double-edged sword as rheumatologists can be at a loss as to which to prescribe, a problem exacerbated by the near constant emergence and approval of new biologics, furtherer muddying the waters, though National Institute of Health and Care Excellence (NICE) guidance does provide some clarity (NICE 2021). Symptoms observed by patients themselves and by clinicians can establish the severity of the disease and the effectiveness of treatment, but unfortunately, they cannot elucidate the precise biological pathways causing disease. Similarly, levels of RF and ACPAs give no information nor do the various risk alleles previously described. Biomarkers which could inform as to the nature of the inflammatory pathway in an individual patient would be of great assistance in the prescription of biologics and would circumvent any trial-and-error approach when devising treatment regimens (Choy et al. 2013).

1.4.2.1 Cytokine-targeting Agents

Cytokines are core to the pathogenesis of RA. They mediate inflammation, the autoimmune response and joint destruction. Cytokines like IL-1, IL-6, and Tumour Necrosis Factor (TNF)α can promote the recruitment and activation of immune cells to sites of disease. IL-1 and

TNFα also promote the production of MMPs which damage the cartilage and bone of rheumatic joints. In addition to cell recruitment, cytokines control the phenotype of T-cells and the activation and function of B-cells within the synovium (McInnes and Schett 2007).

Local modulation of synovitis by various cytokines including granulocyte-macrophage colony-stimulating factor (GM-CSF), IL-12, IL-15, IL-17, and IL-18 is of immense importance to the pathology of RA. Several cytokines display hormone-like properties (e.g., IL-6, TNFα) and elicit functions that affect metabolism (e.g., iron, glucose, lipid metabolism), the neuroendocrine system (e.g., the Hypothalamic-Pituitary-Adrenal axis) and physiological processes affecting sleep, fatigue, and mental health (Jenkins et al. 2021). These processes are often dysregulated in IMIDs, and biological drugs often improve these symptoms and are part of their mode of action (Raison et al. 2006). Thus, the therapeutic benefits afforded by biological drugs typically extend beyond their control of joint pathology (McInnes and Schett 2007).

Biological drugs targeting cytokines fall into one of three groups based on their mechanism of action. Blocking agents target a specific cytokine or its receptor, thus preventing signalling. Another class is recombinant decoy antagonists which serve to 'soak up' a cytokine and diminish its signalling capability. Finally, there are inhibitory soluble cytokine receptors which function in a similar manner to the previous group. They are a form of the cytokine's natural receptor, modified to be soluble and thus negate a cytokine effect on cells (Rider et al. 2016).

Monoclonal antibodies (mAbs) can target specific, disease-related, molecules or cell groups. They have been used to treat a variety of diseases including cancer, RA, and other immune-related conditions (Hansel et al. 2010). A mAb targeting TNFα, infliximab, is one of the more ubiquitous biological drugs used against RA, and the first of its kind to do so. Though newer medications such as adalimumab have been developed. By binding to, and thus, blockading the TNFα cytokine directly in the synovium, inflammation, bone, and cartilage destruction is limited. Not all biological drugs are mAbs, Etanercept is a recombinant fusion protein that also targets TNFα and has shown efficacy in treating RA, particularly in combination with methotrexate. However, being administered intravenously, TNF-blocking agents have systemic effects and pose a risk of tuberculosis (TB) infection.

Thus, monitoring for TB and other infections is essential (Haraoui and Bykerk 2007; Mease 2007; Akahoshi and Watabe 2009).

Early on, signalling via IL-6 and its soluble receptor was shown to be of particular importance in driving acute peritoneal inflammation via leukocyte recruitment (Hurst et al. 2001). More evidence now shows this trans-signalling to be the predominant IL-6 based pathological mechanism in inflammatory diseases. To specifically target this, new treatments have been devised. Olamkicept is another recombinant protein biologic drug. It is a soluble (s)gp130Fc variant which selectively inhibits IL-6 trans-signalling (Rose-John et al. 2023).

Tocilizumab is a mAb for IL-6 receptor (IL-6R). IL-6 signalling can occur in two primary ways classical and trans-signalling, and in addition, trans-presentation. These are illustrated in **Figure 1.1.** As tocilizumab binds IL-6R both of these processes, classical and trans-signalling are inhibited (Kaneko 2013; Biggioggero et al. 2019; Jenkins et al. 2021).

TNFα and IL-6 are not the only targets of cytokine-blocking drugs in the treatment of RA. IL-17 is targeted by drugs such as secukinumab and ixekizumab (Genovese et al. 2014; Blanco et al. 2017) and a number of other drugs targeting cytokines have been in development. Unfortunately, despite promise in experimental models, Phase II and III clinical trials have shown these IL-17A inhibitors to possess insufficient clinical efficacy (Genovese et al. 2013; Tlustochowicz et al. 2016; Blanco et al. 2017; Tahir et al. 2017; Kim et al. 2022). Furthermore, the successes of TNFα targeting biologics have made them the benchmark to which new drugs are measured. Additionally, despite tocilizumab's current status as the default IL-6 inhibitor used in RA, there remains a problem. Those patients for whom these established drugs are ineffective. In part, this is because the pathways discussed through which these IL-6 and TNFα act can be activated by a host of other cytokines too. Thus, the development of new and more specific anti-cytokine agents remains an area of great interest (Choy et al. 2013).

For example, Granulocyte-macrophage colony-stimulating factor (GM-CSF) is another cytokine shown to regulate a wide range of inflammatory processes during RA, including stimulation of the production of monocytes dendritic cells and macrophages and enhanced production of a range of inflammatory cytokines (Wicks and Roberts 2015). When used in combination with methotrexate; Otilimab, an anti-GM-CSF monoclonal antibody, has proven effective in RA treatment (Taylor et al. 2023).

Despite the role IL-12 and IL-23 play in RA, which has been discussed previously, (Najm and McInnes 2021) blockade of the cytokines has not been effective in treating RA. However, biological drugs targeting them have shown success in treating psoriatic arthritis (Schurich et al. 2018).

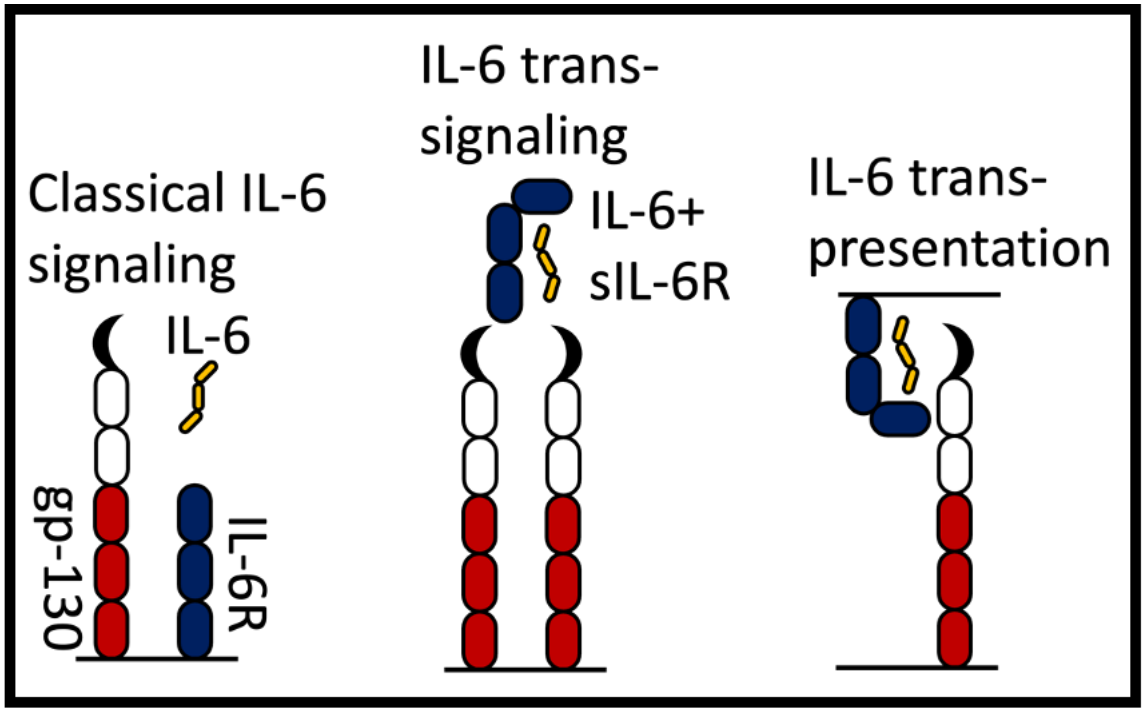


Figure 1.1: Models of IL-6 Signaling. Classical signaling occurs in cells expressing IL-6 receptor (IL-6R) and glycoprotein 130 (gp-130). Trans-signaling occurs when IL-6+circulating soluble IL-6R (sIL-6R) complexes binds to cells expressing only gp-130. Trans-presentation allows cells expressing an IL-6+IL-6R complex to active adjacent cells expressing gp-130. (Adapted from Jenkins *et al.,* 2021.)

1.4.2.2 Lymphocyte-targeting Agents

Activated T and B cells contribute to the pathology of RA in numerous ways. For instance, they generate autoantibodies and secrete pro-inflammatory cytokines. Thus, the cells themselves are therapeutic targets and biologics have emerged as a result (Choy et al. 2013; Smolen et al. 2018).

Rituximab is a blocking monoclonal antibody which depletes autoreactive B-cells by targeting the cell surface protein CD20 (an atypical tetraspanin expressed by mature B cells). Although cell death through antigen-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity appear to be the most likely mechanism for B-cell depletion, additional mechanisms can't be excluded (Edwards et al. 2005). In RA, virtually all patients show marked B-cell depletion, with biological response correlating with the degree of B-cell depletion and the longer-term maintenance of low numbers of CD20-positive cells. However, the response to therapy is complex and only 60% exhibit any improvement in clinical symptoms. (Vos et al. 2007; Vital et al. 2010).

Abatacept, another biologic, interferes with the immune activity of T-cells. It is a fusion protein of the extracellular domain of cytotoxic T-lymphocyte-associated protein 4 (CTLA-4; CD152) to the Fc region of IgG1. T-cells are activated in a two-step process, first antigen recognition followed by co-stimulation. This co-stimulation step requires the engagement of CD28 and CD86. CTLA-4 itself is a structural homologue for CD28, and thus is an antagonist for T-cell activation (Lenschow and Bluestone 1993; Shevach 2008; Buch et al. 2009).

1.4.3 A Biological Cocktail

A key difference between biological drugs and more traditional RA treatments, such as methotrexate and glucocorticoids, is their relative specificity to a particular cytokine, cell or other molecule as opposed to more indiscriminate immune suppression. Though biologics have shown great success there are still patients for whom disease remission is not achieved (Katchamart et al. 2010). A possible option therefore is to combine biologics to target multiple immune mediators while retaining a relatively high degree of specificity. This strategy, however, has its own pros and cons. The financial cost of biologics is considerably more than traditional treatments, prescribing multiple drugs only exacerbates this problem.

However, the caveat to this financial burden is that combination therapies could act synergistically and thus a lower dose of each medication would be required, lowering cost. The cumulation of negative side effects from using multiple drugs must also be considered, especially if a combination treatment is of little efficacy (Choy et al. 2013).

1.4.4 Biological Agents Last or First

The window of opportunity for treatment which exists during the beginning of clinically declinable RA has been mentioned previously. Indeed, it has been shown that a more aggressive regime of immunosuppression (via steroids) in conjunction with more traditional DMARDs is more effective than DMARDs alone in alleviating symptoms, even after steroids are ceased (Landewé et al. 2002). The question therefore arises; if during this disease stage a treatment that would stop the problematic autoimmune and inflammatory pathways in their tracks, could drug-free remission (or something close to it) be achieved?
Perhaps biologics hold the answer. However, as has been described there is difficulty in determining which biologic to prescribe and to whom. Thus, what is required for biologics to become the first option in the treatment of RA is patient stratification coupled with a detailed understanding of disease mechanisms.

1.4.5 Small-molecule Inhibitors

A relatively new class of chemically derived drugs are small-molecule inhibitors (also referred to as target synthetic DMARDs; tsDMARDs), which target intracellular cytokine signalling intermediates (Tanaka 2021). Beyond their therapeutic efficacy, they can be administered orally and are relatively cheap to manufacture (Van De Laar et al. 2020). Treatment with traditional biological drugs requires patients to attend hospital appointments to receive intravenous or subcutaneous drug infusions (typical doses range from 4-8mg/kg). These oral drug inhibitors, therefore, offer improvements in patient care and considerable advantages in health economics (Massalska et al. 2020).

Small-molecule inhibitors cross plasma membranes to target signalling molecules controlling cytokine action (Morris et al. 2018; Tanaka 2021). The JAK-STAT signalling pathway, which has been briefly mentioned previously, is of great relevance to inflammation and immunity. A summary illustration of the pathway is shown in **Figure 1.2**.

The JAK-STAT signalling pathway is a crucial cellular communication system that plays a vital role in various physiological processes, including immunity, cell growth, differentiation, and homeostasis. This pathway involves the activation of two families of proteins: JAKs and STATs. The pathway is initiated when a ligand, (typically a cytokine or growth factor) binds to its corresponding cell surface receptor (Aaronson and Horvath 2002). This binding event induces the multimerisation of the receptor subunits, bringing the associated JAK proteins into close proximity. The JAKs then undergo trans-phosphorylation, activating their kinase domains. The activated JAKs then phosphorylate specific tyrosine residues on the receptor, creating docking sites for the STAT proteins. The STATs are recruited to the receptor, where they are also phosphorylated by the JAKs. This phosphorylation event triggers the dimerisation of the STAT proteins, which then translocate to the nucleus and bind to specific DNA sequences, leading to the transcriptional activation of target genes (O'Shea and Plenge 2012).

The JAK-STAT pathway is tightly regulated by numerous mechanisms, including the suppressor of cytokine signalling (SOCS) proteins. SOCS proteins act as negative regulators of the JAK/STAT pathway, inhibiting the activity of the JAKs and preventing the phosphorylation of the STAT proteins. This regulation is crucial in maintaining the balance and homeostasis of cellular signalling, as dysregulation of the JAK/STAT pathway has been implicated in many diseases including RA, other IMIDs, cancer and neurodegenerative conditions (Collins et al. 2013; Kershaw et al. 2013).

As opposed to targeting molecules which utilise this signalling pathway JAK family kinases can themselves be targeted (Pesu et al. 2008). The first three developed are tofacitinib, baricitinib and upadacitinib. Each has shown similar efficacy to the biological drug benchmark of TNFα inhibitors (Harrington et al. 2020).

Given the previously discussed relationship between RA and CVD (Radner et al. 2017) questions have been raised about the safety of JAK inhibitors, in relation to the potential increased risk of death or complications related to CVD. RA patients receiving tofacitinib, who also had a history of CVD, were shown to have a risk of adverse cardiovascular outcomes in a randomised controlled trial. However, this increase was not statistically significant. There was no evidence of increased risk of cardiovascular outcomes in patients without a history of CVD (Khosrow-Khavar et al. 2022). Furthermore, observational cohort and population-based cohort studies have found no statistically

significant risk of venous thromboembolism (VTE) in patients receiving tofacitinib versus those taking the more established treatment; TNFα inhibitors (Desai et al. 2019; Desai et al. 2022). Selecting suitable patients for this treatment and providing appropriate follow-up is necessary to ensure better outcomes. Likewise, it remains important that patients are made aware of these risks and that their rheumatologist assess the risk/benefit of these medications (Massalska et al. 2020).

There are other signalling pathways which have been targeted for the treatment of RA. As they do not relate to the work of this thesis to such a degree as JAK inhibitors, they will be discussed only briefly. The mitogen-activated protein kinase (MAPK) is another important cell signalling pathway. It is involved in numerous processes including cell survival, proliferation, and growth (Braicu et al. 2019). Unfortunately, in a number of studies various MAPK inhibitors have not proved successful in treating RA (Hill et al. 2008; Cohen et al. 2009; Damjanov et al. 2009), in part due to their pleiotropic effects on immunity(Ferguson and Gray 2018).

Iguratimod (T-614), a nuclear factor κ-light-chain-enhancer of activated B cells (NF-κB) inhibitor, is a treatment for RA(Xie et al. 2020). Iguratimod effectively combats RA by regulating T cell subsets, suppressing antibody-secreting cells, and inhibiting bone resorption (Xu et al. 2015; Xie et al. 2020). Clinical studies confirm its safety and efficacy in both monotherapy and combination therapy with methotrexate (Hara et al. 2014; Ye et al. 2019).

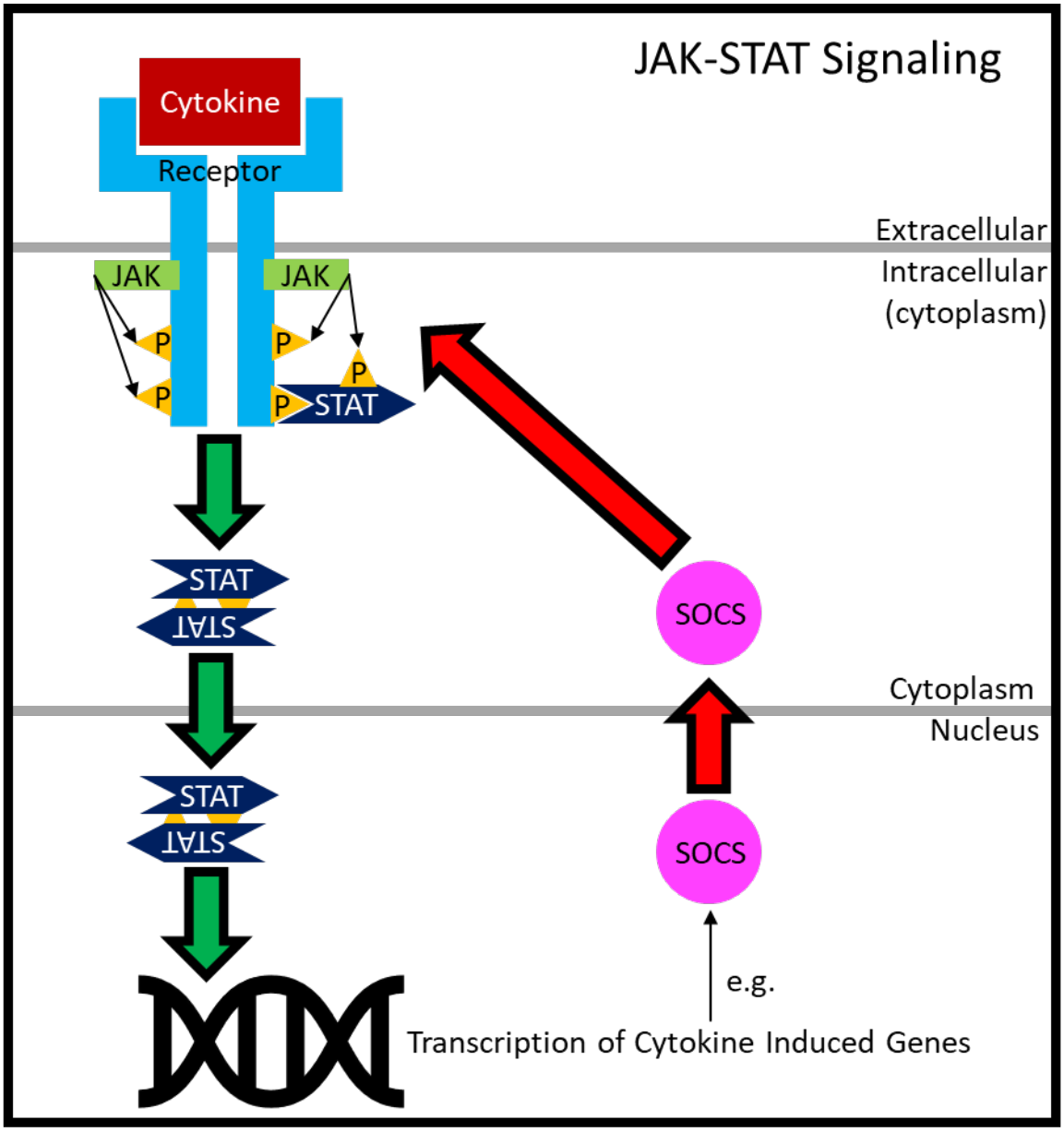


Figure 1.2 Summary of JAK-STAT Signaling. A cytokine binds to its receptor. This results in transactivation of the associated Janus Kinases (JAKs). The activated JAKs are then able to phosphorylate (orange triangles) the receptor. This recruits Signal Transducer and Activators of Transcription (STATs). These STATs form dimers and move to the nucleus. In the nucleus STATs upregulate certain cytokine-responsive genes. One such group of genes are suppressors of cytokine signaling (*SOCS*). The SOCS proteins target STATs to inhibit the signaling pathway via negative feedback. (Adapted from Morris *et al.,* 2018.)

1.5 Studying the Heterogeneity of Rheumatoid Arthritis

The numerous genetic risk alleles, epigenetic modifications and contributary environmental factors alongside the variety in disease presentation and the complexity of determining treatment in patients not achieving remission all attest to the heterogeneous nature of RA. As is the case for all diseases and health issues a strong understanding of the nature of the condition and its idiosyncrasies is imperative to devising treatments so that better patient outcomes may be achieved. With this in mind the study of RA's heterogeneity can serve to aid in the stratification of patients, the discovery of biomarkers and determining possible treatment targets.

1.5.1 Synovial Biopsies

Arthroscopic and ultrasound-guided biopsies are well tolerated by patients, with minor complications (e.g., infection of the biopsy site, discomfort) occurring in a small number of patients. Synovial biopsies are carried out for two reasons: clinical diagnosis and research. In a clinical trial setting histological assessments of joint biopsies have been used to classify disease pathology, including infections, inflammatory and non-inflammatory causes. Technological advances in the assessment of joint biopsies have significantly enhanced our understanding of disease heterogeneity in RA and provided greater insights than more traditional studies of synovial fluids (Sitt et al. 2016; Just et al. 2018; Ingegnoli et al. 2021; Saraiva 2021).

Histological treatment and analysis of the tissue recovered from ultrasound-guided synovial biopsies reveal differing pathophysiological phenotypes characterised by the prevalence and histological organisation of different cell populations within the inflamed synovium. Three histopathological phenotypes have been defined, they are Follicular, Diffuse and Pauci immune. Follicular is characterised by the development of ectopic lymphoid structures and the general abundance of lymphoid cells. Diffuse has a large proportion of myeloid cells, though there can be some overlap with the follicular phenotype. Finally, Pauci immune is notable for its general lack of infiltrating cells. As it can be observed during active, early synovitis and untreated it is indeed a disease phenotype as opposed to partially resolved or 'burnt out' inflammation (Pitzalis et al. 2013).

To enhance understanding of disease pathophysiology, treatment prediction, and novel therapeutic targets synovial biopsied can be used in conjunction with additional technologies. The aforementioned histopathological phenotypes correspond to distinct disease characteristics, aiding in patient stratification and predicting responses to various treatments, including biological interventions (Rivellese et al. 2022). Various technologies including, RNA sequencing (RNA-seq), mass spectrometry and flow cytometry unveil previously unknown cell subpopulations, offering fresh insights into disease mechanisms and potential therapeutic avenues. Recent advancements also shed light on specific macrophage subpopulations in synovial tissue, bearing transcriptomic signatures enriched in anti-inflammatory regulators. These subpopulations are associated with disease remission, suggesting their potential as therapeutic targets (Johnsson and Najm 2021; Andreozzi et al. 2022).

1.5.2 Murine Antigen-Induced Arthritis

There exist several animal models of inflammatory arthritis. While models using rats and even primates have been established (Choudhary et al. 2018; Na et al. 2020), the following will primarily focus on murine models. The most common model is collagen-induced arthritis (CIA). Mice are immunised with an emulsion of type II collagen and complete Freund's adjuvant. This leads to the mouse producing autoantibodies against type II collagen, resulting in disease manifestation with inflammation in the paws (Brand et al. 2007).

CIA and the Antigen Induced Arthritis (AIA) model of arthritis used in this work (discussed below) are both forms of induced arthritis. However, there are alternatives. Genetically manipulated spontaneous arthritis models utilise genetically modified mice, which will innately develop arthritis (Asquith et al. 2009). One example is a transgenic mouse which will over-express human TNFα. This model is particularly useful in the study of RA's chronicity as unlike other models like CIA it is not self-limiting (Keffer et al. 1991).

Despite the various limitations common to all animal models methylated bovine serum albumin (mBSA) Antigen Induced Arthritis (AIA) has enabled a greater study of RA and presents the opportunity to utilise novel assays and techniques to examine its inflammation in new ways. AIA has been used in the lab previously and proved successful.

The protocol for the model is explained in the next chapter. Briefly, mice are immunised with mBSA, which is then followed with an intra-articular injection with mBSA to elicit an immune response localised to the knee joint (Jones et al. 2018).

The use of cytokine receptor knockout animals in AIA illuminates the role those cytokines play in inflammation, furthermore, they are analogous to the three previously described histopathologies. *Il6ra*$^{-/-}$ show an inflammatory pattern similar to Pauci Immune arthritis, *Il27ra*$^{-/-}$ are similar to Fibrotic arthritis and Wild type (*Wt*) are similar to Diffuse arthritis (Jones et al. 2013; Dennis et al. 2014; Jones et al. 2015; Lewis et al. 2019; Twohig et al. 2019). Histology images of these three murine and three human disease phenotypes are shown in **Figure 1.3.** These analogues highlight the importance of IL-6 and IL-27 and their associated signalling pathways (for instance JAK STAT) in not only the pathology of RA but its heterogeneity particularly and afford us the opportunity to study the role of these particular cytokines in RA's pathology.

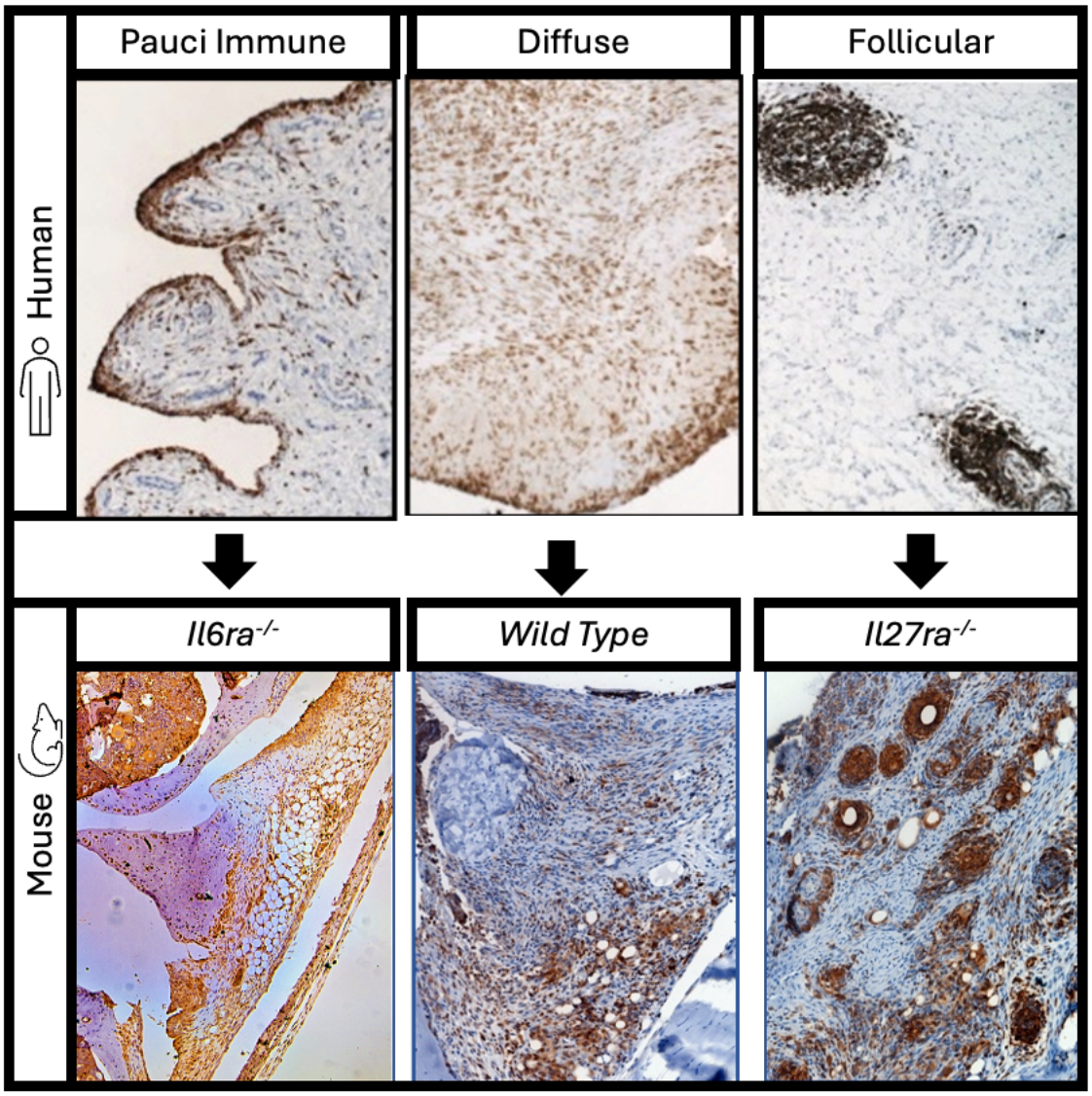


Figure 1.3: Histology of Human and Murine Synovitis. The human pauci immune, diffuse and follicular disease phenotypes share inflammatory features with *Il6ra*$^{-/-}$, *Wt* and *Il27ra*$^{-/-}$ murine synovitis respectively. (Adapted from Pitzalis *et al.*, 2014 and Hill, 2019).

1.6 Surveying the Epigenetic Landscape of Synovitis

Next-generation sequencing (NGS) has all but replaced older technologies such as Sanger sequencing. NGS is less expansive and faster while retaining accuracy and precision (Metzker 2009). NGS data, specifically RNA sequencing (RNA-seq), has been generated from the murine AIA model previously and has shown transcriptional differences relating to disease pathology between the three mouse genotypes. While this data is of immense value to our understanding of RA's heterogeneity it cannot tell us the way these transcriptional differences arise. Therefore, it is necessary to return to determinates of transcription, specifically chromatin accessibility and transcription factor binding and to use functional NGS technology assays to explore these areas.

1.6.1 Chromatin Accessibility

Chromatin is the complex of DNA and various nuclear proteins, principally histones. The primary function of chromatin is to package long DNA molecules into a densely compact structure that protects genomic DNA from enzymatic attack or damage during mitosis and meiosis (Felsenfeld 1978). However, the structural organisation of chromatin is highly dynamic and tightly regulated to enable the transcriptional regulation of gene promoters. The cellular mechanisms accounting for the structural rearrangement of chromatin are currently poorly defined. Yet, advances in molecular biology methods provide exciting opportunities to understand how cells respond to environmental cues that steer physiology or pathophysiology.

The accessible genome is very small, comprising of only 2-3% of an organism's total DNA. Despite its meagre size, however, the accessible genome encompasses more than 90% of regions to which transcription factors bind. Furthermore, with very few exceptions the vast majority of transcription factors will bind exclusively to open or permissive chromatin (Thurman et al. 2012).

Nucleosomes are the foundational repeating sub-unit of chromatin. They constitute the primary degree of high-order packaging of chromosomal DNA by histone proteins. Each nucleosome consists of around 200 DNA base pairs (the length varies from eukaryotic species to species) wrapped around a histone core. This wrapping constitutes an approximately sixfold reduction in the length of the DNA. The core is made up of an octamer

of pairs of four histones, they are H2A, H2B, H3 and H4. They are related by a single dyad axis (Olins and Olins 2003).

There exist myriad other non-histone macromolecules which bind directly or indirectly, dubbed chromatin-binding factors, which can occlude access to DNA. These, alongside the topological organisation and occupancy of nucleosomes, influence the degree to which nuclear macromolecules (for instance transcription factors) can make physical contact with genomic DNA. These interactions are the defining influencers of chromatin accessibility (Klemm et al. 2019).

There are a variety of mechanisms in which post-translational modifications and the composition of nucleosomes, reflecting distinct functional states, are able to influence and regulate chromatin accessibility. For example, the modulation of a nucleosome's affinity for active chromatin remodelling and the steric hindrance (arising from the spatial arrangements of atoms) lead to altered transcription factor binding (Allis and Jenuwein 2016; Dann et al. 2017).

Across the genome, the density and distribution of nucleosomes are not uniform. They do not occur at regular intervals, after a given number of base pairs, instead their positioning along the DNA is associated with said DNA's function. At regulatory loci, enhancer regions, insulator elements and transcribed gene bodies histones are depleted to relatively low densities. However, nucleosomes are more abundant and therefore denser in constitutive and facultative heterochromatin (Lee et al., 2004; Thurman et al., 2012). Similarly, the amount of time a length of DNA remains coiled around a nucleosome core can vary greatly across the genome. This phenomenon is known as nucleosome occupancy (Klemm et al. 2019).

There is not a dichotomy of open and closed chromatin but rather chromatin accessibility exists on a continuum. For ease of understanding degrees of openness can be broadly categorised into one of three states: closed chromatin, permissive chromatin, and open chromatin. Closed chromatin is totally inaccessible to nuclear macromolecules. However, permissive chromatin is sufficiently dynamic to allow transcription factors to bind and begin sequence-specific accessibility remodelling. This progresses to fully open chromatin and, if applicable to the locus, gene regulation or transcription (Klemm et al. 2019).

The causes of the varying degrees of chromatin accessibility are themselves varied. DNA within the nucleus is frequently bound to a range of molecules. Architectural proteins, a group including the nucleosome's core histone proteins as well as linker and isolator proteins, constitute a major part of this range. They alongside RNA polymerases and transcription factors facilitate higher-order chromosome organisation. This, in combination with the previously discussed non-uniformity of nucleosome distribution and occupancy, are the drivers determining chromatin accessibility (Klemm et al. 2019).

There are several different techniques that can be utilised to assess chromatin accessibility, including micrococcal nuclease digestion with sequencing (MNase-seq), DNase I hypersensitive site sequencing (DNase-seq), Formaldehyde-Assisted Isolation of Regulatory Elements and sequencing (FAIRE-seq) and the Assay for Transposase Accessible Chromatin using sequencing (ATAC-seq) (Tsompana and Buck 2014). However, the basic principle is an almost universal constant; the quantification of the susceptibility of chromatin to cleavage of the DNA from which it is constituted or its susceptibility to enzymatic methylation. Though it may seem likely that this susceptibility, and subsequent measurement of chromatin accessibility, would vary depending on the assay employed there is in fact a high degree of conservation of chromatin accessibility observed across a wide range of molecular probes (Buenrostro et al. 2013). Therefore, the use of one method (in the case of ATAC-seq the probe in question is a transposase enzyme) does not compromise the validity of the data.

The discovery and development of ATAC-seq came about due to two observations. Previously transposase enzymes, which had been preloaded with sequencing adapters had been used to generate so-called 'tagmentation' libraries for high throughput sequencing. The Tn5 enzyme used was simultaneously able to fragment and tag genomic DNA (Adey et al. 2010). The Tn5 enzyme was also able to insert itself into regions of the genome devoid of nucleosomes (Gangadharan et al. 2010). A summary of the ATAC-seq method is shown in **Figure 1.4**.

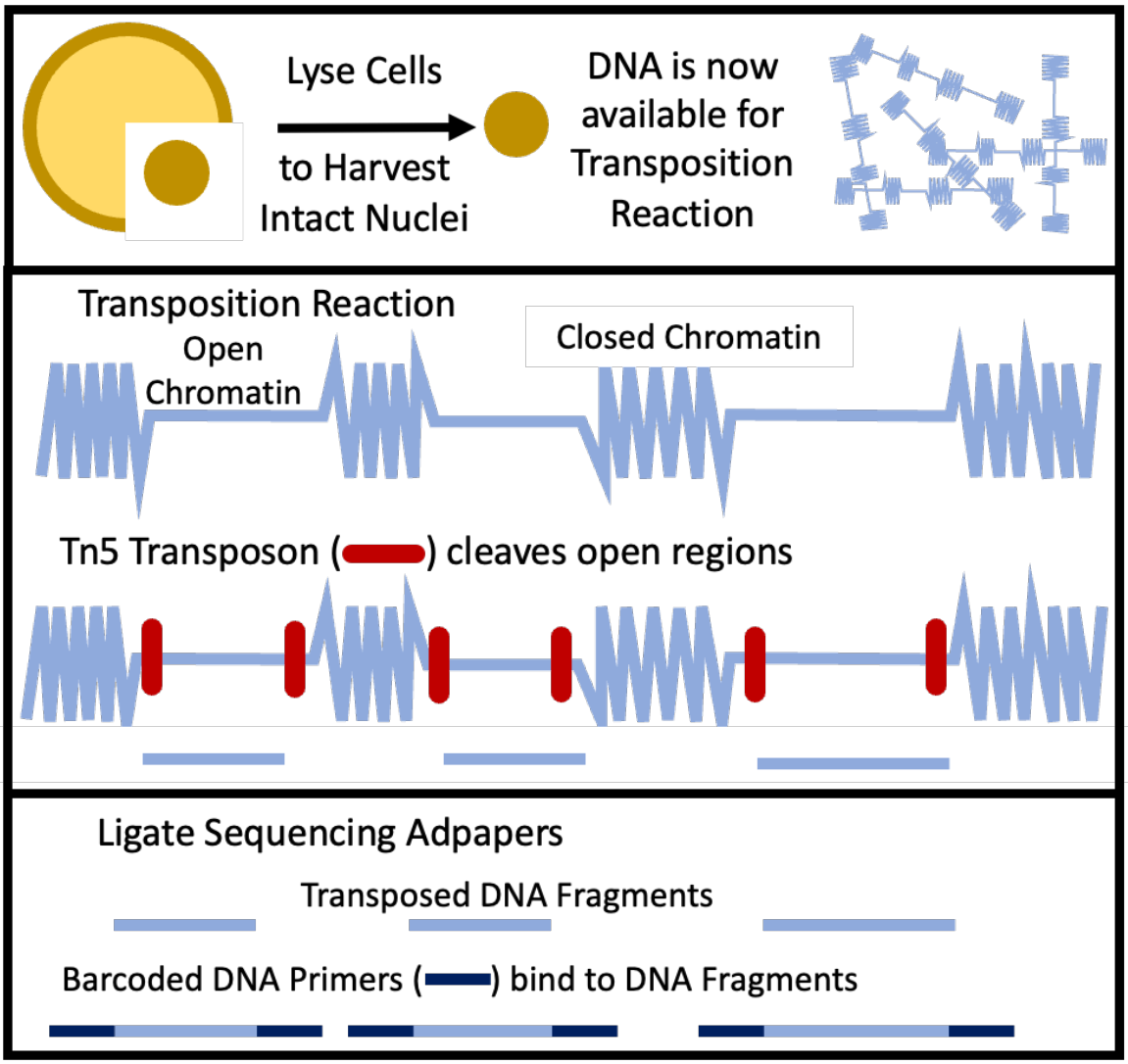


Figure 1.4: Summary of the ATAC-seq Method. Cells are lysed to isolate the nuclei. During the transposition reaction the mutant Tn5 Transposon cleaves the open chromatin regions. To these transposed DNA fragments barcoded DNA primers and ligated and libraries generated for NGS.

1.6.2 Transcription Factor Binding

Though there are a variety of mechanisms by which the expression of genes can be regulated, transcription, the process by which DNA is copied to messenger RNA, and its regulation play a leading role (Casamassimi and Ciccodicola 2019). There are other DNA-binding proteins, such as histones which have been previously described, however, transcription factors are the most relevant to transcription itself (Latchman 1997).

Within the genome, there are specific, short DNA sequences which enable the binding of transcription factors. These short sequences are common across a group of genes which are regulated by a shared transcription factor. When a transcription factor binds to DNA it is not a prerequisite to gene expression as these regulatory proteins can regulate transcription positively and negatively. Furthermore, transcription factor binding sites are not always adjacent to genes in terms of genomic sequence, these loci are known as enhancer regions. However, the three-dimensional nature of DNA enables open chromatin regions to loop around, thus seemingly distant enhancer regions are able to be directly involved in gene regulation (Latchman 1997; Panigrahi and O'Malley 2021). By their very nature, both transcription factors which negatively regulate transcription and enhancer elements are ignored by RNA-seq however, both phenomena are of interest to the understanding of gene regulation and the subsequent understanding of RA's complexity. Thus, assessing transcription factor binding both in and of itself and in conjunction with other methods gives a more comprehensive view of the epigenetic landscape of synovitis. To that end, the employment of Chromatin Immunoprecipitation sequencing (ChIP-seq) has enabled such assessment.

During the ChIP-seq protocol, a summary of which is shown in **Figure 1.5**, DNA-protein complexes are crosslinked. An antibody specific to the DNA-bound protein (typically a transcription factor) is utilised to precipitate the DNA-protein complexes of interest. The sample is purified and sequenced. Unlike RNA-seq and ATAC-seq which sequence all transcribed genes and open regions respectively, ChIP-seq is a far more targeted assay (Park 2009).

This specificity brings its own challenges, primarily to understand the relationship between two or more DNA binding proteins multiple experiments must be carried out. As with ATAC-seq, which shows only if a genomic region is Tn5 accessible, ChIP-seq only shows

protein binding. Therefore, it is of the utmost importance it is remembered that open chromatin and or transcription factor binding does not equal transcription or gene expression.

The AIA experiments included *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ animals as their respective cytokines are known to influence inflammation and RA's heterogeneity. IL6 and IL27 act through the JAK-STAT signalling pathway. The interplay between STAT1 and STAT3 is of particular interest with regard to their modulation of inflammation (Jones et al. 2015; Twohig et al. 2019) and thus STAT1 and STAT3 have been the focus of ChIP-seq experiments.

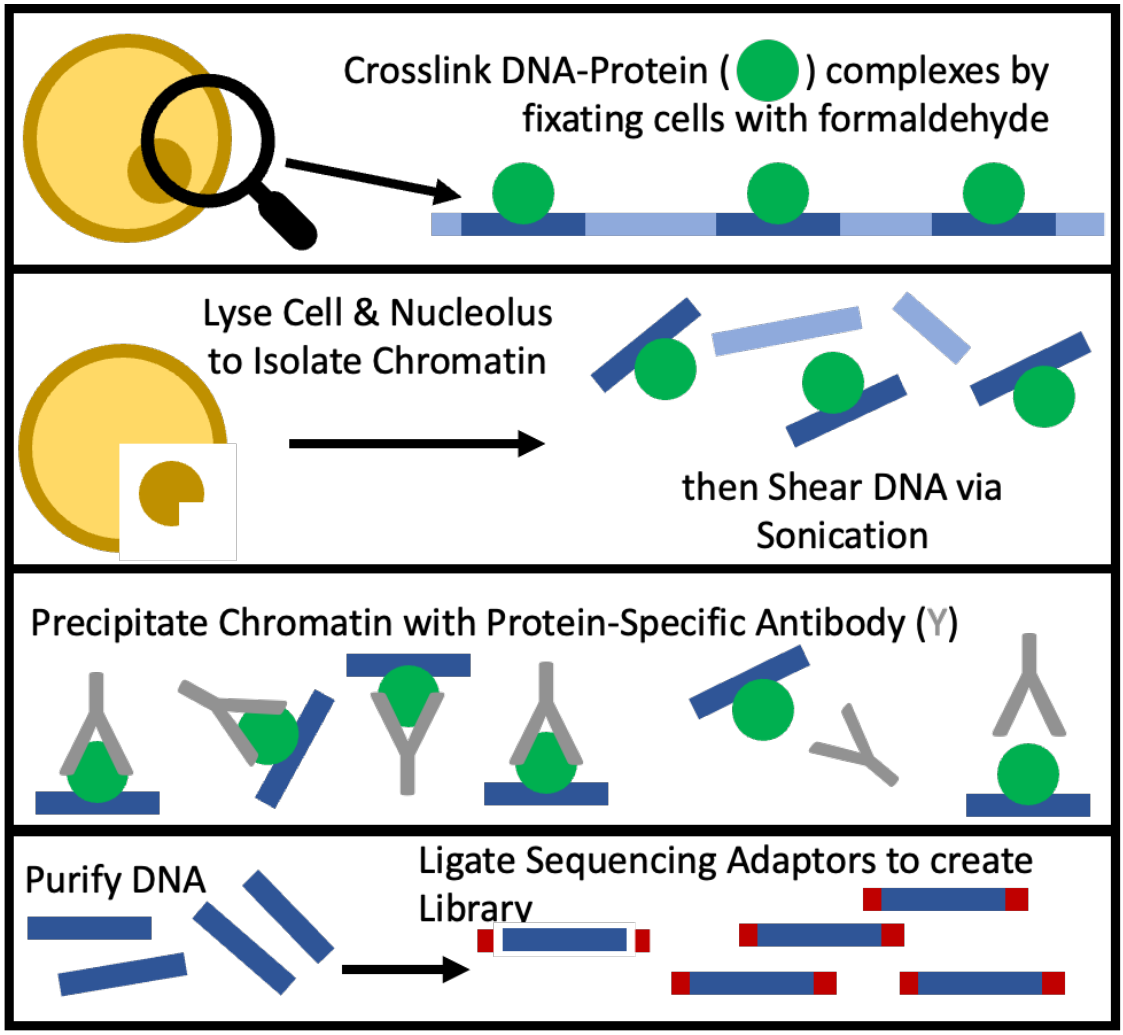


Figure 1.5: Summary of ChIP-seq Method. Cell isolates are fixed with formaldehyde to cross-link DNA-Protein complexes. Cells and nuclei are lysed, and genomic DNA extracted prior to the disruption of chromatin by sonication. Transcription factors bound to specific chromatin fragments were precipitation using protein-specific antibody. DNA is purified from immunoprecipitated complexes, adaptors ligated for library generation and downstream NGS.

1.7 Hypothesise and Aims

The heterogeneity of RA has been described above. Differences in histology of the inflamed synovium have demonstrated the important role IL-6 and IL-27 play in shaping disease outcomes. NGS technologies in combination with the AIA model of joint inflammation will be used to explore this in greater detail. I hypothesise that differences in the pattern of synovitis seen in Wt, Il6ra-deficient and Il27ra-deficient mice with antigen-induced arthritis arise through alternate epigenetic mechanisms affecting synovial gene regulation. I further hypothesise that a detailed characterisation of the transcriptional mechanisms (specifically STA1 and STAT3 mechanisms) controlling antigen-induced arthritis offer interesting insights into the development of synovitis in human rheumatoid arthritis.

To test these, I aim to generate ATAC-seq data of the inflamed synovium. This will involve AIA experiments being carried out on *Wt, Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice. Synovial tissue will be collected on Day 3 and Day 10 post-arthritis induction. Samples will undergo ATAC-seq to show genomic regions of chromatic accessibility under synovial inflammation. Additionally, I aim to generate STAT1 and STAT3 ChIP-seq data of the inflamed synovium. this will also involve AIA experiments being carried out on *Wt, Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice. Synovial tissue will be collected on Day 3 and Day 10 post-arthritis induction. Samples will then undergo STAT1 and STAT3 chip seq. This will show STAT1 and STAT3 transcription factor binding activity under synovial inflammation.

Building upon the first of these aims I hypothesise that the varying histological presentations of synovitis in WT, Il6ra-deficient and Il27ra-deficient mice with antigen-induced arthritis are influenced by variations in chromatin accessibility between the genotypes and over the time course of the disease. From the ATAC-seq data, the patterns of chromatin accessibility during synovitis can be characterised. Furthermore, the way in which this varies due to IL-6 and IL-27 receptor deficiency will elucidate how these cytokines can influence the epigenetic landscape of synovitis. Ultimately, by exploring these changes we will see how they relate to gene expression and downstream effects in relation to disease. Thus, I aim to identify changes in gene accessibility in the face of synovial inflammation and IL-6 and IL-27 receptor deficiency.

With regards to the second key aim I hypothesise that STAT1 and STAT3 transcription factors share a complex regulatory interplay that shapes the histological

presentation of synovitis in WT, Il6ra-deficient and Il27ra-deficient mice with antigen-induced arthritis. The role of STAT1 and STAT3 in RA pathology has been described. Characterising the behaviour of these transcription factors in synovitis, particularly in relation to IL-6 and IL-27 receptor deficiency will show how the interplay between STAT1 and STAT3 shape the disease. Thus, I aim to identify changes in STAT1 and STAT3 gene binding in the face of synovial inflammation and IL-6 and IL-27 receptor deficiency.

Chapter 2:

Methods and Materials

2.1 Methods

2.1.1 Mice

Experiments involved male C57BL/6 *Il27ra*$^{-/-}$, *Il6ra*$^{-/-}$, and *WT* mice. *Il27ra*$^{-/-}$ mice were originally sourced from The Jackson Laboratory (line B6N. 129P2-*Il27ra*^{tm1Mak}/J)(Yoshida et al. 2001). An exon which encodes part of the extracellular fibronectin type II domain of the IL-27 receptor was replaced with a neomycin resistance (neo) cassette. *Il6ra*$^{-/-}$ mice were originally generated by GlaxoSmithKline (Stevenage, UK) *via* a traditional replacement vector. This deleted exons 4, 5 and 6 to disrupt IL-6 recognition of its cognate receptor (Jones et al. 2010). All knockout mice were bred and housed at Cardiff University's Joint Biological Services (JBIOS) facility under specific-pathogen-free (SPF) conditions. *WT* mice were purchased from an accredited UK supplier (Charles River) and acclimatised for at least one week in JBIOS before entering any experimental procedures. All experiments were carried out in accordance with Home Office-approved project licences PB3E4EE13 (2019) and PE8BCF782 (2020 onwards). Due consideration was given to the concepts of Replacement, Reduction and Refinement when conducting experiments involving mice.

2.1.2 T-cell Isolation from Murine Splenocytes

The spleens from 1-2 mice were isolated and kept on ice in supplemented culture media (**Table 1**). The spleens were homogenised and passed through a 40µm strainer to obtain a single-cell suspension. Cells were resuspended in 10mls of phosphate-buffered saline (PBS) and centrifuged at 400xg for 5 minutes at 4^{o}C. The supernatant was decanted, and the cells were resuspended in 5ml of ice-cold Red Blood Cell Lysis solution before incubation on ice for 1 minute. Supplemented RPMI-1640 was to neutralize the lysis buffer and the cells were pelleted by centrifugation.

T-cells were purified by magnetic-activated cell sorting (MACS) using a CD4^{+} T cell isolation kit (Miltenyi Biotec). Following tissue homogenization and the isolation of splenic cells, cells were resuspended in 400µl of MACS buffer (Table 2.2). As per the kit's instructions, cells were incubated (5 minutes, 4^{o}C) with the provided biotin-conjugated monoclonal antibody cocktail targeting CD8a, CD11b, CD11c, CD19, CD45R (B220), CD4b (DX5), CD105, MHC-class II, Ter-119 and γ/δ TCR (50µl per spleen). Cells were incubated for 10 minutes at 4°C with a

conjugated anti-biotin monoclonal antibody (isotype: mouse IgG1) labelled with MicroBeads (20 µL per 10^7 cells). Next, 20 ml of MACS buffer (**Table 2**) was added, and the cells were centrifuged at 400xg or 10 minutes. During this spin, an LS positive selection column was placed into the magnetic stand and 3ml of MACS buffer was applied to equilibrate the extraction column. After removing the supernatant, the cells were resuspended in 500µl of MACS buffer and transferred to the LS column. The column was washed three times with 1ml of MACS buffer and the eluted $CD4^+$ T-cells counted. The cells were then divided into a 24-well plate, with 2 x 10^6 cells/well. The wells were then topped up to 500µl.

2.1.3 Cytokine Stimulation of CD4 T-Cells

IL-6 (R&D Systems) was then added for a final concentration of 20ng/ml. The cells were incubated at 37°C for 30 minutes. The cells were collected, pooled, and subjected to centrifugation at 400xg for 5 minutes at 4°C. They were washed once with cold PBS and centrifuged as before. The cells were resuspended in serum-free media, counted and chromatin immunoprecipitation performed as described in Section 2.1.6.

2.1.4 Murine Antigen-Induced Arthritis

Antigen-induced arthritis (AIA) was induced in male mice, aged 8-12 weeks (Jones et al. 2018). All injections were carried out on mice anaesthetized with isoflurane. Care was taken to closely monitor the mice during anaesthesia induction and recovery.

At the beginning of the model (Day -21) mice underwent a 100µl subcutaneous (s.c.) injection of an emulsion of methylated BSA (mBSA) and complete Freund's adjuvant (CFA) utilising a 1ml syringe and 25-gauge needle. The emulsion was prepared by forcibly mixing equal parts mBSA (2mg/ml) and CFA with a 50ml syringe and an 18-gauge needle on ice. Immediately following this injection, the mice were subjected to an intraperitoneal (i.p) injection of *Bordella pertussis* toxin (1.6 ng/µl) using a 1ml syringe and 25-gauge needle. Seven days later (Day -14), the mice received an immunological booster by re-injection with mBSA/CFA as before. It should be noted that the first injection was always on the left flank of the animal and the second on the right. After 14 days, arthritis was induced (Day 0) by intra-articular (i.a.) injection (both knees) with 10µl of mBSA (10mg/ml) using an insulin

syringe. Arthritis progression and knee swelling were monitored with a POCO 2T micrometer (Kreoplin). After 3- or 10-days post-induction mice were culled via CO_2 and synovial tissue collected.

2.1.5 Harvest of Synovial Tissue and Collagenase Digestion

Synovial tissue was dissected from the mouse knee-joint cavity. The skin was removed from the leg; the patella detached from the tibia and the articular capsule opened. The synovial membrane and infrapatellar fat pad were excised from the joint. The method of synovial tissue collection was identical for both treatment naïve mice and mice primed for AIA.

To facilitate the extraction of genomic DNA, all synovial tissue was treated with collagenase to disrupt the extracellular matrix supporting the synovial cells. Tissue was collected and placed into 1mg/ml Collagenase IV (Worthington) in RPMI-1640 on ice. For ChIP-seq all mouse knees (of a given time point and genotype) would be pooled in ~30ml of collagenase-supplemented media. For ATAC-seq both knees of individual mice would be pooled in ~10ml of collagenase supplemented media. The tissue was treated with the collagenase media for 2 hours at 37^{o}C. Every 20 minutes the tubes were removed from the water bath and mixed using a vortex. Following digestion, the media was passed through a 40µm strainer to remove any undigested tissue debris or fur contamination. The samples were then subjected to centrifugation at 350xg for 5 minutes. Finally, cells were resuspended in serum-free RPMI-1640 media and counted.

2.1.6 Chromatin Immunoprecipitation (ChIP)

2.1.6.1 Crosslinking and preparation of cell nuclei

ChIP-seq and or quantitative polymerase chain reaction (qPCR) analysis typically requires 4-7 x 10^6 cells per antibody and efficient sonication (described later) is performed at this cell number range. The following protocol was used for both AIA synovial cells (Section 2.1.4-2.1.5) and splenocytes (Section 2.1.2-2.1.3). Thus, after the cell count described at the end of sections 2.1.3 and 2.1.5 cells would be split according to their number and how many antibodies were to be used.

Cells were then resuspended in 10 ml of serum-free RPMI-1640 media. Crosslinking of the protein-DNA complexes was achieved by fixation of the cells with 270µl of 38% (v:v) formaldehyde for 10 minutes at room temperature. The formaldehyde was neutralized by the addition of 1.25 ml of 1M glycine (pH 2.5) and incubated at room temperature for 5 minutes. Treated cells were centrifuged at 310xg for 8 minutes at 4°C and washed with 2 ml of ice-cold PBS (pH 7.4) and centrifuged at 860xg for 5 minutes at 4°C. The cell pellet was resuspended in 1ml of Cell Lysis Buffer (CLB) (**Table 3**) and transferred into 1.5ml tubes. Tubes were incubated for 10 minutes on ice and subsequently centrifuged at 500xg for 5 minutes at 4°C. The supernatant was carefully decanted, and the nuclei resuspended in 275µl of nuclei lysis buffer (NLB) (**Table 4**). Samples were incubated on ice for a further 10 minutes before the addition of 165µl of immunoprecipitation dilution buffer (IPDB) (**Table 5**).

2.1.6.2 Fragmentation

The samples were sonicated with a Bioruptor (Diagenode) for 90 continuous cycles of 30 seconds on and 30 seconds off. To check the efficiency of the fragmentation, 20µl of sample was removed and treated with 1.4 µl of 5M NaCl, 2.6µl ultrapure water and 1µl Proteinase-K (20mg/ml) (Invitrogen). Following a 2-hour incubation at 65°C, samples were separated by electrophoresis on a 2% (w:v) agarose gel to evaluate DNA shearing. The rest of the sample was flash-frozen in liquid nitrogen and kept at -80°C. If shearing was insufficient the samples were defrosted on ice and the sonication repeated as described for an additional 30-60-cycles depending on the observed degree of shearing.

2.1.6.3 Immunoprecipitation

Fragmented DNA samples were diluted with 935µl of IPDB. A 1ml sample was subjected to antibody immunoprecipitation with 100µl removed for the Input control. The input sample was returned to -80°C for storage. For samples destined for qPCR analysis 275µl of sample was removed and treated with an Immunoglobulin G (IgG) Antibody. Antibodies against transcriptional regulators (or IgG) (**Table 12**) were added to each sample and incubated at 4°C on a rotating wheel overnight. To recover antibody-antigen complexes, 40µl of Pierce

Protein A/G Magnetic Beads (Thermo Scientific) were pre-washed with 1ml of ice-cold PBS and recovered by attraction to a magnet. Beads were resuspended in PBS and 40µl added to each antibody-treated fraction. Samples were incubated for 2 hours at room temperature on the rotating wheel. Antibody-antigen complexes were isolated using a magnet and washed twice with 1ml of immunoprecipitation wash buffer 1 (IPWB1) (**Table 6**). Fractions were then subjected to two washes with 1ml of IPWB2 (**Table 7**) and two washes with 1ml of Tris-EDTA (TE) buffer (**Table 8**). Following each wash step, antibody-antigen complexes were recovered using a magnet. Finally, the beads were resuspended in 100µl of TE buffer and transferred to a new Eppendorf LoBind 1 ml tube. The input sample was removed from -80^{o}C and allowed to thaw and transferred to a LoBind tube (note, all 1.5 ml tubes used from this point on are LoBind). Next, 10µl of 10% (w:v) sodium dodecyl sulphate (SDS), 6µl of 5M NaCl and 2µl of RNaseA (Thermo Fisher) was added to each sample and incubated at 65^{o}C for 2-4 hours. This treatment allows the release of the genomic DNA fragments from the immunoprecipitated transcriptional regulator. Samples were briefly mixed with a vortex and returned to the magnetic stand to separate protein components from the genomic DNA fragments. The supernatant was transferred to a new tube. The beads were then washed with 100µl of TE, which was then combined with the supernatant. Samples were then incubated at 45^{o}C overnight.

2.1.6.4 DNA Purification

Following the aforementioned overnight incubation, 200µl of Phenol/Chloroform/Isoamyl alcohol (25:24:1) was added to the samples and centrifuged at 16,000xg for 5 minutes at room temperature. Taking care not to disturb the tubes, the top aqueous phase was removed and transferred to a new tube. To this 10µl of glycogen, 20µl of 3M Sodium acetate (pH 5.2) and 500µl absolute ethanol were added and the samples were mixed with a vortex. The samples were then incubated for two hours at -80^{o}C to precipitate. Samples were centrifuged at 16,000xg for 20 minutes at 4^{o}C. The supernatant was removed, and the pellet was washed with ice-cold 70% (v:v) ethanol. Samples were centrifuged at 16,000xg for 10 minutes at 4^{o}C. The supernatant was carefully removed, and the pellet was allowed to air dry at room temperature for 15 minutes. The DNA pellets were finally resuspended in 50µl ultrapure water and stored at -80^{o}C prior to library preparation or qPCR analysis.

2.1.7 Quantitative PCR (qPCR) analysis

qPCR was performed on a Thermo Fisher Scientific ViiA 7 Real-Time PCR System. Samples were analysed in triplicate for each of the four oligonucleotide primer sets used. TaqMan oligonucleotide primers corresponding to the promoter regions for *Irf1* (Interferon Regulatory Factor 1), *Socs3* and *Stat3* were designed using the "Custom Plus TaqMan Assay design tool" (**Table 13**). These sequences were designed to flank identified STAT transcription binding motifs within each of the promoters.

Each amplification reaction contained 5µl of TaqMan Fast Advanced Master Mix, 4.5µl of water, 0.5µl of primer, and 1µl of sample genomic DNA. Each reaction was repeated in triplicate. The samples included those which had undergone immunoprecipitation with STAT1, STAT3 or p300 antibodies as well as the IgG and input controls. For the control wells (one per primer) the sample was replaced with an additional 1µl of water. The enrichment achieved was an average taken from the three replicates. This was then normalised first against the input control sample, followed by the IgG control sample.

2.1.8 ChIP-Seq Library Preparation

Library Preparation utilised the Illumina TruSeq ChIP Kit (Part number: 15023092). Details of the contents of the kit and their preparation are described in **Table 14**.

2.1.8.1 Repair of damaged or incompatible 5' and 3' protruding DNA sequences

The 50µl of samples frozen after the DNA purification step (Section 2.1.6.4) were thawed on ice and transferred to the first row of a 96-well plate. To each sample, 10µl of Resuspension Buffer (RSB) and 40µl of End Repair Mix (ERP) were added. This reaction mix contained an optimized combination of T4 DNA polymerase, Klenow fragment and T4 polynucleotide kinase for efficient blunt-ending of the fragmented DNA and appropriate phosphorylation of the DNA ends. The plate was sealed with an adhesive seal and placed on a preheated thermocycler at 30^{o}C (lid set to 100^{o}C) for 30 minutes. After incubation, the plate was removed from the thermocycler and the seal removed from the plate. To this,160µl of AMPure XP Beads were added and gently mixed by pipetting up and down 10 times. The

plate was incubated at room temperature for 15 minutes. Next, the plate was placed on the magnetic stand for 15 minutes. Using a 200µl pipette 127.5µl of supernatant was removed and discarded. This step was repeated to remove all the supernatant. With the plate still on the magnetic stand 200µl of 80% (v:v) ethanol was added, left for 30 seconds before being removed and discarded. This ethanol wash step was repeated twice. The plate was then allowed to air dry at room temperature for 15 minutes and removed from the magnetic stand. The pellet in each well was resuspended with 17.5µl RSB and gently mixed by pipetting. The plate was incubated at room temperature for 2 minutes. The plate was then replaced on the magnetic stand and left for 5 minutes. Finally, 15µl of clear supernatant was removed and transferred to the next row of the plate. If not immediately progressing to the next step the plate was sealed with an adhesive seal and stored at -20^{o}C overnight.

2.1.8.2 Sample preparation for 3' polyadenylation

If the samples and PCR plate were frozen at the end of the last step, they were first thawed at room temperature and then subjected to centrifugation at 280xg for 1 minute. The adhesive seal was then removed. To each sample, 2.5µl of RSB and 12.5µl A-Tailing Mix was added. The plate was sealed with an adhesive seal and placed in a thermocycler programmed with the following: Lid pre-heated to 100^{o}C, 37^{o}C for 30 minutes, 70^{o}C for 5 minutes, and hold at 4^{o}C. When the thermocycler reached 4^{o}C the plate was removed to proceed immediately to the next stage of the kit.

2.1.8.3 Ligate Adapters

Stop Ligation Buffer (STL) and selected RNA adapters provided in the kit were centrifuged at 600xg for 5 seconds. The seal was removed from the plate and 2.5µl of RSB was added to each sample well. Immediately prior to use the Ligation Mix was removed from -20^{o}C storage and 2.5µl was added to each sample well. The Ligation Mix was then immediately replaced to -20^{o}C storage. Next 2.5µl of the selected RNA adapter was added to its corresponding sample (the sample number and its adapter were recorded during this step) and gently mixed by pipetting up and down 10 times. The plate was then sealed with an adhesive seal and then subjected to centrifugation at 280xg for 1 minute. Next, the plate was placed on a preheated thermocycler at 30^{o}C (lid set to 100^{o}C) for 10 minutes. During

this incubation, the plate was removed from the thermocycler and the seal was removed from the plate. To each sample well 5µl of STL was added and gently mixed by pipetting up and down 10 times. To this 42.5µl of well-mixed (via a vortex) AMPure XP Beads were added, and gently mixed by pipetting up and down 10 times. The plate was incubated at room temperature for 15 minutes. Next, the plate was placed on the magnetic stand for another 5 minutes. Following this, 80µl of supernatant was removed and discarded from each sample well. With the plate still on the magnetic stand 200µl of 80% (v:v) ethanol was added, left for 30 seconds before being removed and discarded. This wash step was repeated for a total of two ethanol washes. The plate was then allowed to air dry at room temperature for 15 minutes and then removed from the magnetic stand. The pellet was resuspended with 52.5µl RSB and gently mixed by pipetting up and down 10 times. The plate was incubated at room temperature for 2 minutes before being returned to the magnetic stand for 5 minutes. Next, 50µl of supernatant was removed and transferred to the next row of the plate, which was then removed from the magnetic stand. To this 50µl of well-mixed (via a vortex) AMPure XP Beads were added, and gently mixed by pipetting up and down 10 times. The plate was incubated at room temperature for 15 minutes. Next, the plate was placed on the magnetic stand for another 5 minutes. Following this, 95µl of supernatant was removed and discarded from each sample well. With the plate still on the magnetic stand 200µl of 80% (v:v) ethanol was added, left for 30 seconds before being removed and discarded. This wash step was repeated for a total of two ethanol washes. The plate was then allowed to air dry at room temperature for 15 minutes and then removed from the magnetic stand. The pellet in each well was resuspended with 22.5µl RSB, and gently mixed by pipetting up and down 10 times. The plate was incubated at room temperature for 2 minutes. The plate was then replaced on the magnetic stand and left for 5 minutes. After this, 20µl of clear supernatant was removed and transferred to the next row of the plate. If not immediately progressing to the next step the plate could be sealed with an adhesive seal and stored at -20^{o}C overnight.

2.1.8.4 Enrich DNA Fragments

Stored microtiter plates with processed samples were centrifuged at 280xg for 1 minute to consolidate the starting material. The adhesive seal was then removed. To each sample well

5µl of thawed PCR Primer Cocktail (PPC) and 25µl of thawed PCR Master Mix (PMM) were added and gently mixed by pipetting. An adhesive seal was applied to the plate and placed in a thermocycler programmed as follows: Lid pre-heated to 100°C, 98°C for 30 seconds, 18 cycles of 98°C for 10 seconds then 60°C for 30 seconds then 72°C for 30 seconds, 72°C for 5 minutes, hold at 4°C. The plate was then removed from the thermocycler and the seal removed. To each sample well, 50µl of AMPure XP Beads were added and gently mixed by pipetting. The plate was incubated at room temperature for 15 minutes and transferred to a magnetic stand for an additional 5 minutes. Following this, 95µl of supernatant was removed and discarded from each sample well. With the plate still on the magnetic stand, samples were briefly washed with two exchanges of 200µl of 80% (v:v) ethanol. The plate was then air-dried at room temperature for 15 minutes and removed from the magnetic stand. The pellet was resuspended with 17.5µl RSB and gently mixed by pipetting. The plate was incubated at room temperature for 2 minutes before being returned to the magnetic stand for 5 minutes. Finally, 15µl of supernatant was removed and transferred to the next row of the plate, which was then removed from the magnetic stand. If not immediately progressing to the next step the plate could be sealed with an adhesive seal and stored at -20°C overnight.

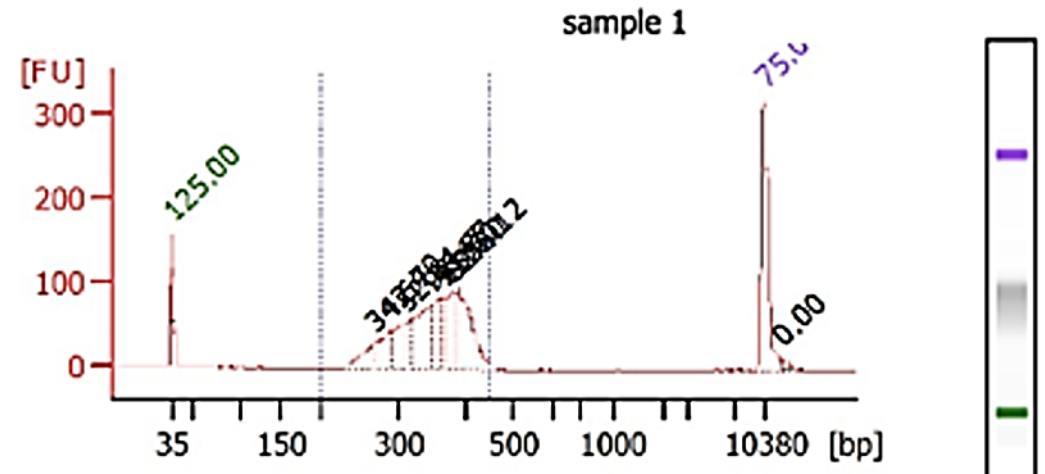
2.1.8.5 Size Selection

This stage of library preparation serves to improve the overall quality of the final sequencing by removing DNA fragments too small or too large which may waste sequencing capacity. Small DNA fragments in particular can harm data quality. These small fragments are typically comprised of adapter-dimers and primer-dimers. Stored microtiter plates were centrifuged at 280xg for 1 minute and the adhesive seal was removed. The Sage Science marker mix and loading solution were removed from 4°C storage and brought to room temperature. To the samples, 15µl of TE buffer (**Table 8**) and 10µl of marker mix were added. The samples were mixed with a vortex and briefly centrifuged for collection. Size selection was performed using a Sage Science BluePippin machine and the proprietary 2% (w:v) agarose gel cassette DF Marker V1 or V2 (Sage Science) DNA fragments with a size range of 200-400 bp were selected. Loading and retrieval of the samples were done in accordance with the

manufacturer's protocol. If not immediately progressing to the next step samples were stored at -20°C.

2.1.8.6 Library Validation, Quantification and Pooling

When preparing the pool of libraries for sequencing it was necessary to quantify the libraries so that the amount of DNA from each library was equal in the final pool. The DNA concentration of the libraries was assessed with an Invitrogen Qubit Fluorometer and Qubit dsDNA HS Assay Kit. Standards (from the kit) were run each time the Qubit was used in accordance with kit instructions. When running the libraries 1µl of sample was added to 199µl of Qubit working solution. To establish the average size (in base pairs) of DNA strands in a library an Agilent 2100 Bioanalyzer System and Agilent Technologies High Sensitivity DNA Chips and DNA Reagents were used in accordance with the manufacturer's instructions. The DNA concentration obtained *via* the Qubit Fluorometer, and the average size of DNA strands obtained *via* the Bioanalyzer were used to calculate library molarity. An example bioanalyzer trace is shown below. The volume of the pool was 100µl and the desired molarity was ~2.5nM, thus the libraries were pooled to achieve this volume and molarity with any difference in volume being made up with RSB (TruSeq ChIP Kit).



Overall Results for sample 1 : **sample 1**

Number of peaks found:	9
Noise:	0.5
Corr. Area 1:	1,223.0

Region table for sample 1 : **sample 1**

From [bp]	To [bp]	Corr. Area	% of Total	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/µl]	Molarity [pmol/l]	Color
200	450	1,223.0	97	347	13.5	459.31	2,055.4	

Figure 2.1: Example Bioanalyzer output trace. In this case the trace is from ChIP-seq DNA library preparation for the Day 3 *Wt* STAT1 sample

2.1.9 Assay for Transposase Accessible Chromatin and Sequencing (ATAC-seq)

2.1.9.1 Cell Lysis

Following AIA and subsequent synovial tissue extraction and processing (Sections 2.1.4-2.1.5) the cells were counted and resuspended in serum free-RPMI-1640 to a concentration of 150,000 cells per ml. ATAC-seq was performed with 150,000 cells.
Cells were centrifuged for 5 minutes at 1000xg at 4°C followed by a 2-second "pulse" at full speed. The pellet was then resuspended in 100µl ice-cold PBS. The cells were again subjected to centrifugation for 5 minutes at 1000xg at 4°C followed by a 2-second "pulse" at full speed. The supernatant was removed and discarded, and the cells were resuspended in cold ATAC Lysis Buffer (**Table 9**) before incubation on ice for 10 minutes and centrifugation (as before). The supernatant was discarded, and the samples processed for the transposition reaction.

2.1.9.2 Transposition Reaction and Purification

The transposition reaction utilized the Illumina Tagment DNA Enzyme and Buffer Kit. Samples were resuspended in 50µl of Transposition mix, transferred to a 0.2ml tube and incubated at 37°C for 30 minutes in a preheated thermocycler. Every 10 minutes the tubes were removed from the thermocycler and briefly vortexed to increase fragment yield. Immediately following this incubation, the samples were purified using a Qiagen MinElute PCR Purification kit. The DNA was extracted in 10µl of elution buffer. Purified DNA was stored at -20°C to await further processing.

2.1.9.3 PCR Amplification

After defrosting the samples on ice (if frozen at the end of the last step) 37.5µl of the PCR master mix was added. To the samples and master mix 2.5µl of a Barcoded PCR Primer (25 µM) was added individually. The barcode primer added to each sample was recorded immediately. Samples were then placed in a thermocycler, programmed as follows; 5 minutes at 72°C then 30 seconds at 98°C followed by 5 cycles of 10 seconds at 98°C, 30 seconds at 63°C then 1 minute at 72°C, and finally hold at 4°C. The samples were then run again with the thermocycler programmed as follows 30 seconds at 98°C followed by 15

cycles of 10 seconds at 98°C, 30 seconds at 63°C then 1 minute at 72°C, and finally hold at 4°C. Following this reaction, the samples were purified using a Qiagen MinElute PCR Purification kit. The DNA was eluted in 20µl of the elution buffer. The purified DNA could be stored at -20°C to await library quantification and pooling.

2.1.9.4 ATAC- seq Library Validation, Quantification and Pooling

ATAC-seq libraries were validated, quantified, and pooled in the same manner described in section 2.1.8.6 save that the buffer used for dilutions when compiling the pool was the Tagment DNA buffer.

2.1.10 Next Generation Sequencing (NGS) and Sequencing Data Normalisation Workflow

Sequencing was carried out using an Illumina NovaSeq 6000 sequencing system and NovaSeq 6000 S1 Reagent Kits (200 cycles). The raw data of both ATAC-seq and ChIP-seq was normalised using the Supercomputing Wales High-performance computing system Hawk located at Cardiff University. The code used in the following analyses can be found in the **Appendix**.

2.1.10.1 ChIP-seq Workflow

The ChIP-seq data was processed in accordance with the ENCODE guidelines (ENCODE Guidelines for Experiments Generating ChIP-seq Data. 2017). The workflow of the normalisation pipeline was as follows. First, the raw FASTQ sequencing files were trimmed and then quality controlled (QCed). The reads were mapped to the *Mus musculus* GRCm38 reference genome sequence. Duplicate reads were marked and again QCed. Sequencing peaks were then called with the Model-based Analysis of ChIP-seq (MACS) 2 algorithm and the sequences under the peaks were obtained. Finally, the peak-gene overlaps were obtained utilising bedtools intersect for transcription start sites (TSS), introns, exons and stop codons according to the aforementioned reference genome.

2.1.10.2 ATAC-seq Workflow

The normalisation pathway for the ATAC-seq data was primarily based on the paper *'ATAC-seq normalization method can significantly affect differential accessibility analysis and interoperation'* (Reske et al. 2020) with some minor modifications. The same reference genome: *Mus musculus* GRCm38 was used. The stages of the pipeline were as follows. The raw FASTQ sequencing files were trimmed. Both the trimmed and the raw files were then QCed. These QCed files were then combined. The sequences were then aligned to the reference genome with the Burrows-Wheeler Aligner. This is a different aligning tool from the one used in the paper (bowtie2) but no less effective. Next, there was a coordinate, sort, and index step before ENCODE blacklist regions are removed. The next four steps were added to the workflow of the paper. They consisted of removing mitochondrial reads, retaining only proper pairs, calculating duplication frequency, estimating library complexity then finally subsampling binary alignment map (BAM) files so that all have the same number of reads. Subsequently, PCR duplicates were removed, sorted, and indexed. Then BAM was converted to browser extensible data paired-end (BEDPE) format, Tn5 shift and minimal conversion were added to the papers workflow. The final steps were MACS2 peak calling, sequences under peaks and peak-gene overlaps were obtained for transcription start sites (TSS), introns, exons and stop codons according to the aforementioned reference genome (in the same fashion as with ChIP-seq data).

2.2 Ingenuity Pathway Analysis

To assess the biological significance of the genes identified in ATAC-seq and ChIP-seq experiments, QIAGEN's Ingenuity Pathway Analysis (IPA) was used. Gene sets (described in the results chapters) were initially subjected to IPA's 'Core Analysis-Expression Analysis' using the default settings. Once these gene lists had undergone this core analysis various 'comparison analysis' permutations were conducted (also described in the results chapters). These too were conducted using the default settings. Results from these experiments were exported from IPA as .xls files and GraphPad Prism was used to create the various figures shown in the results chapters.

2.3 Materials

2.3.1 Composition of Buffers and Reagents not Ready Made

Table 1: Supplemented Media

Ingredient	Volume
RPMI-1640	500ml
Heat inactivated FCS	50ml
L-glutamine (2mM)	5ml
Penicillin (100µg/ml)	2.5ml
Streptomycin (100µg/ml)	2.5ml
β-mercaptoethanol (55nM)	0.5ml

Table 2: Magnetic-activated cell sorting (MACS) buffer

Ingredient	Volume
PBS	500ml
BSA	2.5ml
EDTA (0.5M)	5ml

Table 3: Cell Lysis Buffer (CLB)

Ingredient	Volume
Ultra-pure Water	9.6 ml
1M Tris-HCl pH 8.0	100 µl
1M NaCl	100 µl
70% v/v NP40	28.57 µl
1M Sodium Butyrate	100 µl
1M PMSF	100 µl
10 mg/ml Leupeptin	10 µl

Table 4: Nuclear Lysis Buffer (NLB)

Ingredient	Volume
Ultra-pure Water	8.1 ml
1M Tris-HCl pH 8.0	500 µl
0.5M EDTA	200 µl
10% w/v SDS	1 ml
1M Sodium Butyrate	100 µl
1M PMSF	100 µl
10 mg/ml Leupeptin	10 µl

Table 5: Immunoprecipitation Dilution Buffer (IPDB)

Ingredient	Volume
Ultra-pure Water	7.9 ml
1M Tris-HCl pH 8.0	200 µl
1M NaCl	1.5 ml
0.5M EDTA	40 µl
100% v/v Triton X-100	100 µl
10% w/v SDS	10 µl
1M sodium butyrate	100 µl
1M PMSF	100 µl
10 mg/ml Leupeptin	10 µl

Table 6: Immunoprecipitation Wash Buffer 1 (IPWB1)

Ingredient	Volume
Ultra-pure Water	45.3 ml
1M Tris-HCl pH 8.0	1 ml
1M NaCl	2.5 ml
0.5M EDTA	200 µl
100% v/v Triton X-100	500 µl
10% w/v SDS	500 µl

Table 7: Immunoprecipitation Wash Buffer 2 (IPWB2)

Ingredient	Volume
Ultra-pure Water	43.7 ml
1M Tris-HCl pH 8.0	500 µl
5M LiCL	2.5 ml
0.5M EDTA	200 µl
70% v/v NP40	714.29 µl
Sodium deoxycholate	500 mg

Table 8: Tris-EDTA (TE) Buffer

Ingredient	Volume
Water	100 ml
1M Tris-HCl pH 8.0	1 ml
0.5M EDTA	200 µl

Table 9: ATAC Lysis Buffer

Ingredient	**Volume**
Water	9.84 ml
1M Tris-HCl pH 7.4	100 µl
5M NaCl	20 µl
1M $MgCl_2$	30 µl
100% v/v Igepal CA-630	10 µl

Table 10: Transposition Reaction Mix

Ingredient	**Volume** (Per Sample)
Nuclease-free Water	22.5 µl
2x Reaction Buffer	25 µl
Tn5	2.5 µl

Table 11: ATAC PCR Master Mix

Ingredient	**Volume** (Per Sample)
Nuclease-free Water	10 µl
PCR Primer (Ad1_noMX) (25 µM)	2.5 µl
NEBNext High-Fidelity 2x PCR Master mix	25 µl

2.2.2 Additional Chromatin Immunoprecipitation Materials

Table 12: ChIP Antibodies

Target	Antibody	Volume added to Sample
STAT1	Stat1 Antibody Cell Signaling #9172	25 µl
STAT3	SANTA CRUZ Stat3 (C-20): sc-482	25 µl
P300	MERCK Anti-P300 clone RW128	5 µl
IgG	Anti-rabbit IgG, HRP-linked Antibody Cell Signaling #7074	5 µl

Table 13: qPCR Primers used in ChIP Validation

	Forward (5'-3')	Target STAT motif.	Reverse (5'-3')
IRF1Promoter	CCTTCGCCGCTTAGCTCTAC	ACAGCCTGATTTCC	CCCACTCGGCCTCATCATT
IRF1 Down Stream	GCCTTGGCGTGACTCTTGAC	ATCTATTAGAAACGCCACCTAA	ACATGACCAAACACCATTTAGCA
SOCS3 Promotor	CTCCGCGCACAGCCTTT	TGCAGAGTAGTGACTAAA	CCGGCCGGTCTTCTTGT
STAT3 Promotor	CAGCAGGACATTCCGCTAAT	TTTGTAAGCTAGGCCTCTGCGCG	ACAAAGCTCTCAGAACAGCC

Table 14: Illumina TruSeq ChIP Kit (Part number: 15023092)

Reagent	**Preparation**
Resuspension Buffer	Thawed from -20°C storage at room temperature. After 1st use stored at 4°C but brought to room temperature for subsequent uses.
End Repair Mix	Thawed from -20°C storage at room temperature. Returned to -20°C storage after use.
A-Tailing Mix	Thawed from -20°C storage at room temperature. Returned to -20°C storage after use.
Ligation Mix	Removed from -20°C storage immediately before use. Returned to -20°C storage immediately after use.
Stop Ligation Buffer	Thawed from -20°C storage at room temperature. Returned to -20°C storage after use.
RNA Adapters	Thawed from -20°C storage at room temperature. Returned to -20°C storage after use.
PCR Master Mix	Thawed from -20°C storage at room temperature. Returned to -20°C storage after use.
AMPure XP Beads	Stored at 4°C, removed at least 30 minutes before use and mixed via a vortex very thoroughly to ensure proper distribution of beads.
80% (v:v) ethanol	Prepared immediately before use.

Chapter 3:

Tracking chromatin accessibility as a response to synovitis in murine antigen-induced arthritis.

3.1 Tracking Epigenetic Mechanisms of Disease Heterogeneity

Prior RNA-seq analysis had identified significant differences in the transcriptional output of synovial tissues from *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice with AIA (Hill, 2019). However, the precise mechanisms by which these changes in gene regulation occur remain unknown. As discussed in **Chapter 1**, gene regulation is controlled at multiple levels, and includes molecular processes affecting chromatin accessibility (Klemm et al. 2019). In this Chapter, I investigate changes in chromatin accessibility as a response to arthritis induction and consider how these events may shape disease heterogeneity in synovitis.

To coordinate the regulation of transcription factor binding to gene promoters and enhancers the chromatin architecture has to be remodelled to allow access for transcriptional regulators and suppressors. Advances in next-generation sequencing now enable a genome-wide appraisal of these modifications and methods such as assay for chromatin transposase-accessible chromatin combined with whole genome sequencing (ATAC-seq) provide a fast and sensitive means to study the epigenome (Reske et al. 2020; Grandi et al. 2022). In this Chapter, I have applied ATAC-seq to understand the epigenetic mechanisms that affect transcriptional decisions in synovitis.

As discussed in **Chapter 1**, patients with RA often display differing efficacies of treatment to certain classes of biological medications (biologics). These differences are likely a reflection of the heterogeneous nature of the disease. Furthermore, difficulties in establishing an efficacious treatment regime for patients negatively impact the quality of life(McInnes and Gravallese 2021). Hence the importance of developing a greater understanding of the various mechanisms which contribute to said heterogeneity. In this chapter one such mechanism, chromatin accessibility, is examined.

The variance in the transcriptome and subsequent variance in disease between the pathotypes observed in *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice with AIA may arise through changes in chromatin accessibility. In this regard, ATAC-seq offers the opportunity to quantify chromatin accessibility via the insertion of sequencing adapters at open genomic loci by mutant transposase enzymes (see **Figure 3.1**).

ATAC-seq has been used extensively to characterise the chromatin accessibility profiles for various cancers. This form of research has identified proximal and distal regulatory elements involved in cancer pathogenesis. Such data has improved the

classification and stratification of tumours, and provides insights into treatment responses, patient outcomes, and potential novel treatments. By assessing the accessible genome holistically, as opposed to the transcriptome only, light is shed on non-coding elements and their potential for improved diagnosis and treatment(Corces et al. 2018).

This chapter discusses the optimisation and high-level assessment of ATAC-seq data from synovial tissues extracted from *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice with AIA. The analysis explored changes in the chromatin accessibility landscape as a response to arthritis induction and differences arising as a response to cytokine receptor deficiency.

3.2 Introduction to ATAC-seq

The materials and methods for ATAC-seq are described in **Figure 3.1** and fully outlined in **Chapter 2 (Section 2.1.9)**. ATAC-seq was conducted at baseline and Day 3 and Day 10 of AIA in *Wt* mice. Parallel studies were performed at Days 3 and 10 of AIA in *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice. Sequencing was performed with a NovaSeq 6000 S1 Reagent Kit v.1.5 (200 cycles).

3.2.1 Generation of ATAC-seq Gene Lists

To enable bioinformatic comparisons between each test condition, it was necessary to generate a single gene list for each mouse genotype. Sequencing peaks from each replicate were, therefore, pooled and ranked according to peak fold enrichment. Next, the number of genes associated peaks for each replicate was determined and the mean for each condition was calculated. This number of peaks from the pooled and ranked list would become the cut-off. Any peaks with a peak fold enrichment score below this threshold were removed, together with any duplicates, creating a final dataset for analysis.

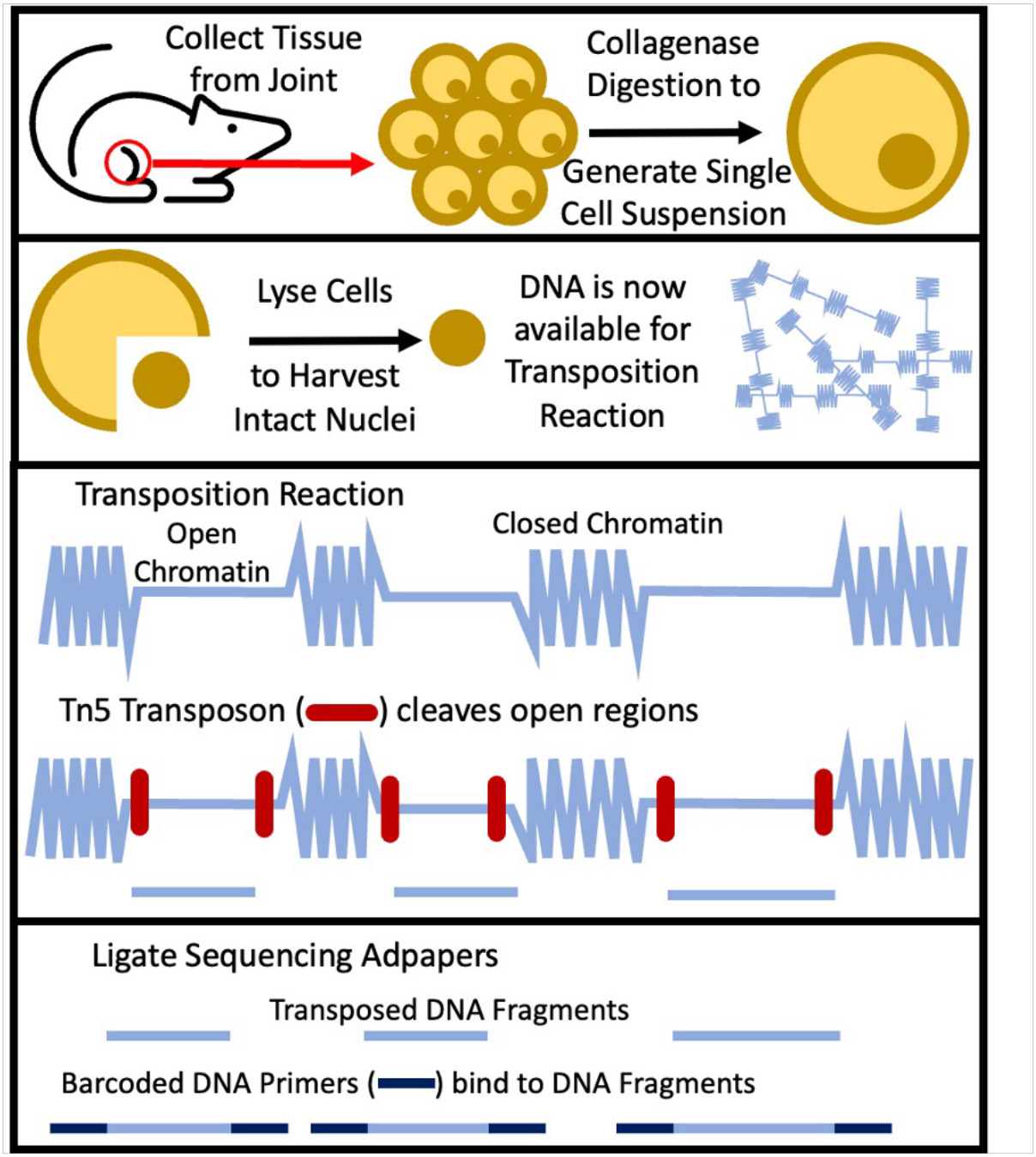


Figure 3.1: Summary of ATAC-seq Murine Synovial Tissue. Tissue is collected from the knee and subjected to a collagenase digestion. Recovered cells were lysed to isolate the nuclei. During the transposition reaction the mutant Tn5 Transposon cleaves the open chromatin regions. To these transposed DNA fragments barcoded DNA primers and ligated and libraries generated for NGS.

3.3. Results

3.3.1 Rationalisation of *Il6ra*$^{-/-}$ ATAC-seq Data

The normalised *Il6ra*$^{-/-}$ ATAC-seq dataset contained many thousand more peaks than *Wt* or *Il27ra*$^{-/-}$ samples (the reasons for this will be discussed in **Chapter 6**). To control for this difference, it was necessary to remove less significant peaks from the *Il6ra*$^{-/-}$ datasets. The peaks were, therefore, ranked according to peak fold enrichment, a measure of statistical significance. The mean number of peaks for all other samples was 20,000 (rounded to the nearest 10,000. Thus, the peak fold enrichment value of the 20,000th peak was used as a cut-off score in the *Il6ra*$^{-/-}$ data sets. Any sequencing mark with a peak fold enrichment lower than that of the 20,000th ranked peak was removed.

3.3.2 Distribution of Open Chromatin Regions Across Chromosomes

ATAC-seq generates large amounts of raw sequencing data, which required a dedicated bioinformatic pipeline providing consistency in the processing and normalisation (described in **Chapter 2**). This included the mapping of sequencing reads to individual chromosomes to ensure all murine chromosomes (n=20, including the X and Y chromosomes) are represented within the output file. On occasions, when information for specific chromosomes was missing, the R-script encoding the bioinformatic pipeline was reviewed and edited to ensure complete coverage of the genome. This methodological approach ensured the quality and integrity of each dataset.

The histological features and transcriptional profile of synovial pathotypes are distinct (Pitzalis et al. 2013). We, therefore, hypothesised that the pathologies displayed by *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice with AIA would possess a series of common and distinct epigenetic features across the course of the disease. These would reflect cellular processes consistent with the development of synovitis and more specific changes in chromatin accessibility at genomic loci important for disease heterogeneity. To understand these

relationships, bioinformatic tools were applied to identify signatures common to arthritis progression and more specific changes linked with differences in synovitis.

Figure 3.2 shows the percentage of sequencing peaks aligned to each chromosome in *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice. The percentage shown for a given chromosome is the mean percentage across replicates. The figures also include a subset of sequencing peaks corresponding to annotated promoters for specific genes (termed gene-associated peaks; GAPs). Likewise, these are the mean percentages of GAPs only across replicates. Chromosomes 1-19, X and Y are represented in addition to peaks not mapped to a particular chromosome: non-associated (NA).

Initial inspection of the data presented in **Figure 3.2** shows a broad similarity in the distribution of peaks across the chromosomes between conditions. For instance, in all conditions (including GAPs) chromosomes 7 and 11 hold a greater share of peaks than 8,9 and 10. It is noteworthy that in all conditions the percentage of peaks which are non-associated with a chromosome is ~0 when only GAPs are considered. This reduction is to be expected as the chromosome on which the vast majority of identified genes are on is known.

3.3.2.1 Identification of Chromosomes of Interest

Building upon the aforementioned observations regarding the similarity of peak distribution across chromosomes, statistical comparison analysis was performed to determine if the percentage of peaks on a given chromosome varied significantly between differing genotypes and timepoints. Pairwise comparisons were carried out between all conditions using the Wilcoxon Ranksum Test, with correction for multiple hypothesis testing using the Benjamini and Hochberg method available in MATLAB_R2022a. Chromosomal differences were deemed significant with a false discovery rate <0.05.

Below, **Figure 3.3** shows the chromosomes, of each condition, displaying open chromatin accessibility in which there is a significant difference (between experimental conditions) in the percentage of peaks. In five chromosomes (3, 11, 13, 16 and 17) there are differences between some conditions when all peaks are taken into consideration. When only GAPs are considered chromosomes 3, 5, 11, 13 and 17 show significant differences between some (notably identical) conditions.

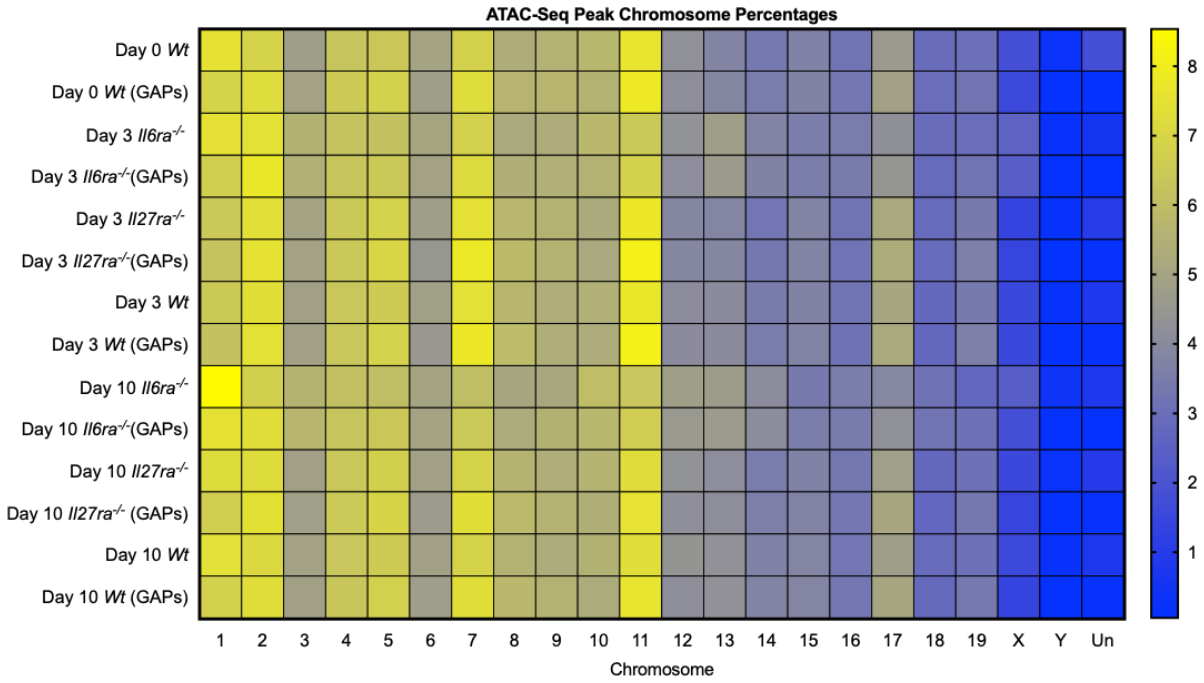


Figure 3.2: The Percentage of Peaks Across a Particular chromosome Across ATAC-Seq Conditions. The mean percentage of peaks on a given chromosome (1-19, X, Y and Unassigned (Un)) was calculated from each condition's replicates. No chromosome exceeded a 10% share of the peaks, as such the gradient of the heatmap is delineated accordingly. Both total peak and gene associated peak (GAPs) percentages are shown in the figure.

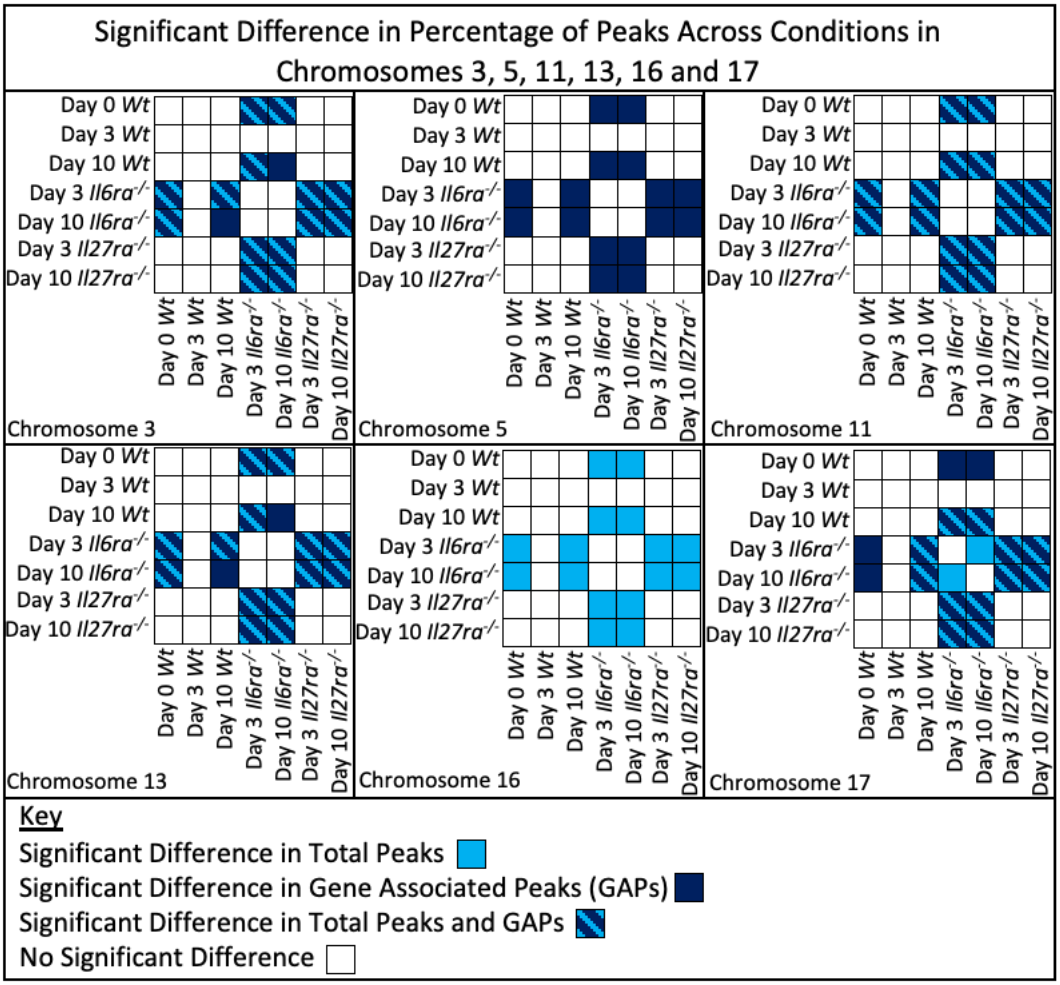


Figure 3.3: The Chromosomes Identified as Having a Significant Difference in the Percentage of Peaks Attributed to Them and the Specific Conditions Where this Difference Occurs. The matrices illustrate the significantly different conditions. In chromosomes not mentioned there are no significant differences between any conditions. The light blue squares represent significant difference in the total peaks. The dark blue squares represent significant difference in the gene associated peaks (GAPs) only. The striped blue squares represent significant difference in the total peaks and GAPs. The white squares indicate no significant difference.

3.3.3 Distribution of Open Chromatin Across Genomic Regions

As ATAC-seq peaks were mapped to chromosomes, so too were they mapped to genomic regions. These being Proximal (≤200 bp from a transcription start site; TSS), Distal (<200 bp from a TSS and comprising introns, exons and stop codons) and Intergenic regions (these not being assigned to a specific gene). The percentage of peaks associated with these regions was calculated for each sample and the mean of those percentages was calculated for each replicate. The distribution of peaks across genomic regions is shown in **Figure 3.4**.

As mentioned previously, the *Il6ra*$^{-/-}$ data required the removal of sequencing peaks with a low significance. The similarity in the distribution of peaks across genomic regions between conditions is of particular reassurance regarding the *Il6ra*$^{-/-}$ data and its integrity and validity. Even after the removal of excess peaks, the distribution of the remaining peaks across regions is consistent with *Wt* and *Il27ra*$^{-/-}$ data.

Figure 3.4 shows that ~60% of the sequencing peaks are associated with proximal regions, whereas the remaining proportion is equally distributed between Distal and intergenic. This means that ~80% of peaks are associated with a gene, a strong indication of high-quality data. Furthermore, the accessible genome encompasses more than 90% of regions to which transcription factors bind (Thurman et al. 2012) and as **Figure 3.4** indicates the majority of peaks are in proximal regions, further validating the data.

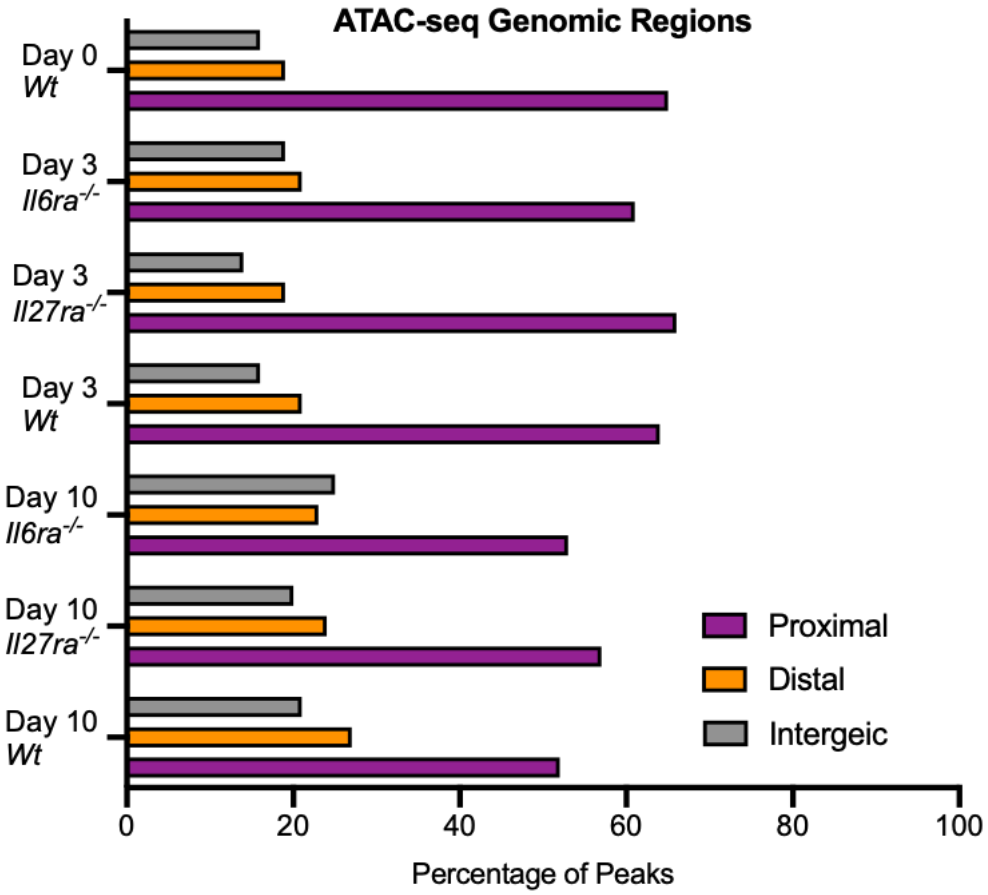


Figure 3.4: The Distribution of ATAC-seq Peaks Across Proximal, Distal and Intergenic Regions. Proximal peaks are defined as being ≤200 bp distant of a TSS. Distal peaks are defined as being >200bp distant of a TSS. They are comprised of Introns, exons and stop codons. Intergenic peaks are not associated with any gene. The mean percentage of peaks associated with a given region was calculated from each condition's replicates.

3.3.4 Accessible Gene Commonality Between Experimental Conditions

The analysis presented above has identified a broad similarity of chromatin accessibility across each of the experimental conditions, as illustrated particularly well by **Figure 3.2**. To assess this further the gene lists presented in **Section 3.2.2** were utilised. Investigations used an open-access Venn diagram tool (https://bioinformatics.psb.ugent.be/webtools/Venn/), to determine which and what proportion of genes (and by extension open chromatin loci) are shared between genetic mouse strains and thus examine the similarity (or lack of) of chromatin accessibility in synovitis. As this analysis concerns only accessible genes it ignores ~20% of peaks mapped to intergenic regions. However, this does not detract from the usefulness of the analysis.

3.3.4.1 Gene Commonality Across Genotypes

The pie charts in **Figure 3.5, Figure 3.6, and Figure 3.7** display the proportion of genes unique to a specific genotype and those shared with at least one other genotype. The left panels show the entire dataset, while the right panels show only the genomic loci identified following disease onset. The genes that are accessible in the non-disease state (Day 0 *Wt* condition), which therefore relate to homeostatic processes, were removed (utilising the Venn tool mentioned previously) from the gene lists of the other conditions. This approach identified a set of gene promoters that are only accessible in the disease state and allowed for an examination of the cytokine-induced changes to the chromatin landscape that influence disease progression and severity.

3.3.4.2 Gene Commonality Across Timepoints

By utilising the same method to compare the chromatin accessibility landscape of *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice it is possible to examine changes in chromatin accessibility across the progression of AIA by comparing Days 0, 3 and 10 of one genotype. In all disease progression comparisons, the Day 0 condition is represented by the *Wt* genotype. These comparisons are shown in **Figure 3.8**.

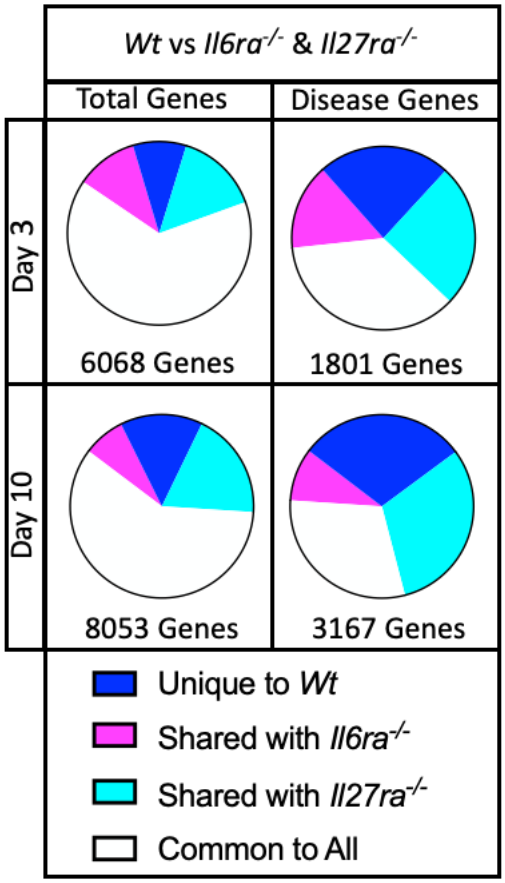


Figure 3.5: The Proportion of Genes Accessible in the *Wt* Condition Verses Other Genotypes. The genes accessible in the *Wt* genotype at Days 3 (top) and 10 (bottom) were compared with the *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ to determine which are uniquely accessible in the *Wt* genotype, which are accessible in both the *Wt* genotype and either *Il6ra*$^{-/-}$ or *Il27ra*$^{-/-}$ and which genes are accessible in all genotypes. The proportion of genes in these four categories in relation to the *Wt* genotype are shown in pie charts. This analysis was carried out in all genes (left) and a disease subset (right) in which homeostatic genes (Day 0 *Wt*) had been removed from the total accessible gene list.

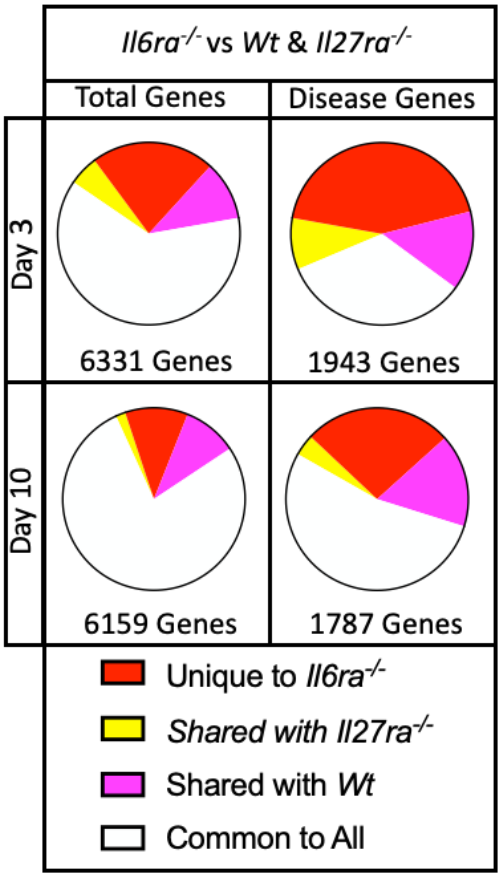


Figure 3.6: The Proportion of Genes Accessible in the *Il6ra*$^{-/-}$ Condition Verses Other Genotypes. The genes accessible in the *Il6ra*$^{-/-}$ genotype at Days 3 (top) and 10 (bottom) were compared with the *Wt* and *Il27ra*$^{-/-}$ to determine which are uniquely accessible in the *Il6ra*$^{-/-}$genotype, which are accessible in both the *Il6ra*$^{-/-}$ genotype and either *Wt* or *Il27ra*$^{-/-}$ and which genes are accessible in all genotypes. The proportion of genes in these four categories in relation to the *Il6ra*$^{-/-}$ genotype are shown in pie charts. This analysis was carried out in all genes (left) and a disease subset (right) in which homeostatic genes (Day 0 *Wt*) had been removed from the total accessible gene list.

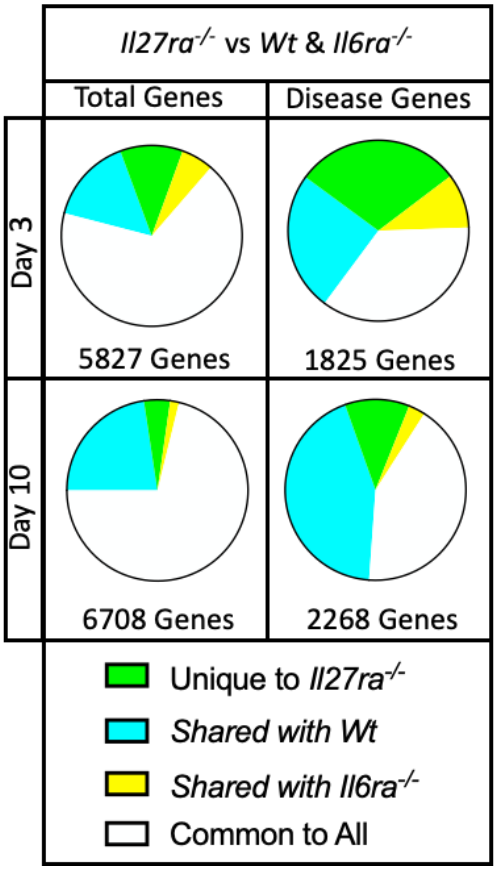


Figure 3.7: The Proportion of Genes Accessible in the *Il27ra*$^{-/-}$ Condition Verses Other Genotypes. The genes accessible in the *Il27ra*$^{-/-}$ genotype at Days 3 (top) and 10 (bottom) were compared with the *Wt* and *Il6ra*$^{-/-}$ to determine which are uniquely accessible in the *Il27ra*$^{-/-}$genotype, which are accessible in both the *Il27ra*$^{-/-}$ genotype and either *Wt* or *Il6ra*$^{-/-}$ and which genes are accessible in all genotypes. The proportion of genes in these four categories in relation to the *Il27ra*$^{-/-}$ genotype are shown in pie charts. This analysis was carried out in all genes (left) and a disease subset (right) in which homeostatic genes (Day 0 *Wt*) had been removed from the total accessible gene list.

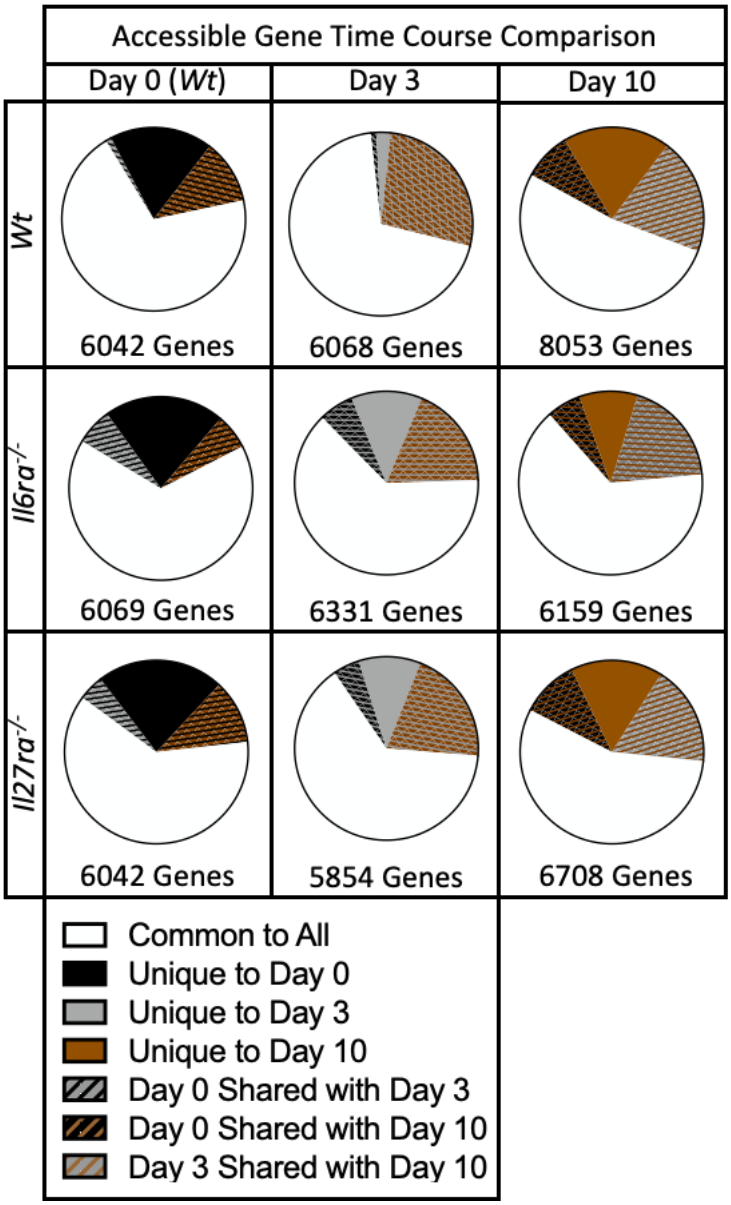


Figure 3.8: Comparisons of Gene Accessibility Through the AIA Time Course in All Genotypes. The genes accessible at a given timepoint were compared with the other two points. This determines which genes are uniquely accessible at that point, shared with one or other of the other timepoints and those which are accessible throughout the time course. The proportion of genes in these four categories in relation to each condition are shown in pie charts. From top to bottom shows the *Wt,* *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ genotypes. From left to right the Day 0, Day 3 and Day 10 timepoints.

3.4 Discussion

A principle aim of this chapter was to systematically assess the ATAC-seq data produced, to contextualise it for further analysis, described in the next chapter. This involved an amalgamation of replicates for ease of assessment and comparison between conditions. Comparisons were distinct and biologically relevant; chromosome and genomic region distribution of accessible genes (and by extension open regions) and a direct comparison of accessible genes between conditions. The major finding of these efforts was that regardless of genotype and timepoint open chromatin regions are predominantly consistent.

3.4.1 ATAC-seq peak distribution indicates conserved accessibility.

Analysis based on open chromatin regions across chromosomes revealed that in both the total peaks and GAPs only six out of 21 chromosomes (22 including unassigned regions) show significant variance in the number of peaks aligned to each. Furthermore, as the many white squares in **Figure 3.3** attest these differences in chromosomal distribution of peaks do not occur in every comparison between conditions. This general lack of variation in chromatin accessibility across conditions is a running theme of this ATAC-seq data. It speaks of a conserved chromatin accessibility landscape and the importance of the subtle changes in the landscape in shaping the disease state. It is important to reiterate that accessibility does not equal expression and thus other regulatory mechanisms must be considered to fully understand the inflammatory heterogeneity. This will be expanded upon in subsequent chapters. Likewise, the aforementioned chromosomes of interest will be examined further in the following chapter to establish their role in arthritis.

3.4.2 Accessible gene commonality suggests a pivotal role for genes with differing accessibility between genotypes.

The Total Gene pie charts in **Figure 3.5, Figure 3.6, and Figure 3.7** indicate that there is a high degree of similarity in chromatin accessibility across all conditions, with the majority of genes being accessible in all conditions. Even in the disease genes subset, the common genes between all conditions retain a significant share. Removal of the homeostatic genes resulted in a considerable decrease in the number of genes across all conditions, suggesting

that most genes are accessible regardless of disease state or genotype (i.e. prior to disease onset).

Based on the activities of the individual cytokine receptor cassettes, it was anticipated that the datasets from *Il27ra* and *Il6ra*-deficient mice would have more in common with findings from the *Wt* mice than with each other. This is demonstrated by the larger magenta and cyan sections in **Figure 3.6** and **Figure 3.7**, respectively, compared to the yellow sections. *Il6ra*$^{-/-}$ mice retain IL-27 receptors, and *Il27ra*$^{-/-}$ mice retain IL-6 receptors, indicating that the chromatin accessibility landscape related to these receptors is conserved where they are present. However, the accessibility of genes that are unique to the *Wt* condition requires both IL-27 and IL-6 receptor signalling, suggesting interplay between the cytokines and their downstream elements in the establishment of the normal disease phenotype.

Interestingly, the *Wt* conditions on Day 3 and Day 10 have a greater degree of shared accessibility with their *Il27ra*$^{-/-}$ counterparts than with the *Il6ra*$^{-/-}$ counterparts. This is particularly evident in the Disease Gene subset, where there is a marked increase in the percentage of shared genes between the *Wt* and *Il27ra*$^{-/-}$ conditions compared to the Total Genes subset. *Il6ra*$^{-/-}$ mice have a low-inflammatory disease phenotype, and of the three genotypes, they are the most phenotypically distinct. This is reflected in the greater similarity between the *Wt* and *Il27ra*$^{-/-}$ conditions.

3.4.3 Temporal accessibility changes suggest a role for genes in initiating inflammation and chronicity.

The broad similarity of open chromatin between conditions is repeated in the pie charts of **Figure 3.8** the majority of genes (shown in white) in all conditions are universally accessible. This indicates that regardless of disease state the chromatin accessibility landscape remains relatively constant.

In all genotypes days 3 and 10 share more accessible genes with each other than they do the day 0 baseline. As both days are in the disease state it is unsurprising that chromatin accessibility is more similar in them than in the non-disease state. Furthermore, in the *Wt* and *Il27ra*$^{-/-}$ conditions there is a greater degree of commonality to Day 0 at day 10 than at day 3. This relationship is mirrored at Day 0 where there is more in common with

Day 10 than 3. This is suggestive of the beginnings of a return to the homeostatic baseline. In AIA the peak of inflammation occurs at day 3 before gradually returning to near baseline by day 10.

It is reasonable to speculate that the genes which are only accessible at day 3 exhibit pro-inflammatory functions which are initially increased early in the disease course but subside by the end of the AIA experiment. Genes accessible only at Day 0 whose usual activity (Day 0 and before) is homeostatic but when the inflammatory challenge is presented at Day 0 their function, and accessibility decreases. As their accessibility is seemingly lost, their absence may play a role in RA's chronicity. Genes shared by Day 0 and 10 only are somewhat similar; initially, they become absent during the acute inflammatory phase but by Day 10, as the inflammatory episode subsides their accessibility returns.

Almost the inverse of those genes whose accessibility is lost after day 0 are those who become accessible on day 3 and remain so till day 10. Likewise, genes are uniquely accessible at day 10. These genes could possess a gain of function role in RA's chronicity. It would be interesting to conduct ATAC-seq at timepoints beyond day 10 to determine if the chromatin accessibility landscape returns to the same state as day 0 or if there are permanent changes following an inflammatory episode.

Chapter 4:

Relating chromatin accessibility to the transcriptional control of genes involved in synovial pathology.

4.1 Introduction

Results presented in Chapter 3 describe the optimisation of ATAC-seq and the initial epigenetic analysis of synovitis in mice with antigen-induced arthritis. Interrogating the ATAC-seq datasets generated from the inflamed synovium of *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice at day 3 and day 10 of disease, I now consider the relationship between chromatin accessibility and changes in gene regulation in directing disease heterogeneity.

Functional genomic studies conducted in cancer and immune cells, and tissues from disease settings have shown that alignment of ATAC-seq with RNA-seq supports understanding of the regulatory mechanisms contributing to tumour progression, the effector characteristics of leukocytes and development of inflammation-induced tissue injury (Corces et al. 2018; Xu et al. 2023; Yuan et al. 2023). In this Chapter, I adopt a similar approach to generate novel insights into the cytokine control of biological pathways linked with synovial hyperplasia, leukocyte infiltration and joint damage.

4.2 Results

ATAC-seq identifies genomic regions which are open and accessible to transcriptional regulation (Grandi et al. 2022). To understand how changes in the chromatin architecture affect transcriptional gene regulation, I aligned corresponding bulk RNA-seq datasets from the inflamed synovial tissues of *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ with antigen-induced arthritis (previously generated by Dr David Hills as part of his PhD studies) to the ATAC-seq datasets presented in Chapter-3. Bulk RNA-seq datasets were obtained at Day 0, Day 3 and Day 10 of the disease and allowed a direct comparison with the newly generated ATAC-seq results. Analysis in the previous chapter identified epigenetic marks linked to the development of synovitis. These changes in chromatin accessibility were not identified in the naïve state (*Wt* Day 0) but were regulated as a response to synovitis at day 3 and/or day 10 of the disease. Analysis constructed in this Chapter focussed on the gene regulation associated with these changes in chromatin accessibility (see **Section 3.2.1** and **Section 3.3.4.1**). The coloured segments of the Disease Genes pie charts in **Figures 3.5, 3.6** and **3.7** in **Section 3.3.4.1** comprise genes exhibiting variable accessibility between genotypes. The white segments of those pie charts indicate genes with common accessibility irrespective of genotype.

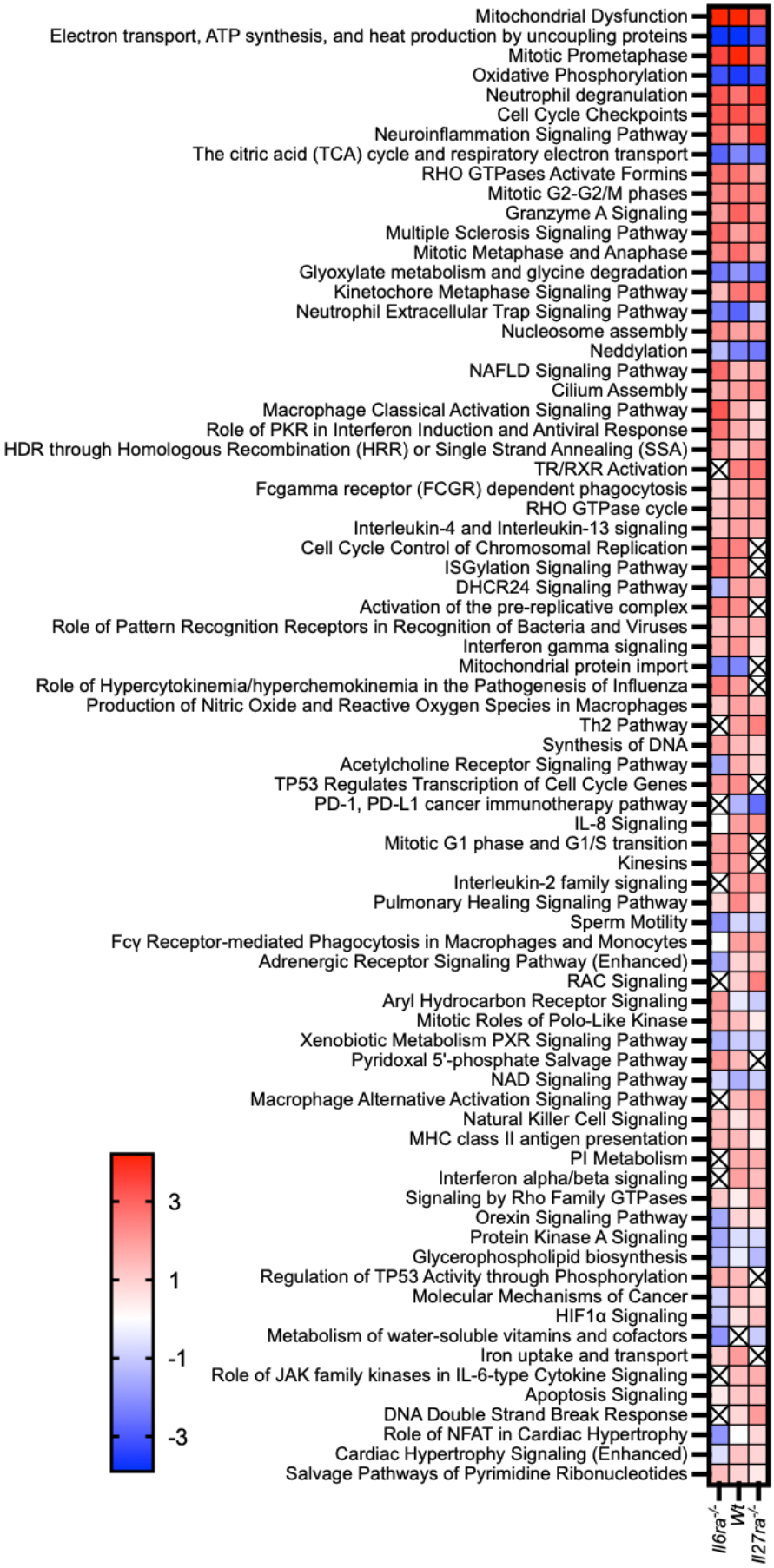
4.2.1 Molecular pathway analysis of RNA transcripts mapped by ATAC-seq.

Bioinformatic analysis using prediction tools in the Ingenuity Pathway Analysis (IPA) software was used to understand how the cytokine control of chromatin accessibility affected biological processes linked to synovitis (**Figure 4.1**, **4.2** and **4.3**). ATAC-seq sequencing peaks were mapped to the promoter regions of annotated genes within the mouse genome. The analysis then considered changes in the relative expression of these gene signatures as a response to arthritis induction in corresponding bulk RNA-seq datasets. Identified gene targets were then subjected to IPA analysis. Results are shown for changes in 'canonical pathways' (**Figure 4.1**), 'diseases and functions' (**Figure 4.2**), and 'upstream regulators' (**Figure 4.3**).

Each condition was first subjected to IPA's 'Core Analysis' using the default settings. These then underwent IPA's 'Comparison Analysis', again using the default settings. The pathways, regulators and diseases identified are ranked by their z-score. Each individual heatmap shows statistics for the top 75 results. Panel **A** of each of these three figures shows the results for Day 3, and panel **B** shows Day 10.

As can be seen in the heatmaps by the numerous differences between genotypes, deletion of the IL-6 or IL-27 receptors and thus inhibition of those cytokines' signalling pathways resulted in changes to gene accessibility, ultimately resulting in changes to the pathophysiology of the synovium. The principle aim of this bioinformatic analysis was to examine how differences between genotypes in the accessible genomic regions, which result from cytokine signalling disruption, contribute to biological differences and ultimately disease phenotype. This aim has been met. The nature of these changes and differences and thus their biological significance is described further in the discussion section of this chapter.

Figure 4.1.A Day 3 ATAC Disease Genes Canonical Pathways



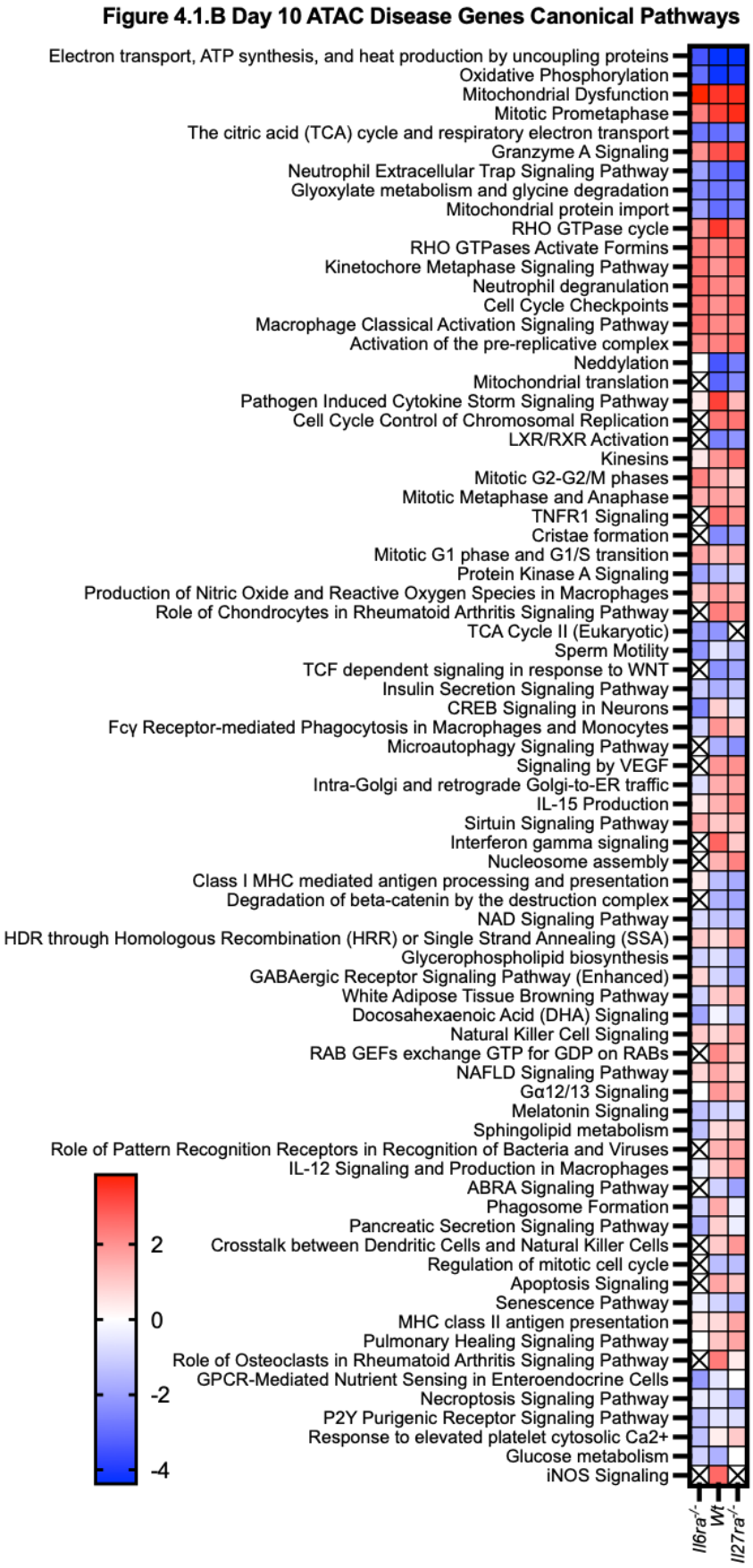


Figure 4.1: Canonical Pathways Identified by IPA Comparative Analysis of Disease Genes Mapped to RNA-seq data. The disease related genes identified by ATAC-seq experiments between genotypes were mapped to RNA-seq data and submitted to IPA for Core then Comparative analysis. Panel **A** shows Day 3 and **B** shows Day 10. The top 75 canonical pathways are exhibited in these heatmaps. They were ranked by their z-score which is displayed in the heatmap. A box marked X indicates a z-score of exactly 0.

Figure 4.2.A Day 3 ATAC Disease Genes Diseases & Functions

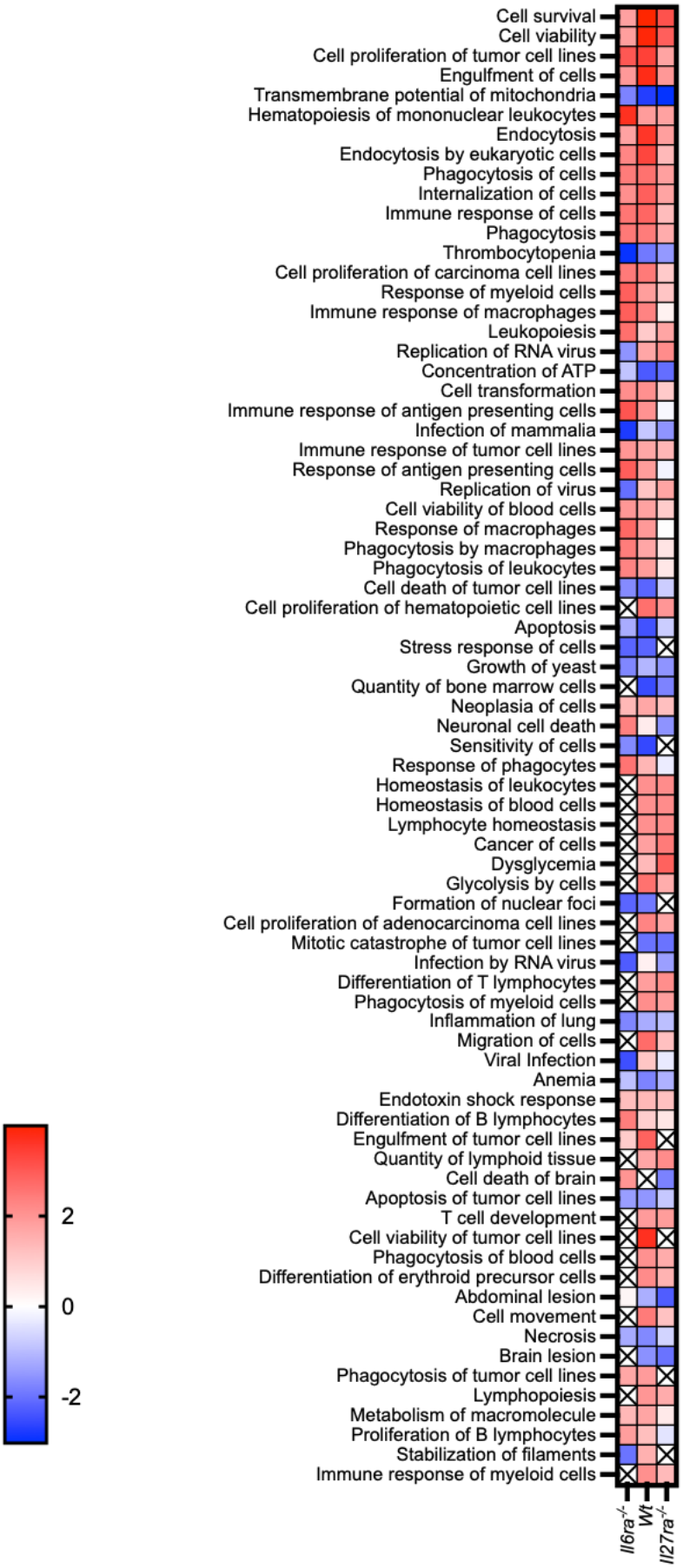


Figure 4.2.B Day 10 ATAC Disease Genes Diseases & Functions

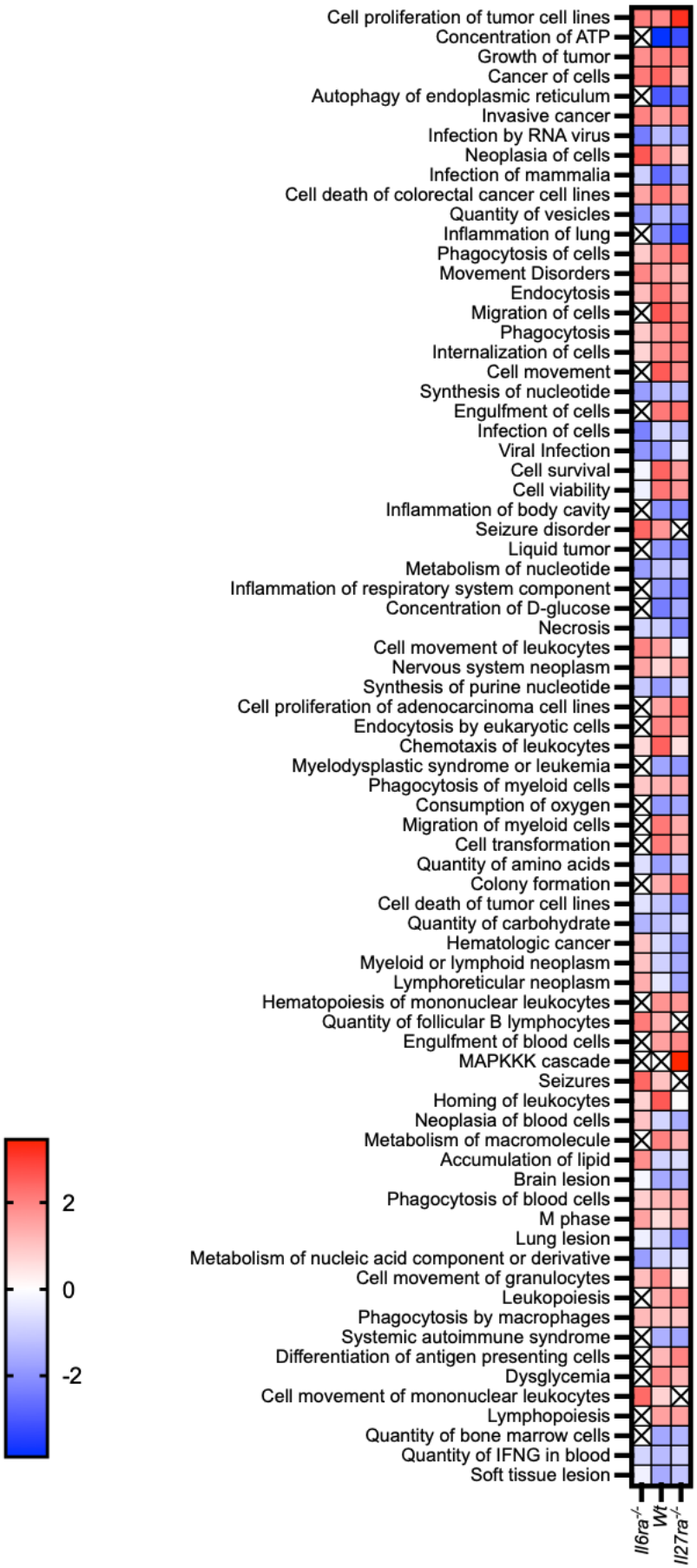


Figure 4.2: Diseases & Functions Identified by IPA Comparative Analysis of Disease Genes Mapped to RNA-seq data. The disease related genes identified by ATAC-seq experiments between genotypes were mapped to RNA-seq data and submitted to IPA for Core then Comparative analysis. Panel **A** shows Day 3 and **B** shows Day 10. The top 75 diseases and functions are exhibited in these heatmaps. They were ranked by their z-score which is displayed in the heatmap. A box marked X indicates a z-score of exactly 0.

Figure 4.3.A Day 3 ATAC Disease Genes Upstream Regulators

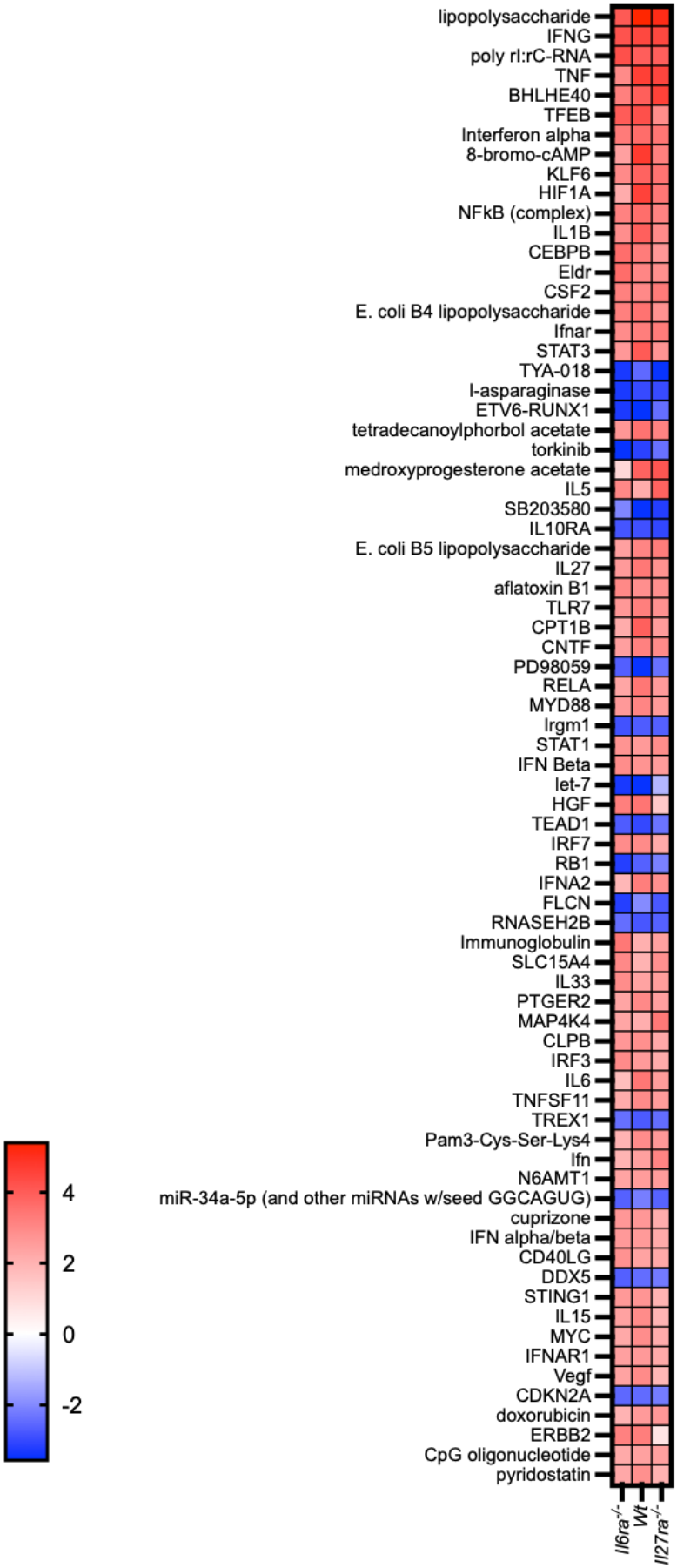


Figure 4.3.B Day 10 ATAC Disease Genes Upstream Regulators

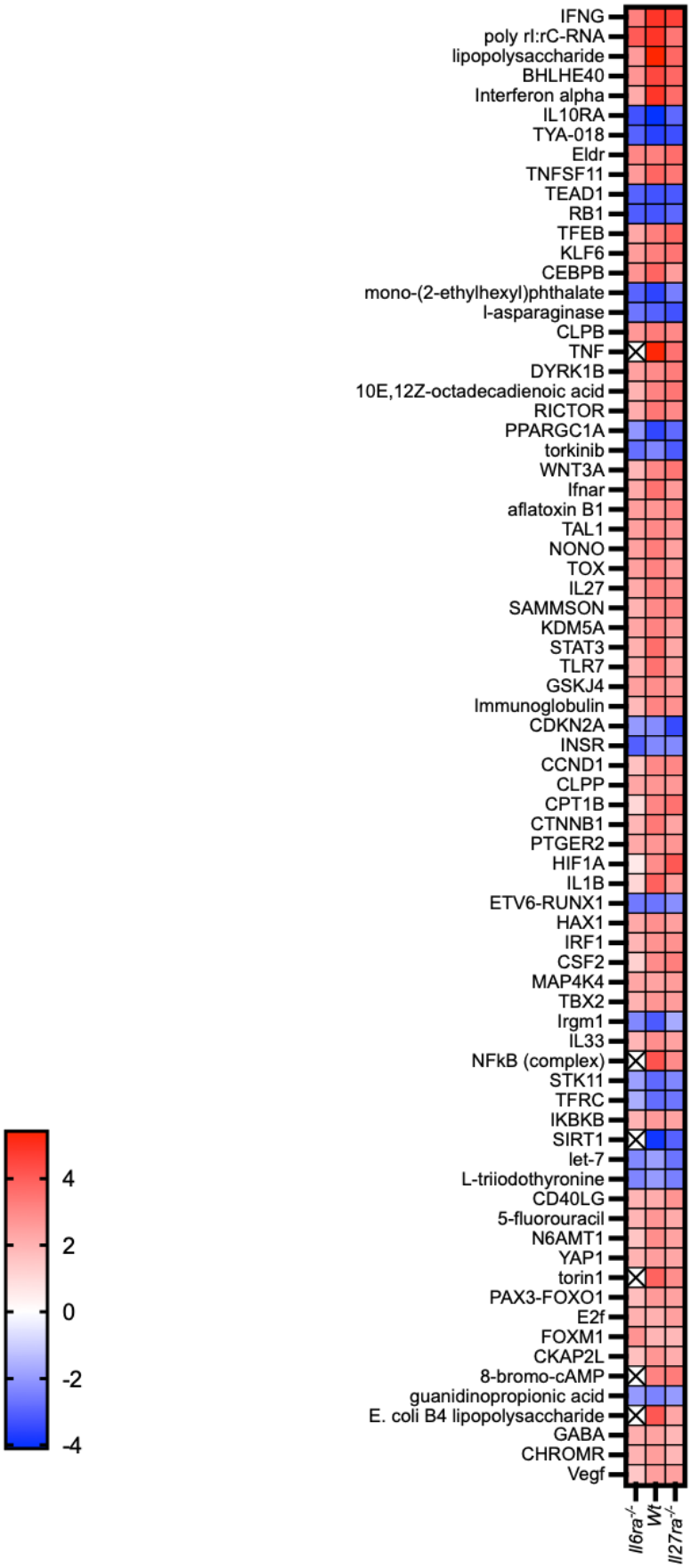


Figure 4.3: Upstream Regulators Identified by IPA Comparative Analysis of Disease Genes Mapped to RNA-seq data. The disease related genes identified by ATAC-seq experiments between genotypes were mapped to RNA-seq data and submitted to IPA for Core then Comparative analysis. Panel **A** shows Day 3 and **B** shows Day 10. The top 75 upstream regulators are exhibited in these heatmaps. They were ranked by their z-score which is displayed in the heatmap. A box marked X indicates a z-score of exactly 0.

4.2.2 Exploring expression changes in disease-associated genes.

By mapping the genes which are only accessible during synovitis to RNA-seq data, it was possible to examine how this accessibility relates to the expression of those genes. A common expression baseline of a naïve synovium (Day 0 *Wt*) was utilised so that comparisons between ATAC-seq conditions could be made. Furthermore, of the disease-associated genes, those with the greatest changes in expression from the naïve baseline could be identified and compared between conditions. In **Figure 4.4** these genes are shown to the left and right of the volcano plots illustrating expression.

The genes are highlighted in accordance with the genotypes in which they are accessible. Note that the genes are highlighted in the same colours as the section of the pie chart they fall into in **Figures 3.5, 3.6** and **3.7** in **Section 3.3.4.1** of the previous chapter. The significance of the relationship between genes and the genotype in which they are accessible has been discussed previously.

Panel **G** of the figure shows the percentage of genes which are significant and significantly altered in their expression (log_2 fold change (FC) ≥1 or ≤-1 and –log adjusted p-value (padj) ≥2). Across conditions less than 50% of genes exhibit a significant change in expression.

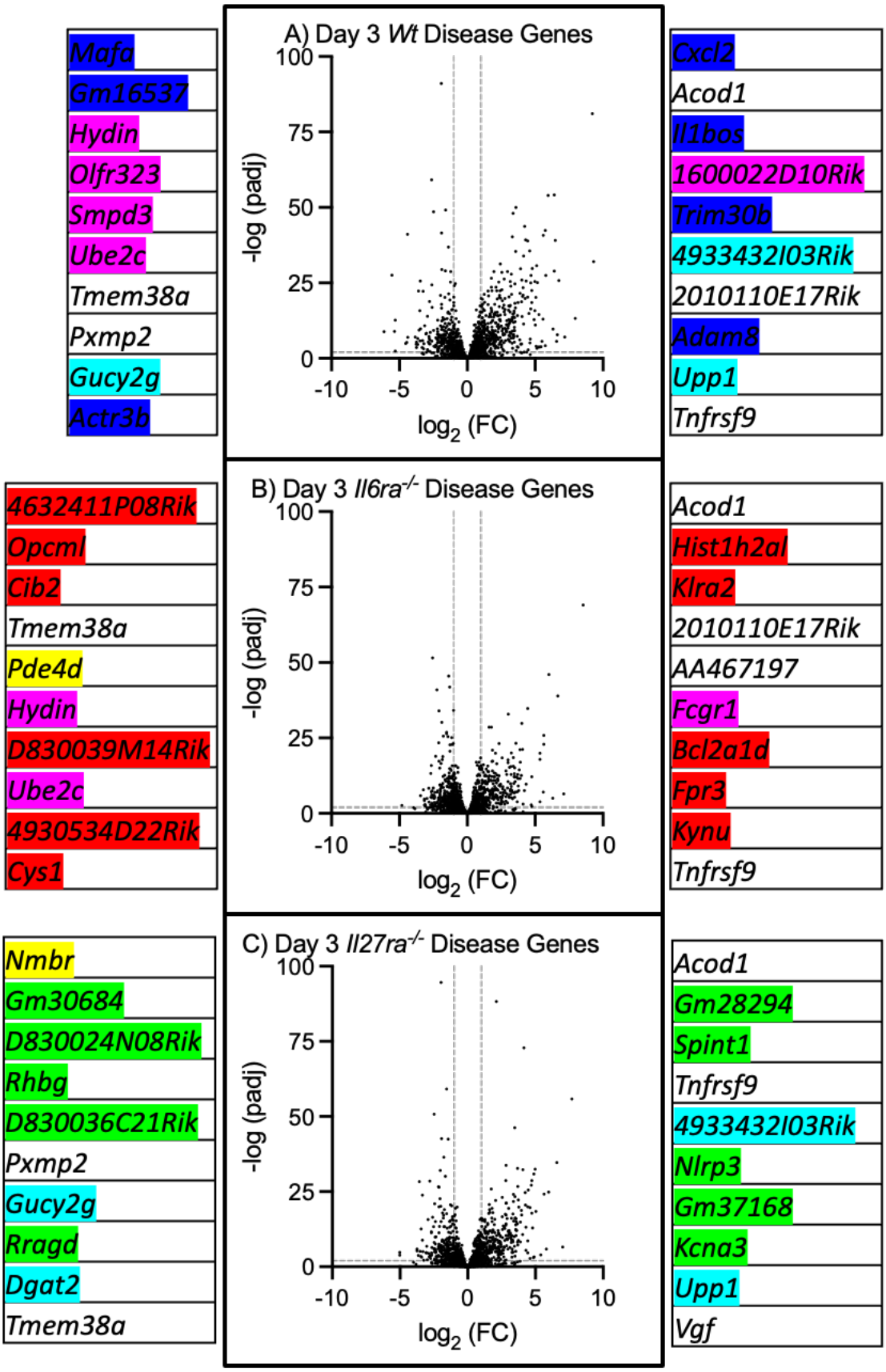


Figure 4.4.A, B & C

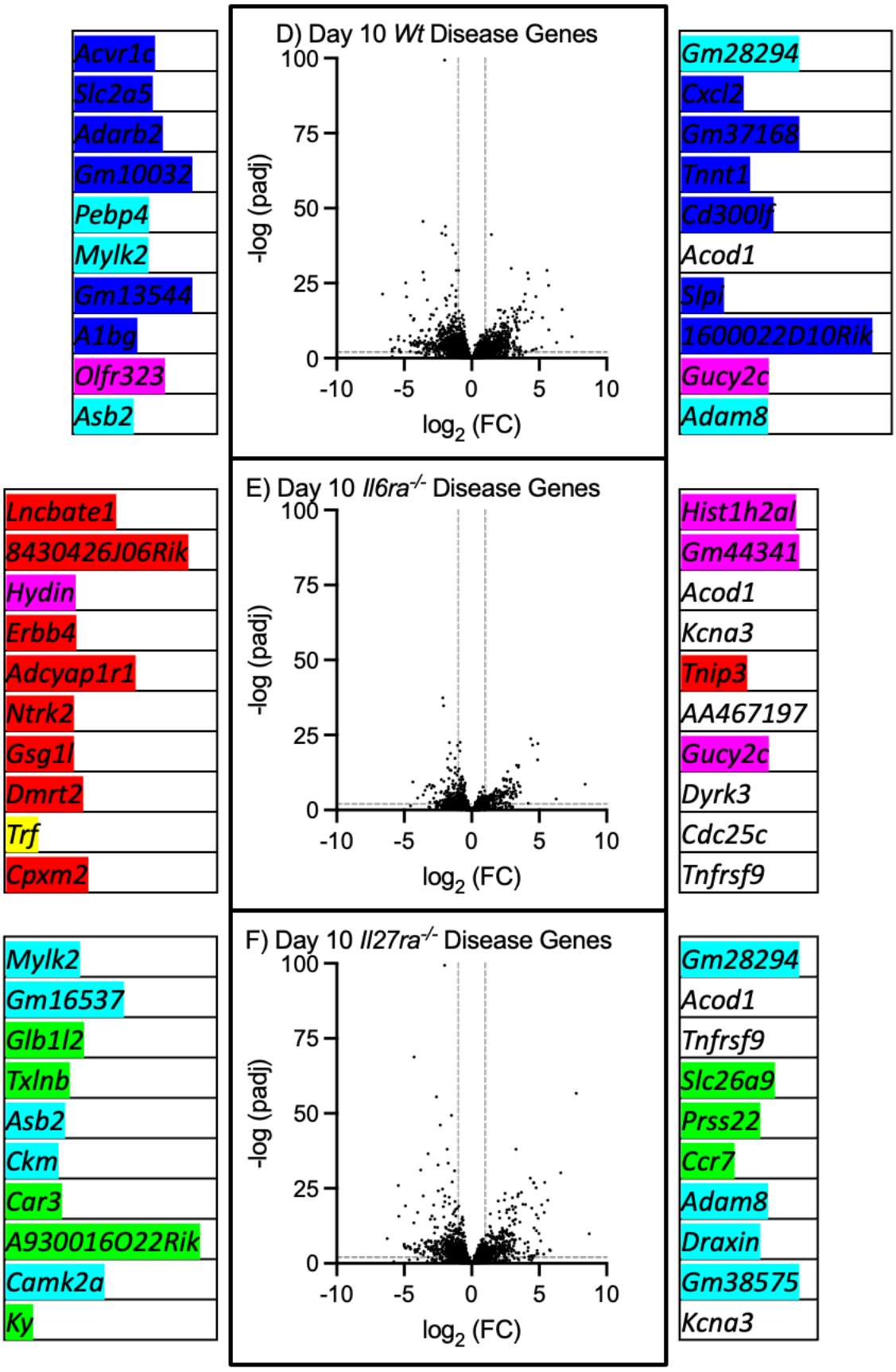


Figure 4.4.D, E & F

Unique to *Wt*
Unique to Il6ra$^{-/-}$
Unique to Il27ra$^{-/-}$
Shared between *Wt* & Il6ra$^{-/-}$
Shared between *Wt* & Il27ra$^{-/-}$
Shared between Il6ra$^{-/-}$ & Il27ra$^{-/-}$
Common to All

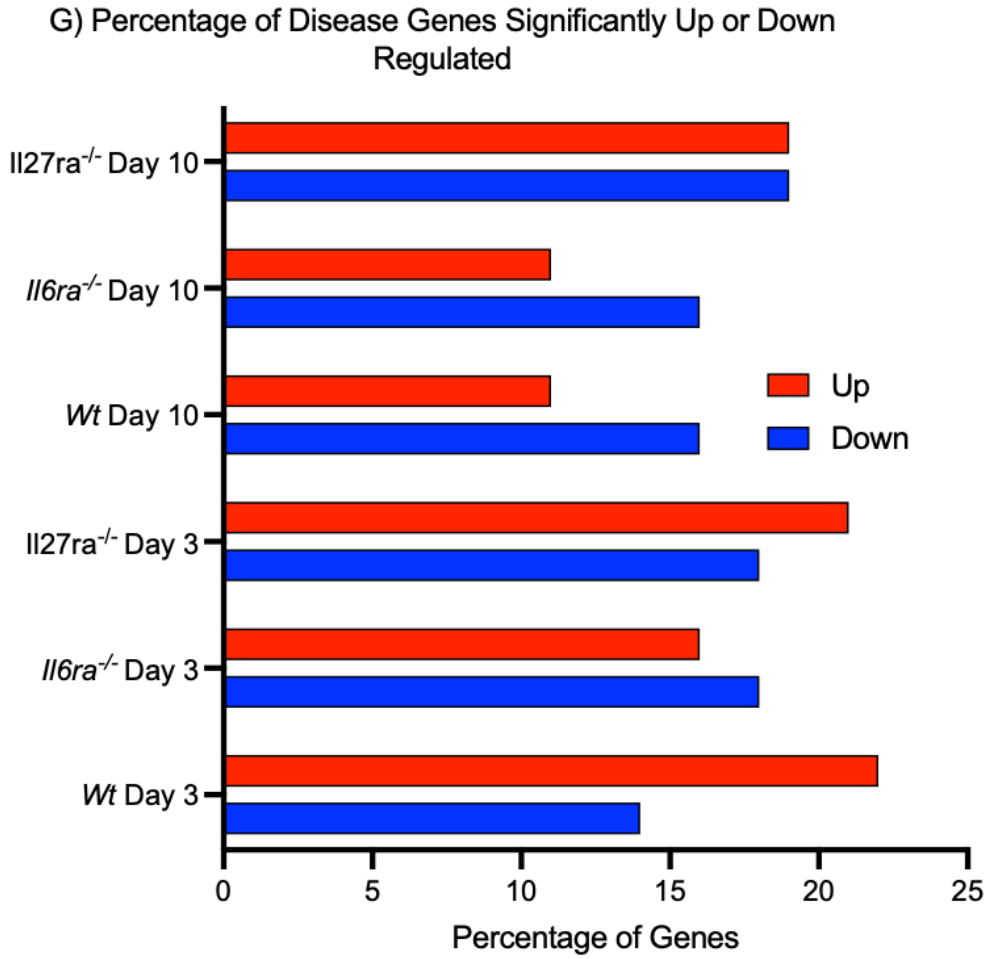


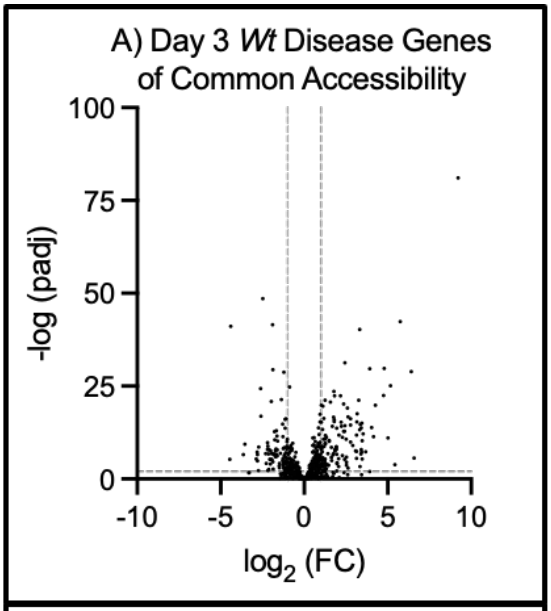
Figure 4.4.G

Figure 4.4: Changes in gene expression in ATAC-seq disease genes mapped to RNA-seq data. Volcano plots in parts **A-F** show disease gene expression profiles across the six conditions. Day 0 (Naïve) *Wt* was used as the baseline across all conditions for the comparison of gene expression. The grey dashed lines indicate significant genes: $\log_2$ fold change (FC) ≥1 or ≤-1 and –log adjusted p value (padj) ≥2. The plots are flanked by the 10 most up (right) and down (left) regulated genes. The genes are highlighted according to the genotypes in which they are accessible at that time point. The graph in panel **G** shows the percentage of genes which are significantly up (red) or down (blue) regulated. Note that the majority of genes are not significant or significantly regulated.

4.2.2.1 Exploring expression changes in disease-associated genes with common accessibility and those with variable accessibility.

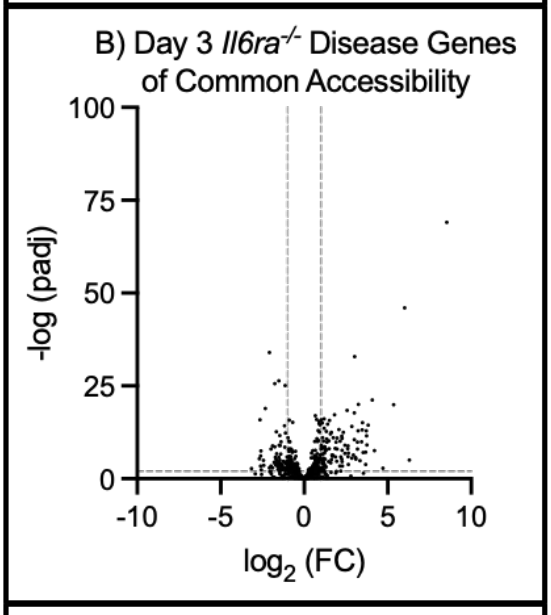
As described in **Chapter 1**, gene expression is controlled at multiple levels and gene accessibility is not equivalent to expression (Casamassimi and Ciccodicola 2019). In the previous chapter, a comparison analysis of accessible genes revealed those which are accessible irrespective of genotype and those which vary in their accessibility (i.e. either uniquely accessible in one genotype or shared across two genotypes). These differences revealed the degree to which cytokine receptor inhibition could manipulate the chromatin accessibility landscape. I, therefore, split the ATAC-seq identified disease genes into these two groups (genes accessible in all genotypes and genes of variable accessibility) and repeated the analysis expression analysis described above. In doing so it is possible to judge the impact of gene accessibility on altering the transcriptome. Likewise, differences in expression found in genes with shared accessibility indicate that there are other mechanisms involved in driving these changes in expression and altimetry downstream differences in disease activity. **Figures 4.5** and **4.6** illustrate the expression changes in commonly accessible genes and variably accessible genes, respectively. As in **Figure 4.4,** the volcano plots displayed are flanked by the top ten genes, which exhibit the greatest change in expression as compared to the naïve baseline: on the right the greatest increase and on the left the greatest decrease. In **Figures 4.5** and **4.6** genes shown in bold are part of the total disease gene top ten.

Tmem38a
Pxmp2
Epm2a
L2hgdh
Slc37a4
Pdpr
Oip5os1
Map3k20
Gm45447
Calr3



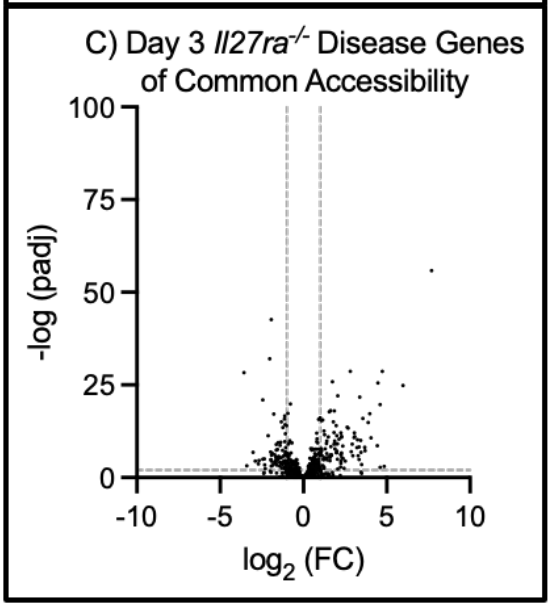
Acod1
2010110E17Rik
Tnfrsf9
AA467197
Gm15473
Samsn1
Gm49067
Rab11fip1
Cers6
Lcp1

Tmem38a
L2hgdh
Pxmp2
Epm2a
Pdpr
Calr3
1700112J16Rik
Oip5os1
Map3k20
Pnpla2



Acod1
2010110E17Rik
AA467197
Tnfrsf9
Gm15473
Gm49067
Rab11fip1
Cers6
Cdc25c
Slc2a6

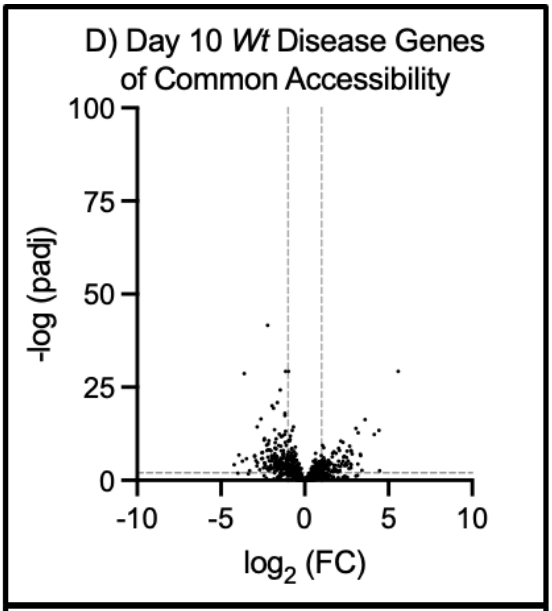
Pxmp2
Tmem38a
L2hgdh
Ppil6
Epm2a
Slc37a4
Pnpla2
Oip5os1
Pdpr
Map3k20



Acod1
Tnfrsf9
Gm15473
AA467197
2010110E17Rik
Samsn1
Rab11fip1
Gm49067
Cdc25c
Lcp1

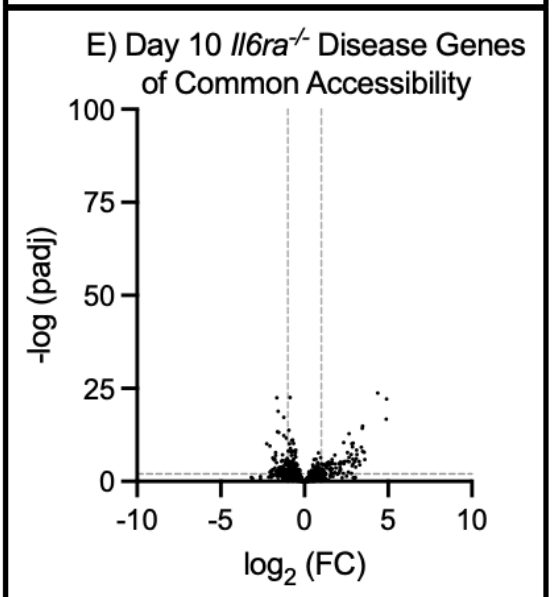
Figure 4.5.A, B & C

Ube2c
Smpd3
2900076A07Rik
Pxmp2
Epm2a
Myf6
Coq9
L2hgdh
Slc37a4
Ptges3l



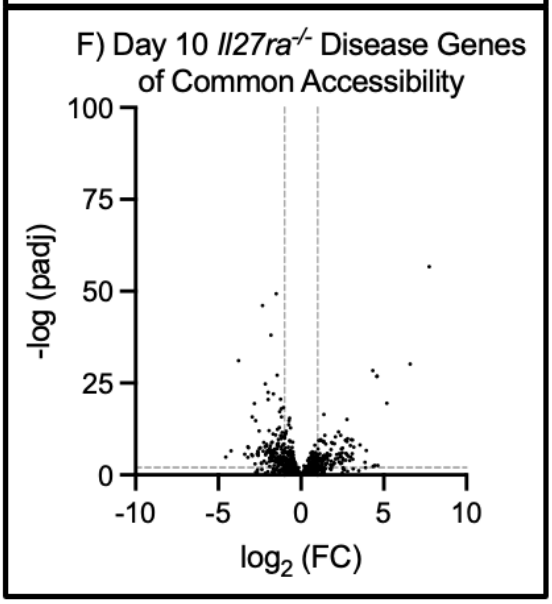
Acod1
Gm15473
Tnfrsf9
Kcna3
Rab11fip1
Gm8091
Derl3
Cdc25c
AA467197
Gm49067

Gpat3
Map3k20
Pxmp2
L2hgdh
Coq9
Pdpr
Aldh4a1
Gm34086
Hacd2
Ppip5k1



Acod1
Kcna3
AA467197
Dyrk3
Cdc25c
Tnfrsf9
Rab11fip1
Dhrs9
Slc2a6
Kif14

Ube2c
2900076A07Rik
Pxmp2
Epm2a
Ptges3l
Coq9
L2hgdh
Slc37a4
Aldh4a1
Map3k20



Acod1
Tnfrsf9
Kcna3
Gm39302
AA467197
Rab11fip1
2010110E17Rik
4933406J09Rik
Prdm1
Gm49067

Figure 4.5.D, E & F

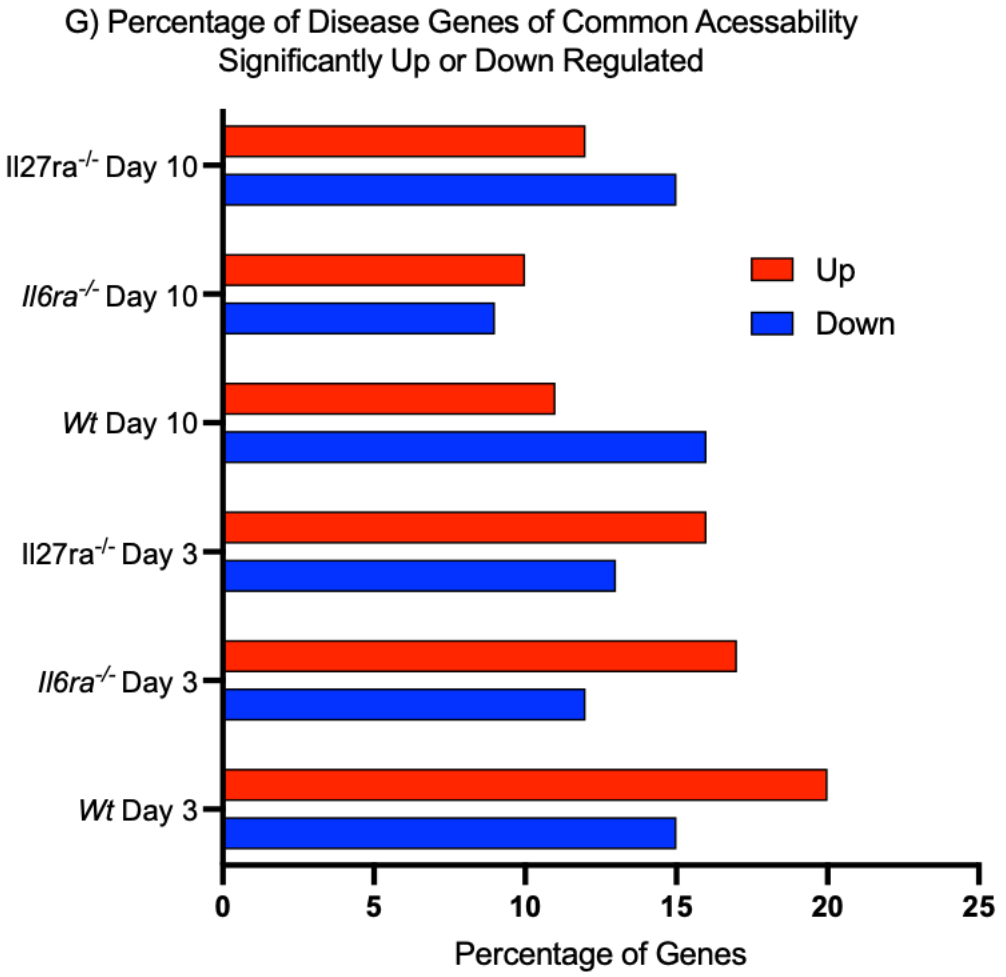


Figure 4.5.G

Figure 4.5: Changes in gene expression in ATAC-seq disease genes of common accessibility mapped to RNA-seq data. Volcano plots in parts **A-F** show expression profiles across the six conditions of disease genes exhibiting common accessibility throughout genotypes at Day 3 or 10. Day 0 (Naïve) *Wt* was used as the baseline across all conditions for the comparison of gene expression. The grey dashed lines indicate significant genes: $\log_2$ fold change (FC) ≥1 or ≤-1 and –log adjusted p value (padj) ≥2. The plots are flanked by the 10 most up (right) and down (left) regulated genes. Genes shown in **bold** are part of the top ten in the total disease gene group (see **Figure 4.4**). The graph in panel **G** shows the percentage of genes which are significantly up (red) or down (blue) regulated. Note that the majority of genes are not significant or significantly regulated.

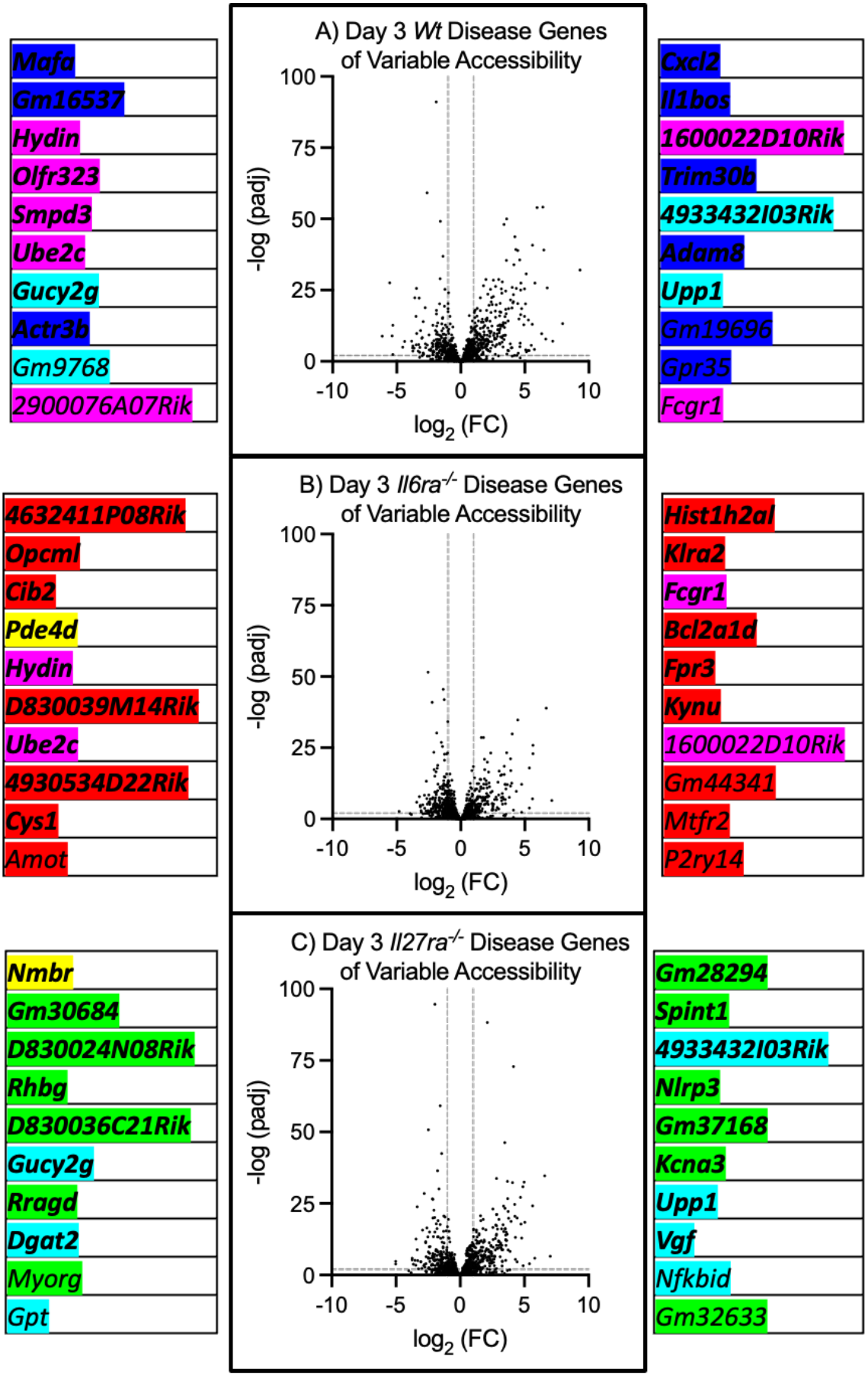


Figure 4.6.A, B & C

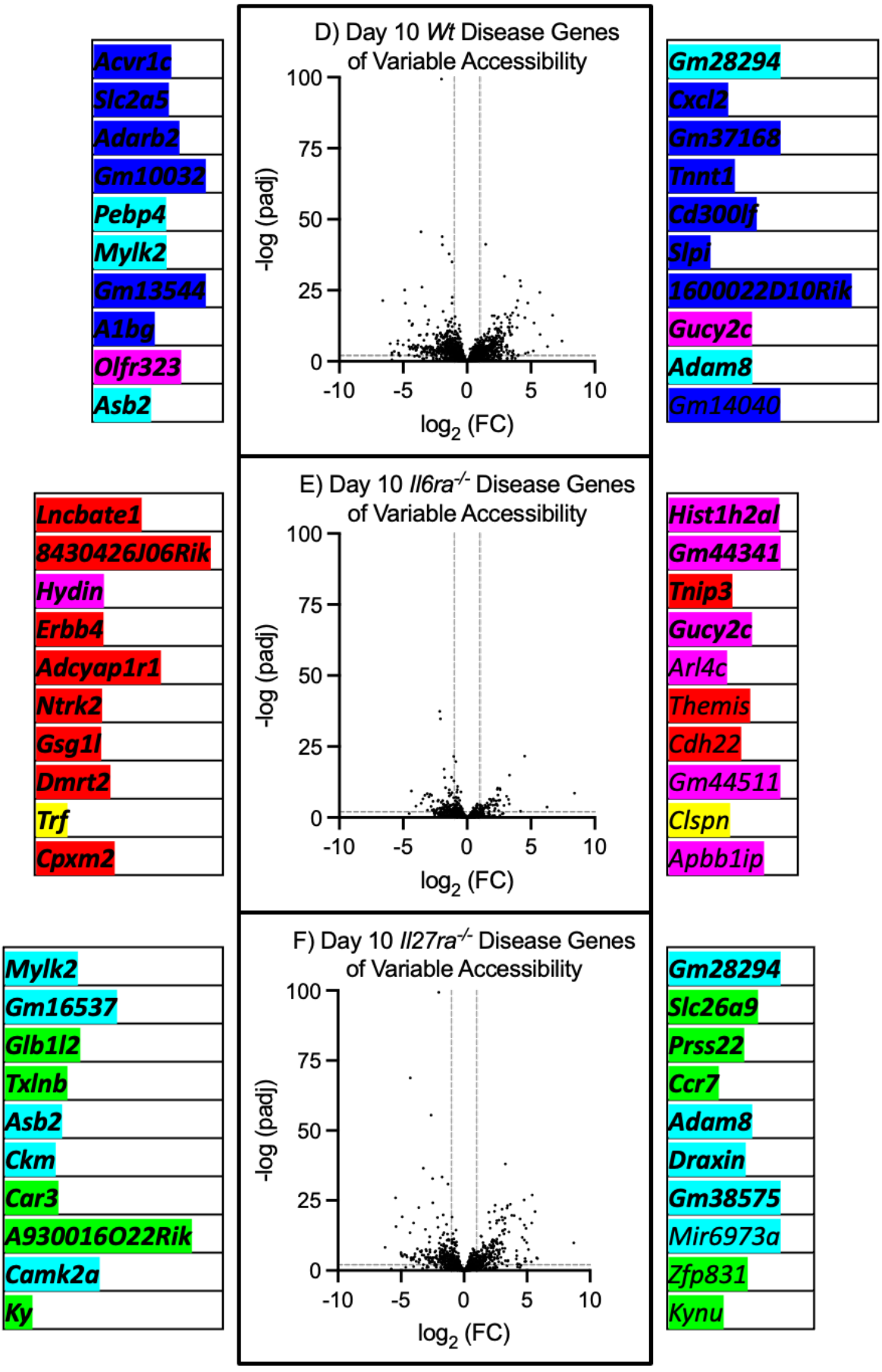


Figure 4.6.D, E & F

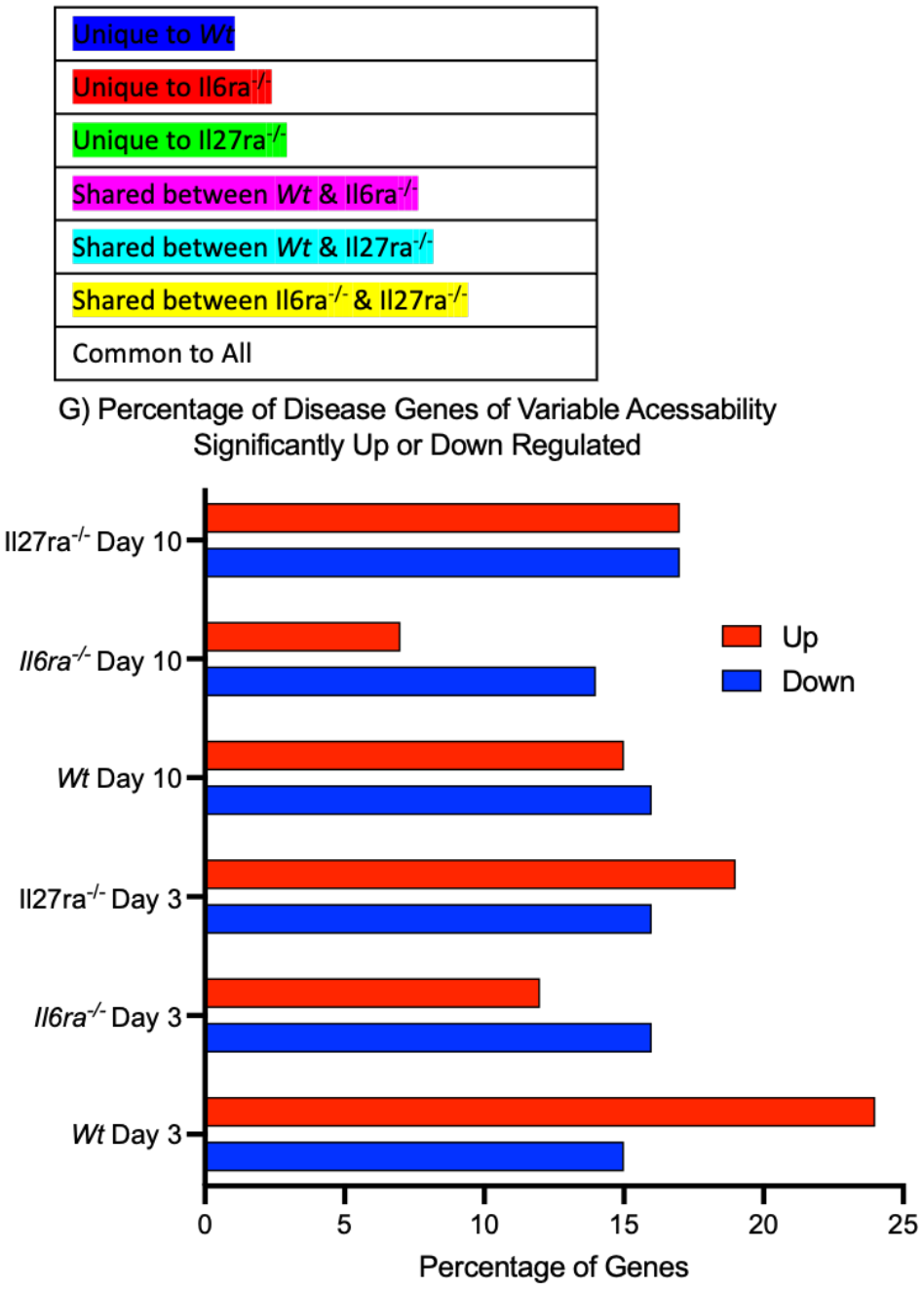


Figure 4.6.G

Figure 4.6: Changes in gene expression in ATAC-seq disease genes of variable accessibility mapped to RNA-seq data. Volcano plots in parts **A-F** show expression profiles across the six conditions of disease genes exhibiting variable accessibility between genotypes at Day 3 or 10. Day 0 (Naïve) *Wt* was used as the baseline across all conditions for the comparison of gene expression. The grey dashed lines indicate significant genes: $\log_2$ fold change (FC) ≥1 or ≤-1 and –log adjusted p value (padj) ≥2. The plots are flanked by the 10 most up (right) and down (left) regulated genes. The genes are highlighted according to the genotypes in which they are accessible at that time point. Genes shown in **bold** are part of the top ten in the total disease gene group (see **Figure 4.4**). The graph in panel **G** shows the percentage of genes which are significantly up (red) or down (blue) regulated. Note that the majority of genes are not significant or significantly regulated.

4.2.3 Translating ATAC-seq disease genes of interest to human rheumatoid arthritis cell populations.

The analysis thus far presented yielded several genes of particular interest. This section focuses on disease genes (both commonly and variably accessible) which showed the greatest degree of upregulation from the naïve baseline. These are the same genes shown in the right columns of **Figure 4.4**. In order to investigate their role in synovitis further, particularly in relation to human disease, the Accelerated Medicines Partnership (AMP) RA Phase I website was used (https://immunogenomics.io/ampra/). This website makes use of work concerning the various cell types involved in human synovitis, which have been identified as potential drivers of inflammation and the disease process generally (Zhang et al. 2019). The website tool uses single-cell RNA-seq data obtained from the synovium of patients with RA. By searching for a gene, the cell groups in which said gene is enriched are displayed. **Figure 4.7** displays screenshots from the AMP website for our genes of interest which appeared in the single-cell RNA-seq. Panel **G** of the figure is another screenshot from the AMP website displaying the cluster annotations used in the other figure panels.

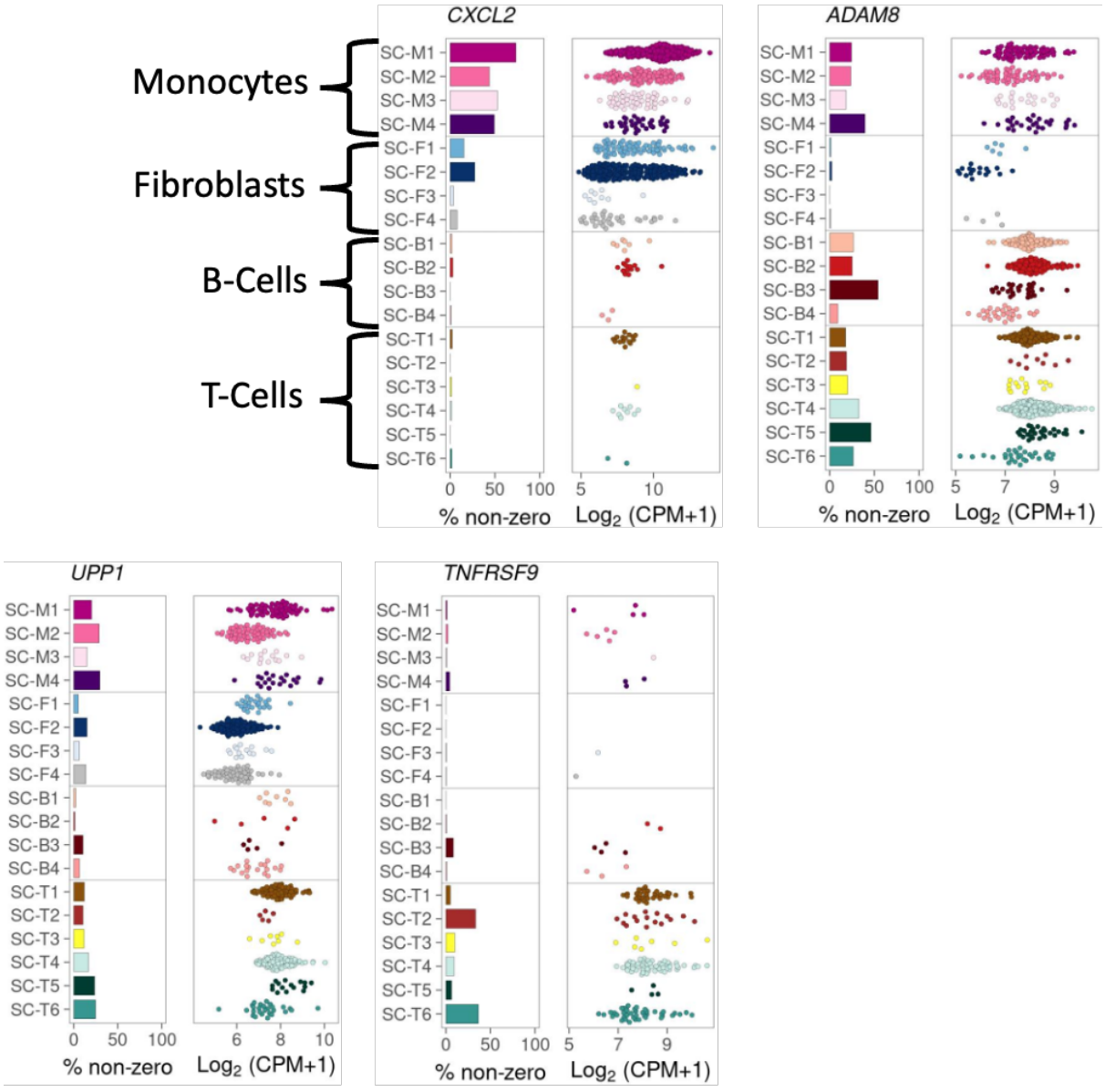


Figure 4.7.A: Day 3 *Wt*

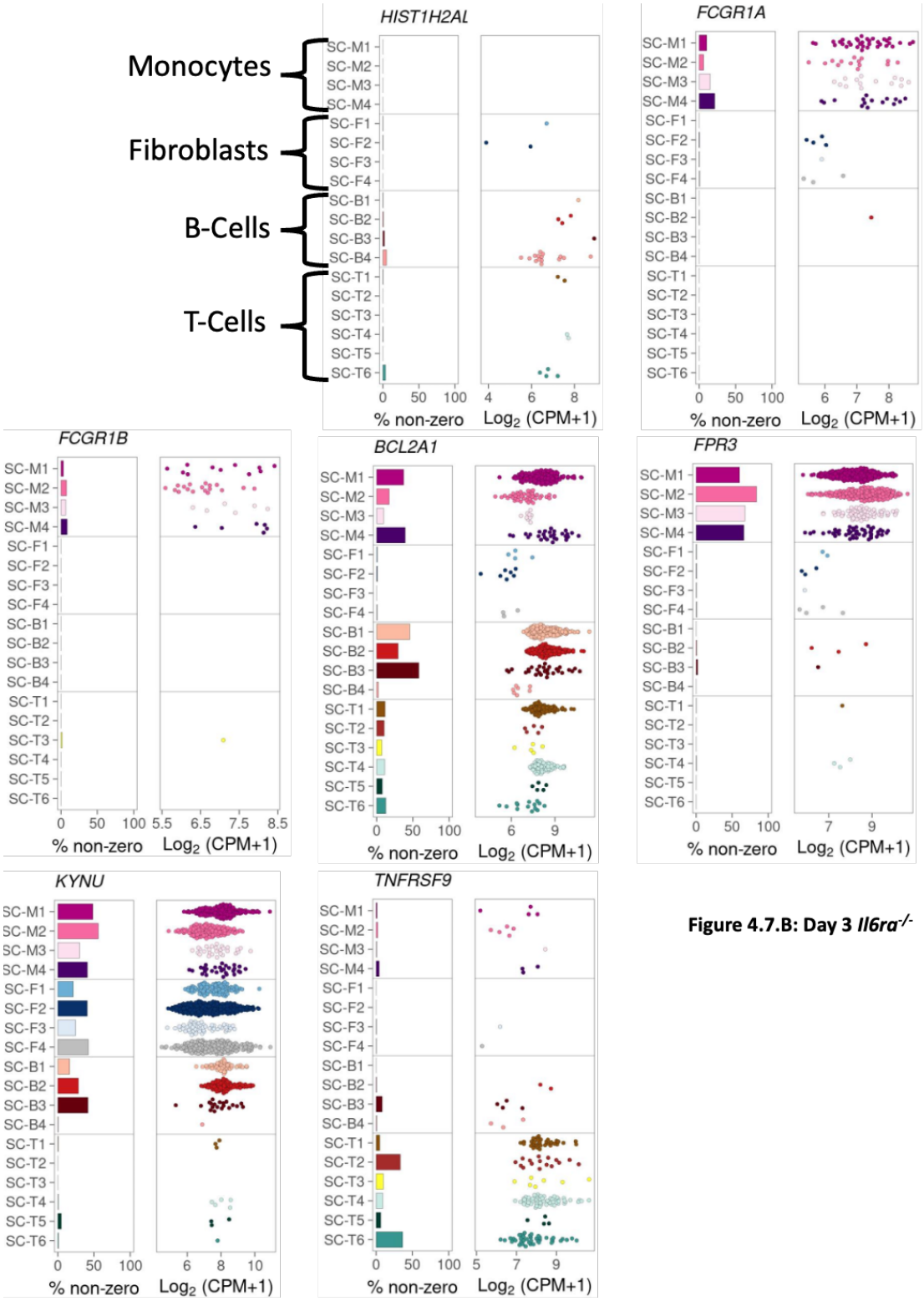


Figure 4.7.B: Day 3 *Il6ra*$^{-/-}$

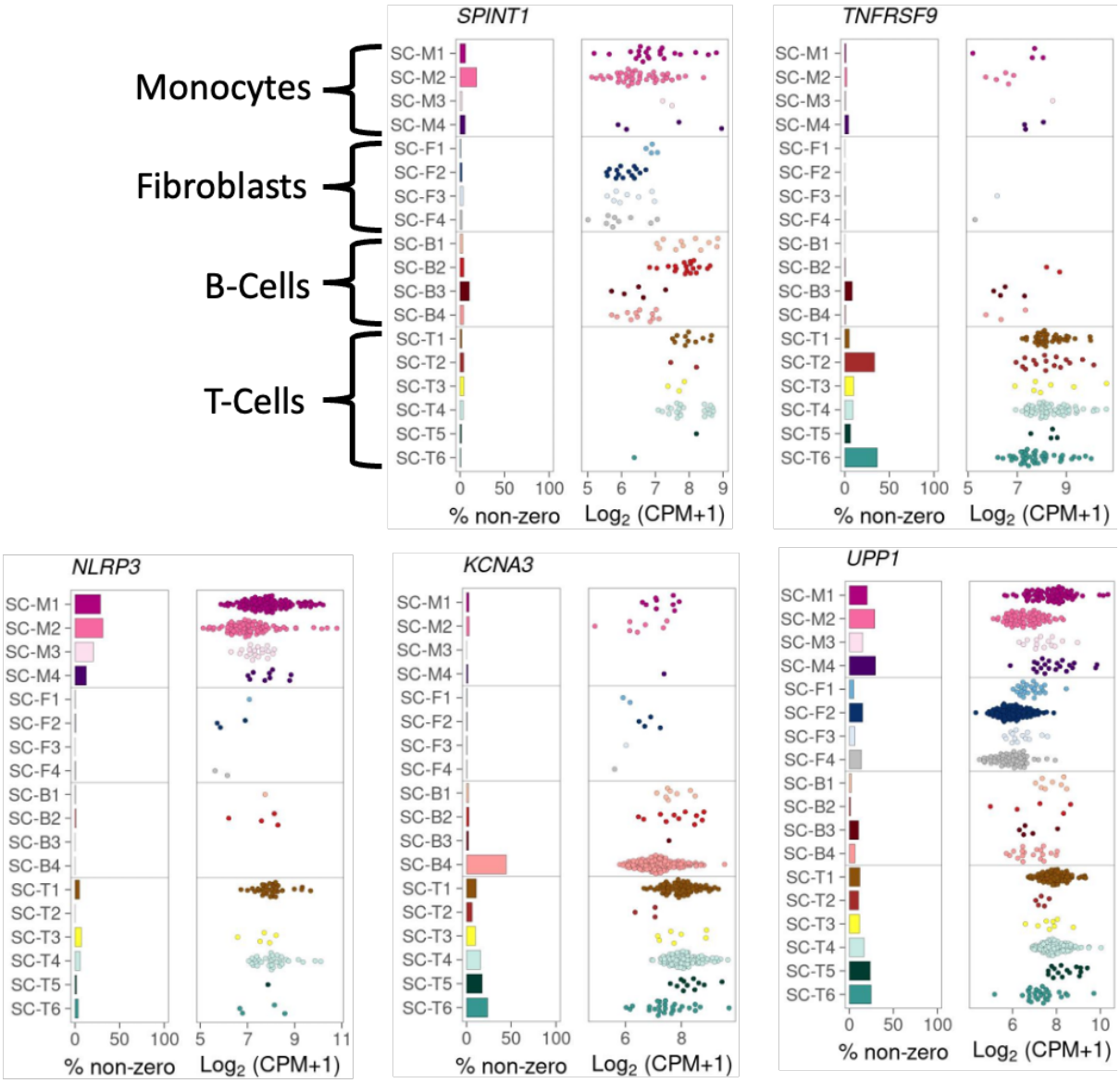


Figure 4.7.C: Day 3 *Il27ra*$^{-/-}$

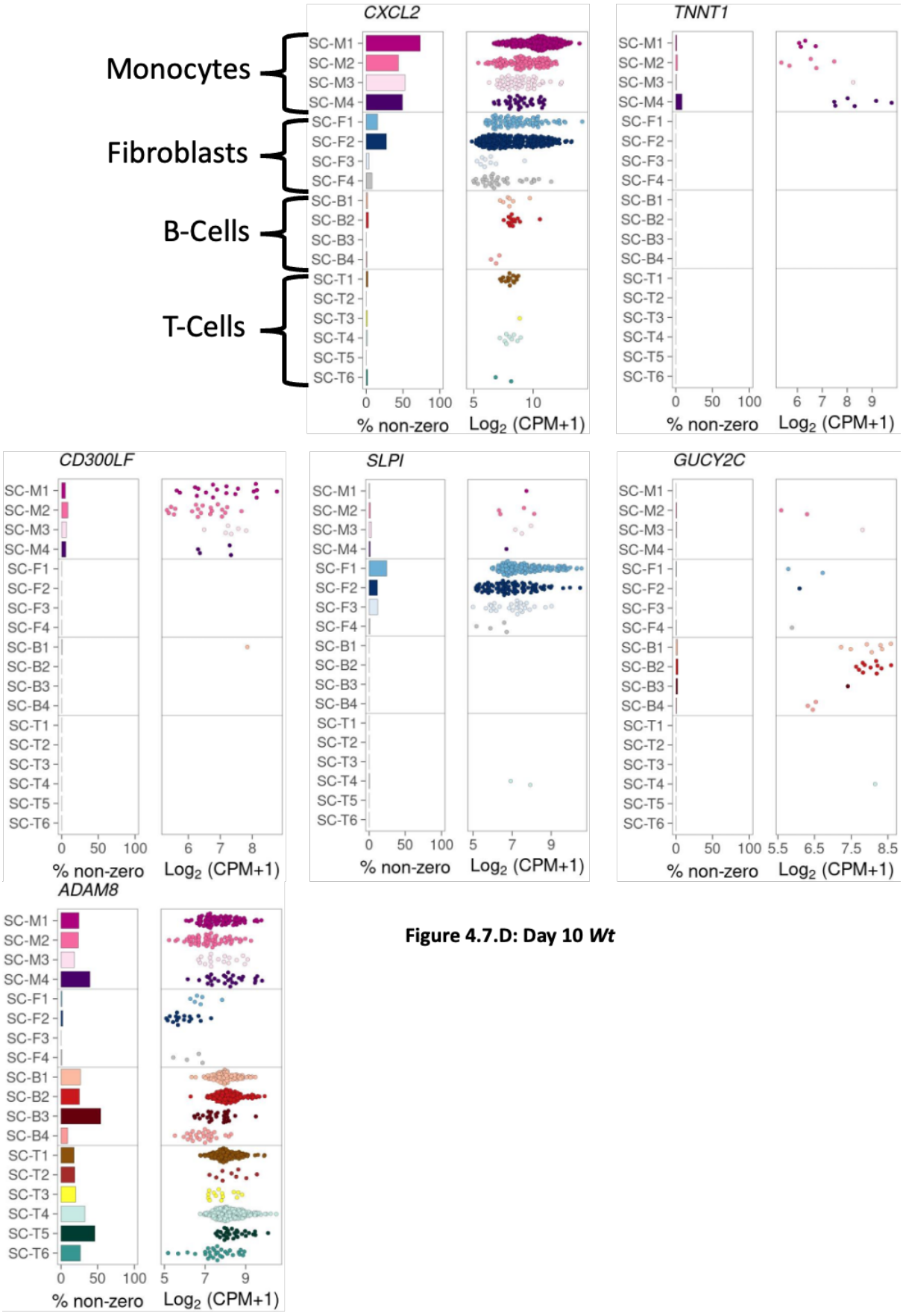


Figure 4.7.D: Day 10 *Wt*

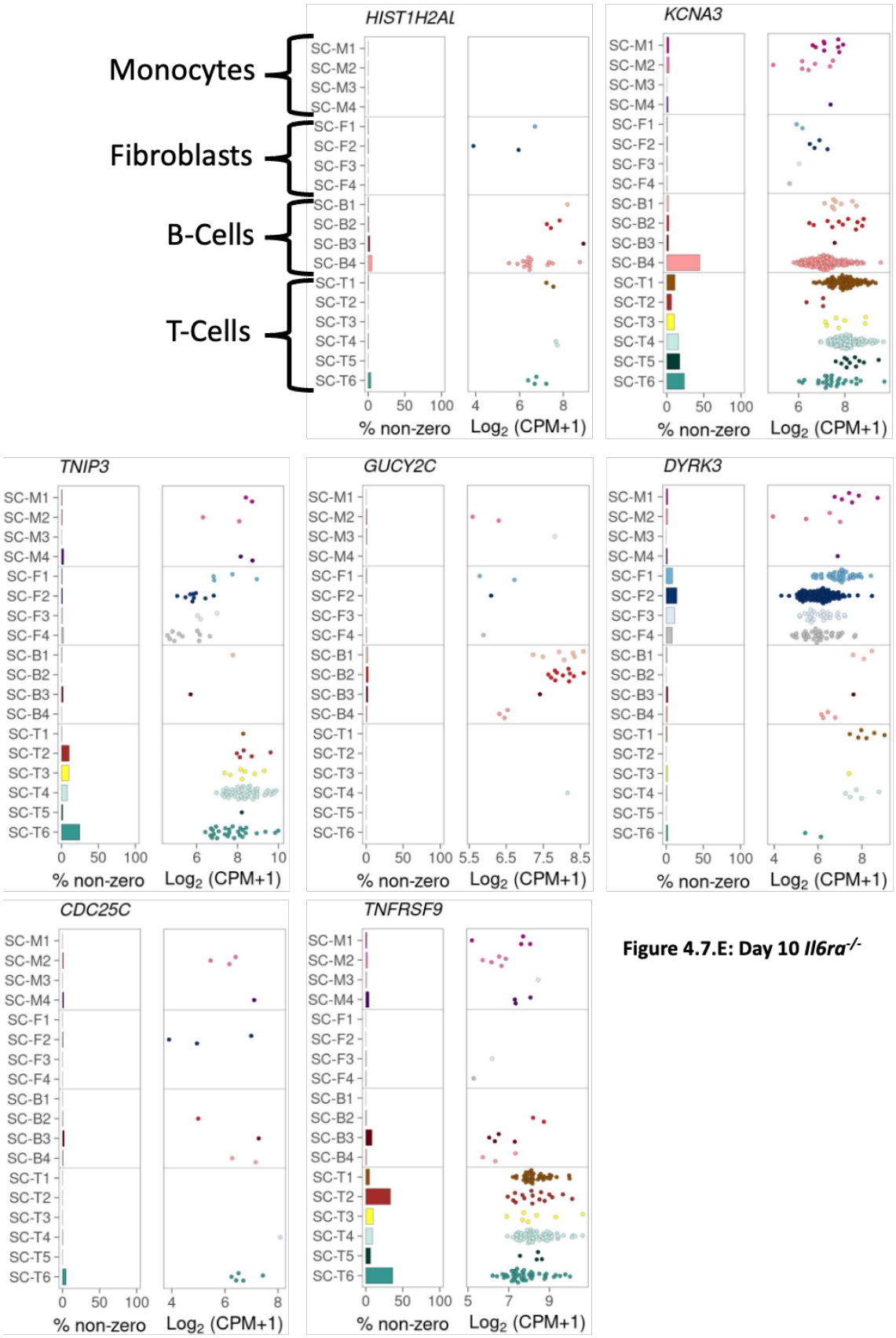


Figure 4.7.E: Day 10 *Il6ra*$^{-/-}$

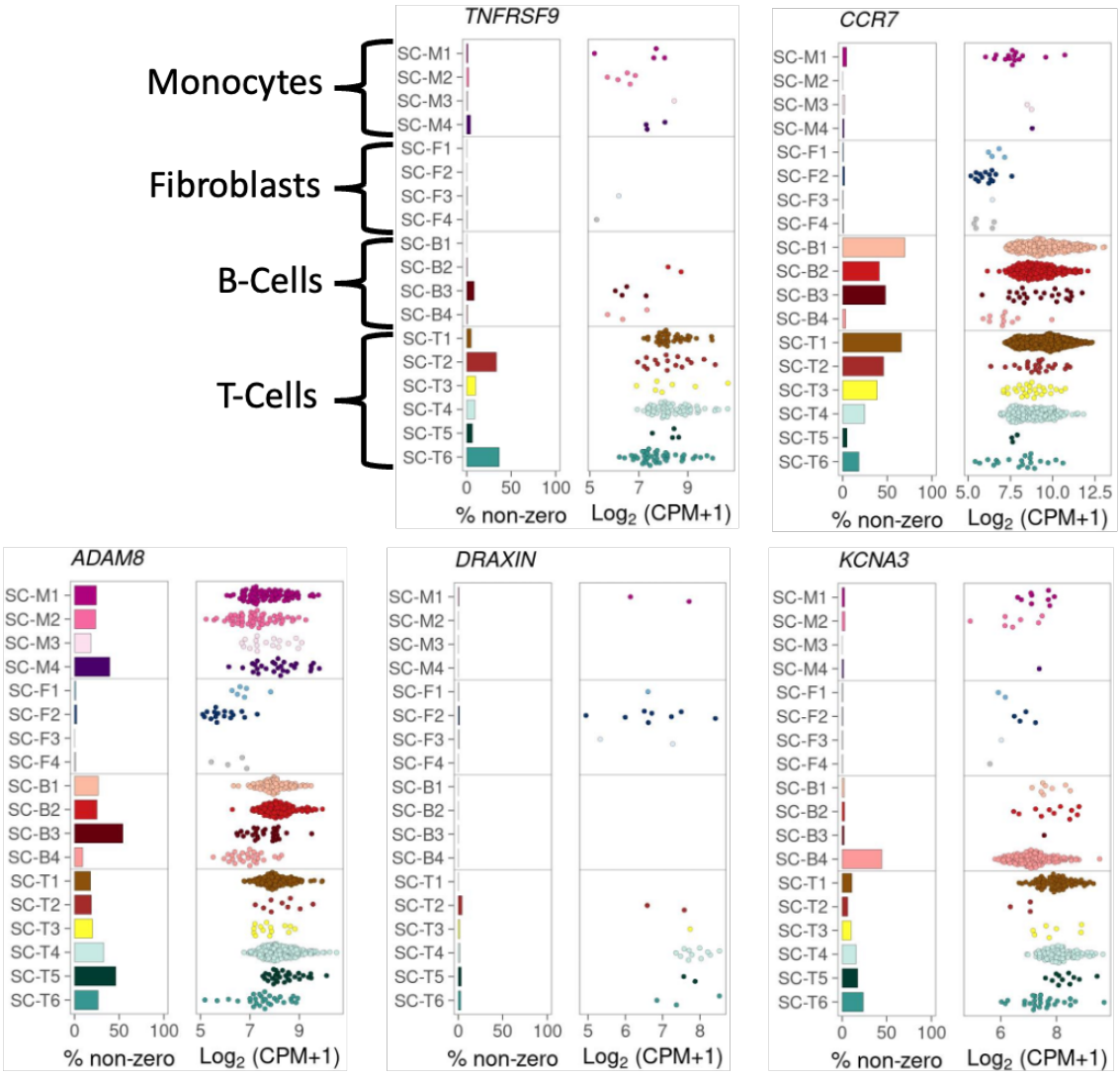


Figure 4.7.F: Day 10 *Il27ra*$^{-/-}$

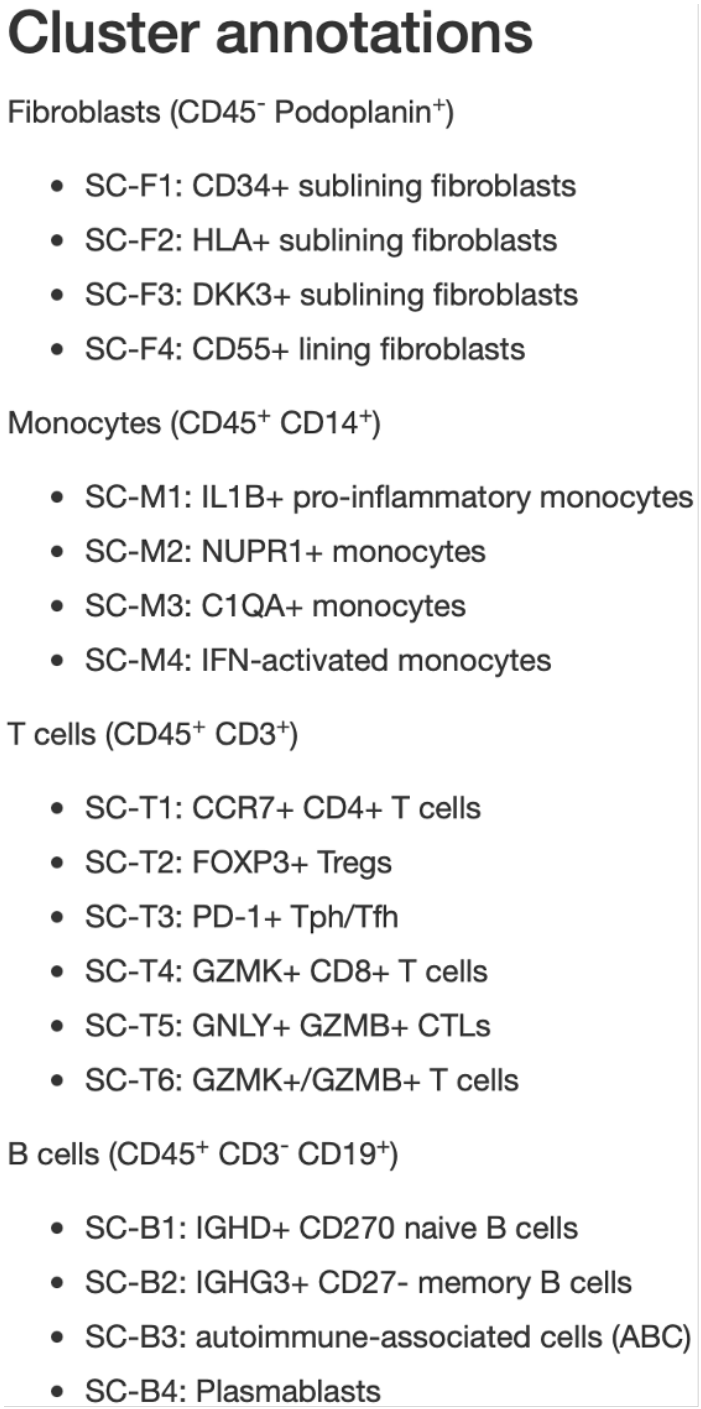


Figure 4.7.G

Figure 4.7: Accelerated Medicines Partnership (AMP) RA Phase I Cell Cluster analysis of ATAC-seq disease genes of interest. ATAC-seq identified genes associated with the disease state. These were mapped to RNA-seq data, to reveal the top ten most upregulated genes as compared to a naïve baseline. Those genes were entered into the AMP website tool. Those which returned results are displayed in panels **A-F**. The cluster annotations are shown in panel **G.**

4.3 Discussion

4.3.1 Interpretation of IPA Analysis

4.3.1.1 Canonical Pathways

When inspecting the heatmaps of **Figure 4.1** a few pathways stood out as particularly interesting. For instance, at Day 3 the pathway; '*Role of JAK family kinases in IL-6-type Cytokine Signalling'* is not enriched in the *Il6ra*$^{-/-}$ genotype. This illustrates the data's validity given the inhibition of the receptor. The genotype with the highest z-score in this pathway is the *Il27ra*$^{-/-}$. To examine further, of the genes enriching this pathway the top three with the highest degree of upregulation in the *Il27ra*$^{-/-}$ genotype were determined. They are *Junb, Socs1* and *Il6ra*. Given the association between the *Il27ra*$^{-/-}$ genotype and a fibrotic form of RA (Jones et al. 2015), as described in **Chapter 1 Section 1.5.2**, *Junb* is particularly interesting as it has been shown to be highly expressed in synovial fibroblast-like cells in RA (Kinne et al. 1995).

At Day 10 two pathways stand out in particular: '*Role of Chondrocytes in RA Signalling Pathway'* and '*Role of Osteoclasts in RA Signalling Pathway*'. In both pathways, the *Wt* genotype is the most enriched. Again, the top three genes were determined. In the chondrocyte pathway, they are *Cxcr4, Nlrp3* and *Nos2*. All three are associated with RA disease processes and increased expression in the RA synovium (Perkins et al. 1998; Peng et al. 2020; Yin et al. 2022). In the osteoclast pathway, the three genes are *Adam8, Nfkbid* and *Fcgr1a*. The first *Adam8,* is associated with increased bone damage in RA (Ainola et al. 2009). Likewise, the other two are associated with RA pathology (Scheinman 2013; Theeuwes et al. 2023).

4.3.1.2 Diseases & Functions

The heatmaps of **Figure 4.2** include many functions related to inflammation and immunity. At Day 3 the top-most of these is *'Cell Survival'.* The most enriched in the *Wt* genome, and the top three genes are *Cxcl3, Fcgr1a* and *Cxcr4.* The latter two have been mentioned previously while *Cxcl3* is involved in the invasion of fibroblast-like synoviocytes into the joint during RA inflammation (Laragione et al. 2011).

At Day 10, the same pathway is again most enriched in the *Wt* genome. Two of the top three genes; *Cxcl3* and *Cxcr4* are conserved from Day 3, highlighting the long-term changes from the naïve baseline. The third gene is *Lilrb4* which is involved in the coordination of the immune response and has been identified as a potential therapeutic target in treating RA (Abdallah et al. 2021).

4.3.1.2 Upstream Regulators

At Day 3 two upstream regulators in particular stand out due to their relevance to the work of this thesis and their importance in the pathology of RA; STAT1 and STAT3 (Jones et al. 2015; Twohig et al. 2019). Numerous other inflammatory mediators are present across both Day 3 and Day 10. The variation between conditions, though subtle in places speaks to the various other mechanisms by which genes can be regulated aside from chromatin accessibility.

4.3.2 AMP Cell Type Analysis

The plots obtained from the AMP website which comprise **Figure 4.7** indicate that there are a wide variety of cell types and subtypes in which the genes submitted are enriched. The *Wt* genotype of all four principal cell types; fibrocytes, monocytes, T-cells, and B-cells, are represented by the genes analysed. In the *Il6ra*$^{-/-}$ genotype, however, while the various cell populations are represented, they are less enriched than the other two genotypes. This behaviour tallies with the genotype's low inflammatory disease phenotype. In the *Il27ra*$^{-/-}$ genotype T-cells and B-cells are the predominate cell types enriched, the cell types which also aggregate to form ELSs (Corsiero et al. 2016). Given the role of IL-27 in inhibiting ELSs (Jones et al. 2015) finding these cell types to be enriched, particularly at Day 10 where inflation is more established, highlights the role of chromatin accessibility with regards to this particular pathogenic mechanism.

The AMP data originates from human synovium showing the validity of our own work in relation to human disease. These investigations, pave the way for future murine experiments involving single-cell ATAC-seq to investigate the chromatin accessibility landscape in more detail.

4.3.3 Interpretation of differential expression analysis and its ramifications regarding the ATAC-seq method

Chromatin accessibility is one of several levels of control over gene expression. By mapping the ATAC-seq data to RNA-seq data the expression of different ATAC-seq gene subsets could be analysed. Across **Figures 4.4, 4.5** and **4.6** it is apparent that the majority of ATAC-seq identified genes vary little in their expression as compared with the naïve baseline with ~50% of genes not exhibiting significant change. This was somewhat surprising as analysis in the previous chapter indicated that these particular genes were not accessible in the naïve condition. Genes which exhibited a decrease in their expression are more surprising as not only had they been shown not to be accessible but are in fact being expressed to a greater degree in the naïve condition.

The scale of chromatin accessibility must be considered. In the previous chapter, it was shown that each condition was comprised of several thousand genes. We, therefore, propose that the role played by differences in chromatin accessibility in causing inflammatory heterogeneity is a relatively subtle one. Furthermore, this analysis highlights the importance of considering other regulatory mechanisms of gene expression. Transcription factor activity for instance is the focus of the next chapter.

In the previous chapter, temporal comparisons of gene accessibility showed changes over the experimental time course. This indicated a dynamic and variable chromatin accessibility landscape even within a genotype. As the RNA-seq data to which this ATAC-seq data was mapped originates from a different set of experiments this dynamism may be overlooked. Further, albeit slight, temporal differences in chromatin being accessibility to it being transcribed to messenger RNA (mRNA) may also account for why we do not see greater variability in expression in genes with variable accessibility.

To address this, it would be of great interest to repeat ATAC-seq experiments at more timepoints during an AIA experiment so that the rate of change in chromatin accessibility during inflammation can be gauged. Likewise, conducting synchrony ATAC-RNA-seq experiments on tissue gathered from the same animal may shed light on the temporal shift in chromatin's accessibility to its translation.

As described in **Chapter 3 Section 3.2.1** the gene lists used in this ATAC-seq analysis are not exhaustive of every gene found to be accessible in ATAC-seq experiments. Peaks,

and the genes within them, which did not meet the cut-off score described were removed. This, and the mechanism described above potentially explains why genes uniquely accessible in one genotype are represented across multiple genotypes in the RNA-seq data.

Finally, ATAC-seq is not the ultimate measure of chromatin accessibility. Rather it is a measure of chromatin's accessibility to transposase (Grandi et al. 2022). Chromatin lies on a spectrum ranging from completely inaccessible and inactive regions to regions bound to DNA-binding proteins composed of gene regulatory areas (Ernst and Kellis 2010). On this spectrum sits the mutant Tn5 of ATAC-seq. Other transcription factors may bind to and function in regions in which Tn5 is not present, resulting in genes identified in RNA-seq experiments not appearing in ATAC-seq analysis.

Chapter 5:

Tracking STAT transcription factor involvement in the joint pathology of antigen-induced arthritis.

5.1 An Introduction to ChIP-seq

Differences in the transcriptional output of mice (*Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$) with AIA have been described elsewhere (Hill, 2019) and support the chromatin accessibility results presented in **Chapter 3** and **Chapter 4**. These findings point towards the importance of cytokine signalling *via* the Jak-STAT pathway in shaping gene expression affecting the course of synovitis and disease outcome (see **Chapter 1;** Latchman 1997; Klemm et al. 2019). To gather further insights into the control of synovitis by STAT transcription factors, experiments in this chapter used chromatin immunoprecipitation sequencing (ChIP-seq) to identify how synovial cells sense and interpret cytokine cues in response to antigen-induced arthritis.

While RNA-seq and ATAC-seq offer opportunities to examine changes in transcriptional and epigenetic gene regulation, ChIP-seq allows a genome-wide analysis of specific DNA-binding proteins. This method relies upon antibody immunoprecipitation of regulatory molecules or transcriptional factors bound to genomic DNA (O'Shea et al. 2011; Hirahara et al. 2015; Villarino et al. 2017). Signal Transducer and Activator of Transcription 1 (STAT1) and STAT3 were examined, as both IL-6 and IL-27 signalling utilise the JAK-STAT pathway. These transcription factors are major determinants in the perpetuation of the inflammation which characterises RA. Additionally, they play a key role in regulating the functions and identity of immune cells, principally T and B cells. Dysfunction of this control is a major factor in RA's disease course (Jones et al. 2015; Twohig et al. 2019). Furthermore, commonly target JAK-STAT cytokine signalling is the target of many biological medicines and small molecule inhibitors used in the treatment of rheumatoid arthritis and associated inflammation (Haraoui and Bykerk 2007; Akahoshi and Watabe 2009; Harrington et al. 2020). This combination of factors relating to STAT1 and STAT3 involvement in the pathology of RA made them particularly interesting candidates for investigation.

STAT transcription factors are regulated at multiple levels via the JAK-STAT pathway and upstream cytokine signals. This allows for a high degree of specificity and variation in transcriptional output (Liu et al. 1998). Aside from the different cytokines which activate the JAK-STAT pathway, there are intracellular factors also influencing transcription. JAK-STAT signalling requires the hydrolysation of tyrosine residues by phosphatases. STATs are involved in the transcription of genes which themselves influence STATs. These include

Suppressor of Cytokine Signalling (SOCS), which works as negative feedback regulators of JAK-STAT signalling and is induced by STAT transcription factors (Morris et al. 2018). Other proteins such as Protein Inhibitor of Activated STAT (PIAS) function in a similar manner to SOCS, curtailing STAT activity (Liu et al. 2004). Additionally, the proteasomal degradation of STATs is driven by ubiquitin ligases. The half-life of mRNA translated by STATs is controlled by various microRNAs and RNA-binding proteins. Finally, STAT activity can be suppressed or promoted by a variety of post-translational modifications (Villarino et al. 2017).

Previous studies conducted by our laboratory have used ChIP-seq protocols to investigate how interferon-gamma (IFNγ) alters the transcriptional output of IL-6 to promote tissue fibrosis following repeated episodes of peritonitis (Millrine et al. 2023). Using a similar strategy, studies described in this Chapter have used ChIP-seq to provide insights into the role of STAT transcription factors in shaping disease heterogeneity in synovitis. This chapter discusses the selection of the antibodies used in ChIP-seq experiments and the assessment of STAT1 and STAT3 ChIP-seq datasets from synovial tissues extracted from *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice with AIA.

5.2 Summary of the ChIP-seq Method

The ChIP-seq protocol is summarised in Figure 5.1 and is fully described in **Chapter 2 Section 2.1.9**. ChIP-seq analysis of STAT1 and STAT3 transcription factors was performed on synovial tissue extracts obtained on Day 3 and Day 10 of AIA in *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice. STAT3 ChIP-seq was carried out in the same samples with the exception of Day 3 *Il27ra*$^{-/-}$ which proved unsuccessful due to issues with DNA library preparation. Library preparations were sequenced using a NovaSeq 6000 S1 Reagent Kit v.1.5 (200 cycles).

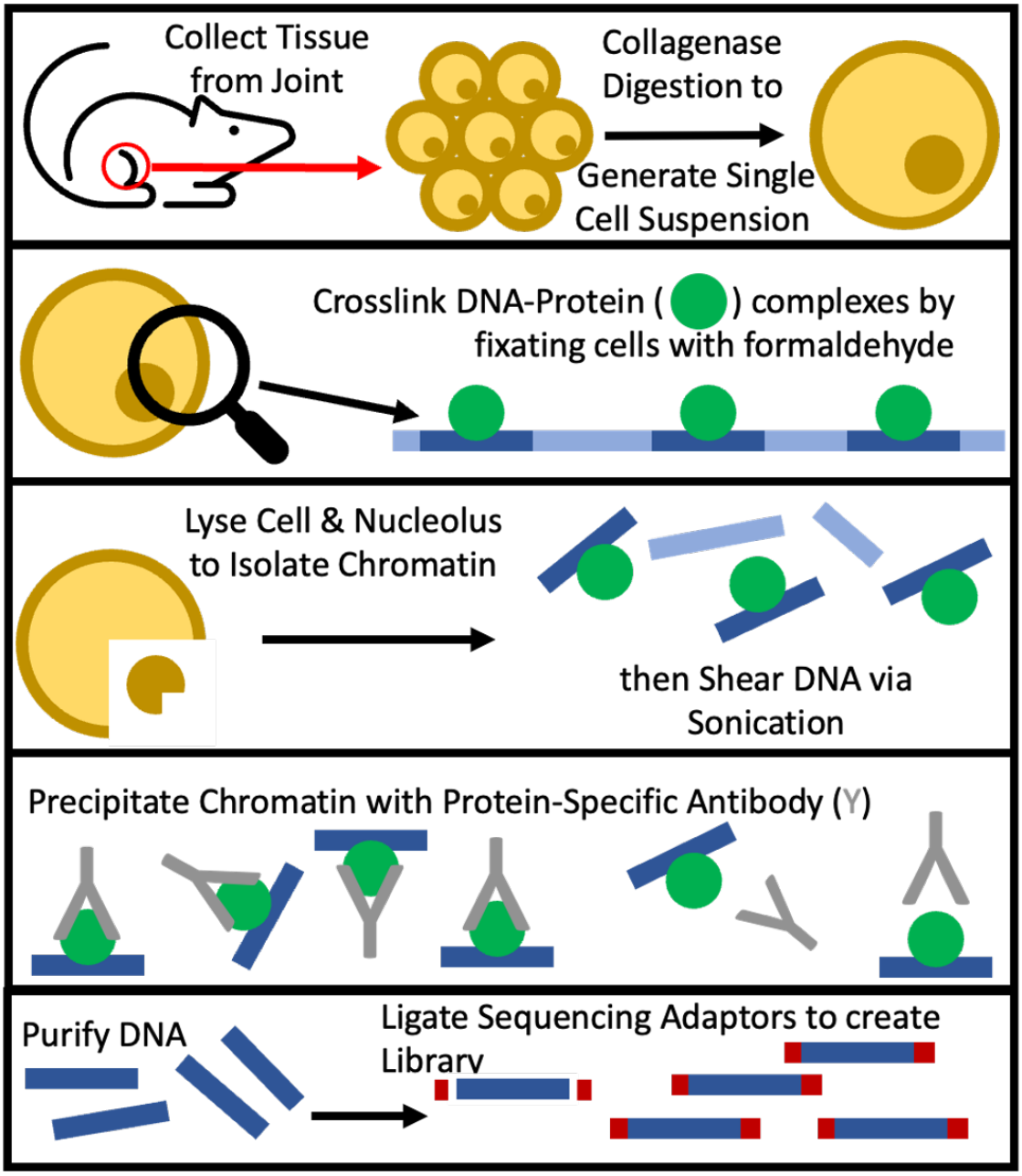


Figure 5.1: Method summary for ChIP-seq analysis of murine synovial tissue. Tissue was harvested from the joints of mice with antigen-induced arthritis and synovial cells recovered following collagenase digestion. Cell isolates were fixed with formaldehyde to cross-link DNA-Protein complexes. Cells and nuclei are lysed, and genomic DNA extracted prior to the disruption of chromatin by sonication. Transcription factors bound to specific chromatin fragments were precipitated using protein-specific antibody. DNA was purified from immunoprecipitated complexes, adaptors ligated for library generation and downstream NGS.

5.2.1 Selection of ChIP Antibodies

Prior to conducting ChIP-seq experiments in mice with AIA, it was necessary to test the efficiency of immunoprecipitation afforded by commercially available ChIP antibodies (detailed in **Chapter 2**, **Sections 2.1.2**, **2.1.3**, **2.1.6** and **2.1.7**). Our prior ChIP-seq analysis of STAT transcription factors used antibodies obtained from Santa Cruz (SANTA CRUZ Stat1 p84/p91 (C-136): sc-464 and SANTA CRUZ Stat3 (C-20): sc-482). However, following alleged Animal Welfare Act violations involving Santa Cruz, we lost access to this STAT1 antibody, necessitating the need for an alternative STAT1 antibody for ChIP. While this legal infringement also affected access to antibodies for STAT3, we had sufficient stocks of the antibody to complete our studies. Splenic lymphocytes were isolated from naïve *Wt* mice and stimulated *in vitro* with 20 ng/ml IL-6 (R&D Systems) for 30 minutes at 37ºC. The ChIP protocol was conducted using several commercial antibodies and their efficacy was assessed via qPCR (ChIP-qPCR). The full method and materials are described in **Chapter 2 Sections 2.1.2, 2.1.3** and **2.1.7**. The results of these initial ChIP experiments are shown in **Figure 5.2**. The antibodies chosen were Stat1 Antibody Cell Signaling #9172 and SANTA CRUZ Stat3 (C-20): sc-482, which our lab was fortunate to have sufficient stock of for these experiments.

5.2.2 Bioinformatic rationale for the analysis of ChIP-seq datasets from *Il6ra*$^{-/-}$ mice

Similar to the normalisation issue encountered with ATAC-seq of synovial tissue from *Il6ra*$^{-/-}$ mice with antigen-induced arthritis (see **Chapter-3**), peak calling of STAT1 and STAT3 ChIP-seq results from *Il6ra*$^{-/-}$ mice led to an over-representation of sequencing peaks judged to be statistically significant. We believe this is in part due to the difficulties in processing *Il6ra*$^{-/-}$ samples. This problem was not observed in datasets from *Wt* or *Il27ra*$^{-/-}$ mice. As the *Il6ra*$^{-/-}$ genotype displays a low inflammatory genotype one knee joint yielded fewer cells whose DNA could be sequenced. This required more animals per experiment than either *Wt* or *Il27ra*$^{-/-}$ mice. This issue is discussed further in the final chapter. Consistent with the method applied to the analysis of ATAC-seq data from *Il6ra*$^{-/-}$ mice, peaks were ranked by peak fold enrichment, a measure of statistical significance. Based on the average number of statistically relevant sequencing peaks identified by ChIP-seq of samples from *Wt* and

Il27ra$^{-/-}$ mice, analysis of datasets from *Il6ra*$^{-/-}$ mice was restricted to the top 350 sequencing peaks (the total number of peaks in each condition can be seen by the black bars in **Figure 5.3**).

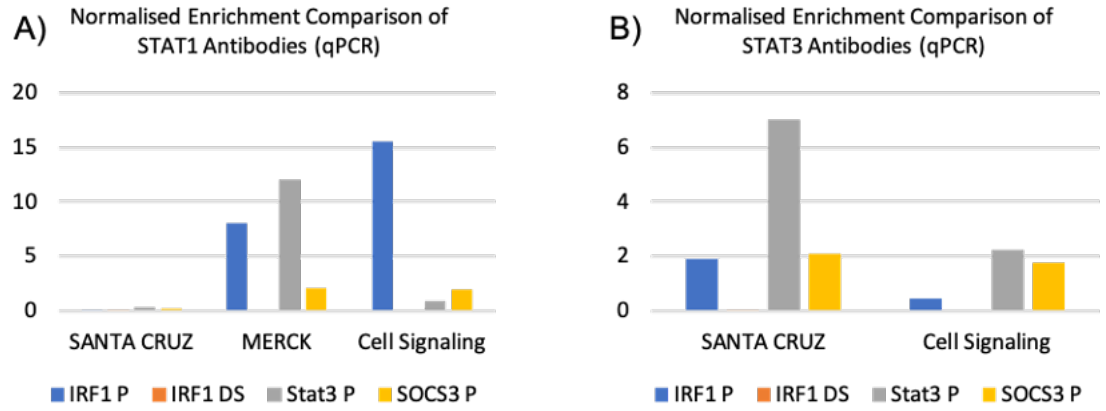


Figure 5.2 ChIP Antibody Selection via qPCR Normalised Enrichment Analysis. Interferon Regulatory Factor 1 Promoter (IRF1 P) is a positive indicator of STAT1 enrichment. IRF1 downstream (IRF1DS) is a negative indicator of STAT1 enrichment. Both suppressor of cytokine signalling 3 promotor (SOCS3P) and STAT3P are positive indicators of STAT3 enrichment. Samples were first normalised against input control samples, then IgG. The IgG control is an indicator of "background" antibody binding as opposed to the specialised binding of STAT1 or STAT3 antibodies. A) Normalised Enrichment Comparison of STAT1 Antibodies (SANTA CRUZ Stat1 p84/p91 (C-136): sc-464) (MERCK Anti-STAT1, CT (rabbit polyclonal IgG) (Cell Signaling STAT1 Antibody #9172). B) Normalised Enrichment Comparison of STAT3 Antibodies (SANTA CRUZ Stat3 (C-20): sc-482) (Cell Signaling Stat3 (79D7) Rabbit mAb #4904).

5.3 Results

5.3.1 Genomic distribution of STAT transcription factors as identified by ChIP-seq of tissues from AIA-challenged mice

All ChIP-seq datasets showed sequencing peaks mapped to specific genomic regions. These peaks were located in proximal (≤200 bp from a transcription start site; TSS), distal (<200 bp from a TSS and comprising introns, exons and stop codons) or intergenic regions (these not being assigned to a specific gene) (**Figure 5.3**).

As explained in **Section 5.2.2** the rationalisation of the *Il6ra*$^{-/-}$ ChIP-seq data utilised the average number of peaks from all other samples. Therefore, in this genotype STAT1 and STAT3 activity, as measured by peak number, cannot be compared to *Wt* and *Il27ra*$^{-/-}$ samples. However, comparisons between *Wt* and *Il27ra*$^{-/-}$ samples can still be made.

Comparing STAT1 transcription factor binding in synovial extracts from *Wt* and *Il27ra*$^{-/-}$ mice with antigen-induced arthritis (analysed on Day 10) there was a reduced number of sequencing peaks recorded in *Il27ra*$^{-/-}$ mice. This reduction in STAT1 was, however, counterbalanced by an increase in the number of sequencing peaks linked with STAT3. **Figure 5.3** shows that despite the relative changes in activity described, STAT1 activity in the *Wt* conditions is greater than the STAT3 activity in the *Il27ra*$^{-/-}$. In essence, the greatest level of STAT1 activity exceeds that of STAT3.

The subsidiary graph in **Figure 5.3** is a supplementary figure. It shows the percentage of peaks associated with each region (samples are in the same descending order in both the main and supplemental figures). The graph indicates that the majority (~60%) of ChIP-seq peaks fall into intergenic regions. Furthermore, in almost all conditions, peaks in proximal regions are in the minority. This lack of binding in a TSS region is possibly indicative of STAT1 and STAT3 binding to other regulatory or enhancer elements.

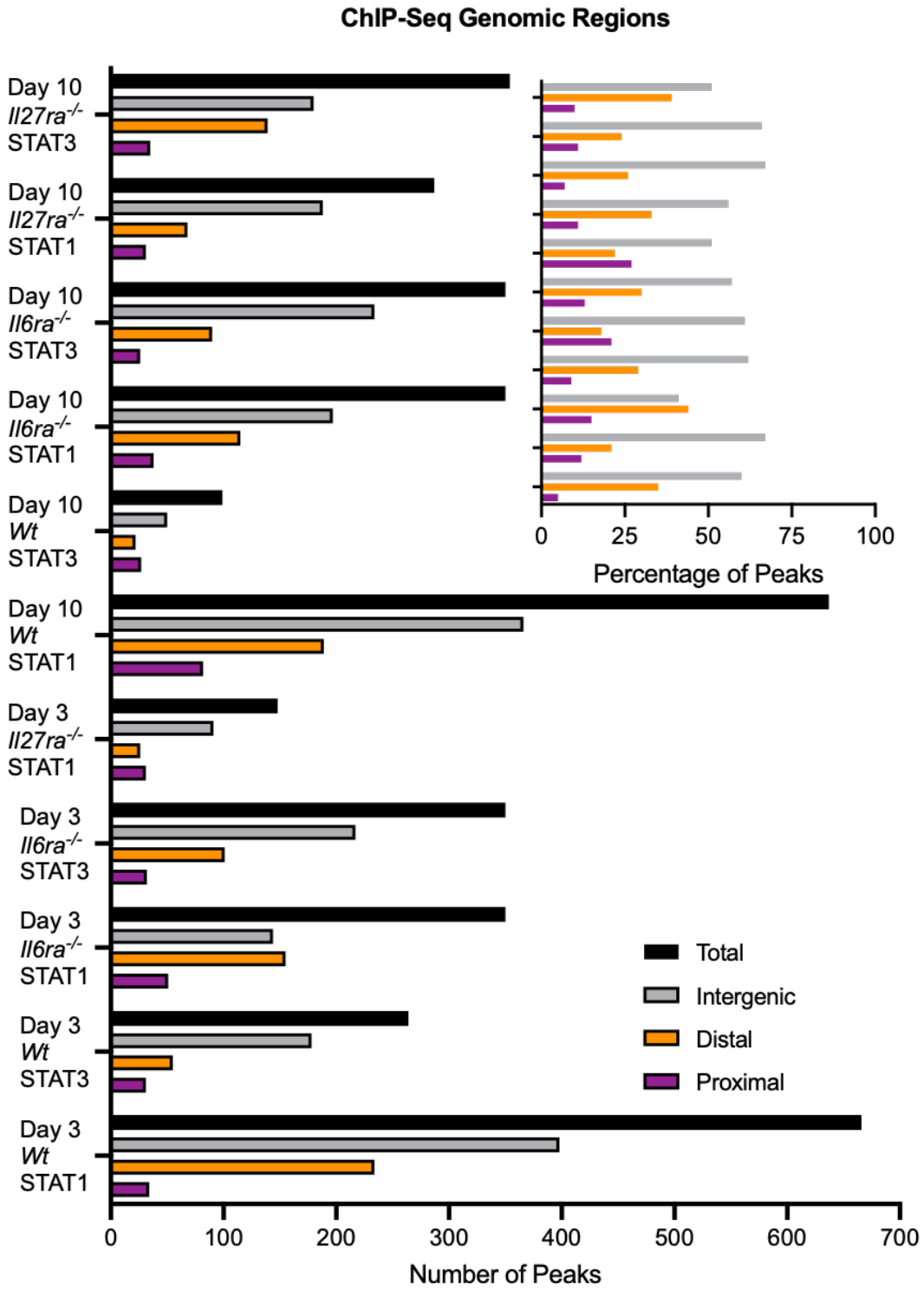


Figure 5.3: The distribution of STAT1 and STAT3 ChIP-seq peaks across Proximal, Distal and Intergenic Regions. Sequencing peaks were mapped to genomic regions as part of the normalisation pipeline. Proximal peaks are defined as ≤200 bp distant of a TSS. Distal peaks are defined as >200bp distant of a TSS. They are comprised of introns, exons and stop codons. Intergenic peaks are not associated with any gene. The principal graph of the figure shows the number of peaks in each region as well as the total for each condition. The subsidiary graph shows the percentage of peaks in each region. The conditions are in the same order in both graphs.

5.3.2 Chromosomal localisation of STAT transcription factors

Consistent with the analysis of ATAC-seq datasets presented previously (**Chapter 3 Section 3.3.2)**, ChIP-seq peaks were mapped to chromosomes (1-19, X, Y and those peaks which are not associated (NA) with a particular chromosome). It was not expected that every chromosome should be represented in every condition. Instead, a transcription factor may have binding activity in all, most or very few chromosomes depending on the condition in question. **Figure 5.4** shows the percentage of sequencing peaks aligned to each chromosome in *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ mice from STAT1 and STAT3 experiments. In addition to the total peaks, the figure includes the percentage of gene-associated peaks (GAPs) on a given chromosome.

5.3.3 ChIP-seq Identified Gene Accessibility

To explore the relationship between transcription factor binding and chromatin accessibility genes identified by ChIP-seq experiments were compared to their timepoint and genotype counterparts in ATAC-seq experiments. The proportion of ChIP-seq identified genes also present in the ATAC-seq gene lists (described in **Chapter 3 Section 3.2.1**) are displayed in **Figure 5.5.** It may be somewhat surprising that not all genes identified in ChIP-seq should also be accessible, in fact, the minority of these genes exhibit accessibility in the ATAC-seq data. This speaks to the highly dynamic nature of the epigenetic landscape of synovitis. It must also be remembered, however, that ATAC-seq is a measure of Chromatin's accessibility by Tn5, which is not analogous to STAT1 or STAT3. As discussed in **Chapter 1 Section 1.6.1** chromatin accessibility is not binary in nature but rather a continuum broadly categorised as closed chromatin, permissive chromatin, and open chromatin (Klemm et al. 2019). It is not entirely certain where on this continuum Tn5, STAT1 and STAT3 sit in relation to one another.

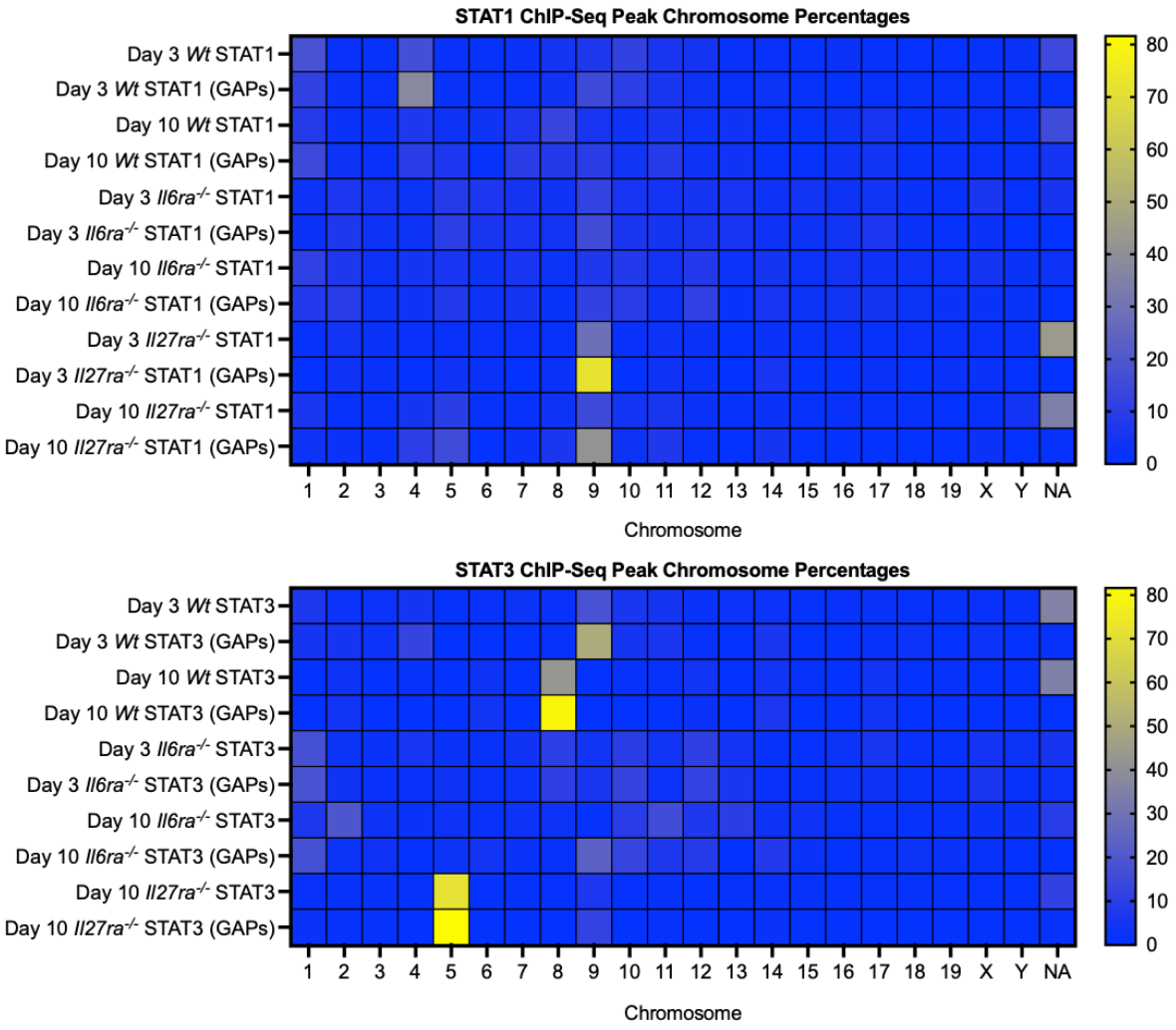


Figure 5.4: The Percentage of Sequencing Peaks Associated with Particular Chromosomes in ChIP-seq Conditions. The percentage of peaks on a given chromosome (1-19, X, Y and those which are not associated (NA) with any chromosome) was calculated. Both total peak and gene associated peak (GAPs) percentages are shown in the figure. STAT1 is shown in the top heatmap and STAT3 is shown in the bottom.

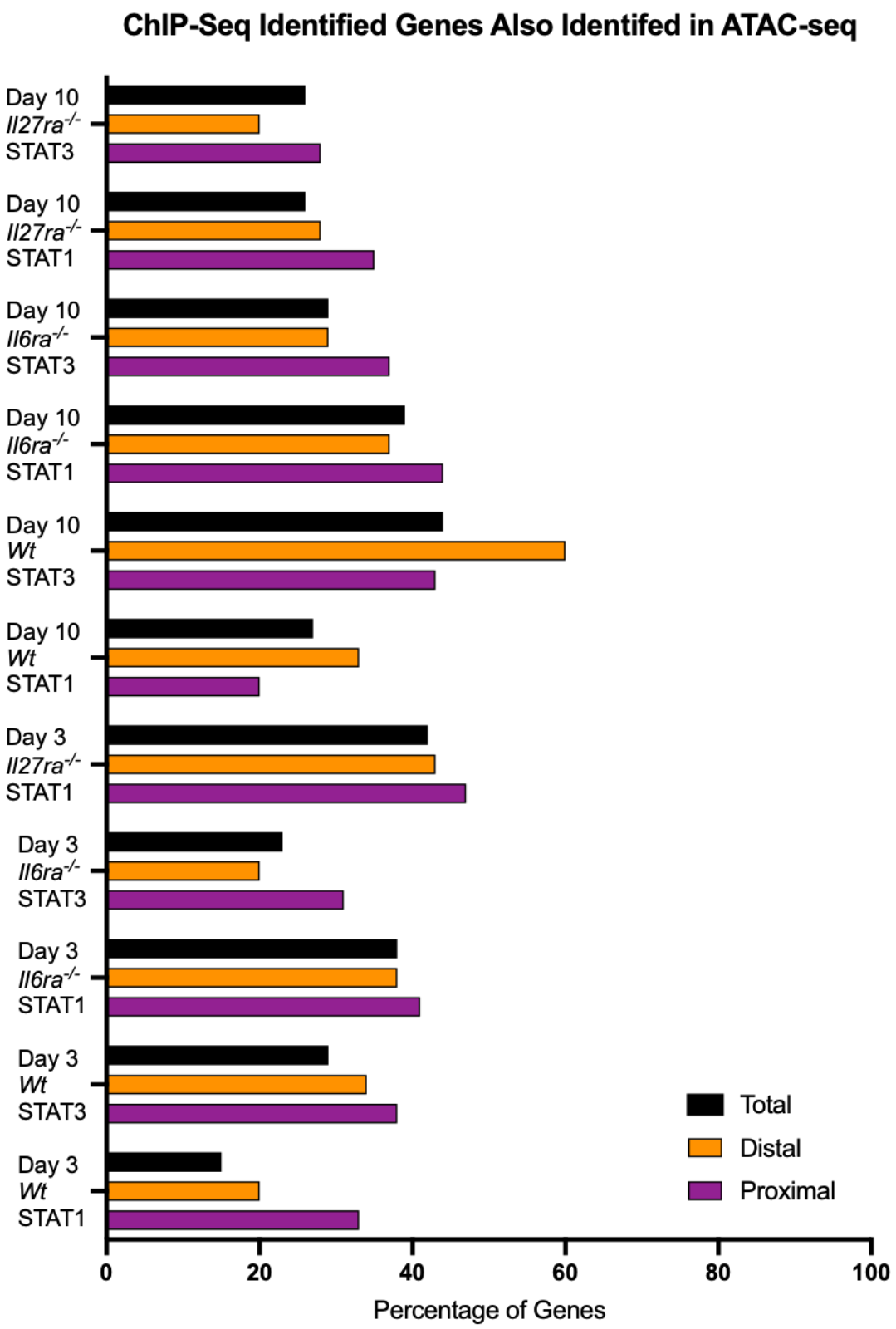


Figure 5.5: The Percentage of Genes Identified in ChIP-seq Experiments Exhibiting Accessibility In ATAC-seq Experiments. Proximal and Distal percentages represent the percentage of that gene subset which are also found in ATAC-seq. Proximal peaks are defined as ≤200 bp distant of a TSS. Distal peaks are defined as >200bp distant of a TSS. They are comprised of introns, exons and stop codons.

5.3.4 Ingenuity Pathway Analysis (IPA) of genes identified in ChIP-seq experiments.

Bioinformatic analysis using prediction tools was next used to understand how the cytokine control of STAT1 and STAT3 affected biological processes linked to synovitis (**Figure 5.6**, **5.7** and **5.8**). To perform this analysis, ChIP-seq gene list datasets for STAT1 and STAT3 were subjected to IPA analysis to identify canonical pathways (**Figure 5.6**), diseases and functions (**Figure 5.7**), and Upstream regulators (**Figure 5.8**). Each condition was first subjected to IPA's 'Core Analysis' using the default settings. All STAT1 and STAT3 datasets then underwent IPA's 'Comparison Analysis', again using the default settings. The pathways, regulators and diseases identified are ranked by their $-\log_{10}$ p-value score. Each individual heatmap shows statistics for the top 50 results. Panel **A** of each of these three figures shows the results for STAT1 at Day 3, panel **B** shows STAT3 at Day 3, panel **C** shows STAT1 at Day 10 and panel **D** shows STAT3 at Day 10.

As can be seen in the heatmaps, deletion of the IL-6 or IL-27 receptors and thus inhibition of those cytokines' signalling pathways resulted in changes to STAT1 and STAT3's behaviour. These changes to the genes to which the transcription factors bound ultimately result in changes to the pathophysiology of the synovium. The nature of these changes and their biological significance are described further in the discussion section of this chapter.

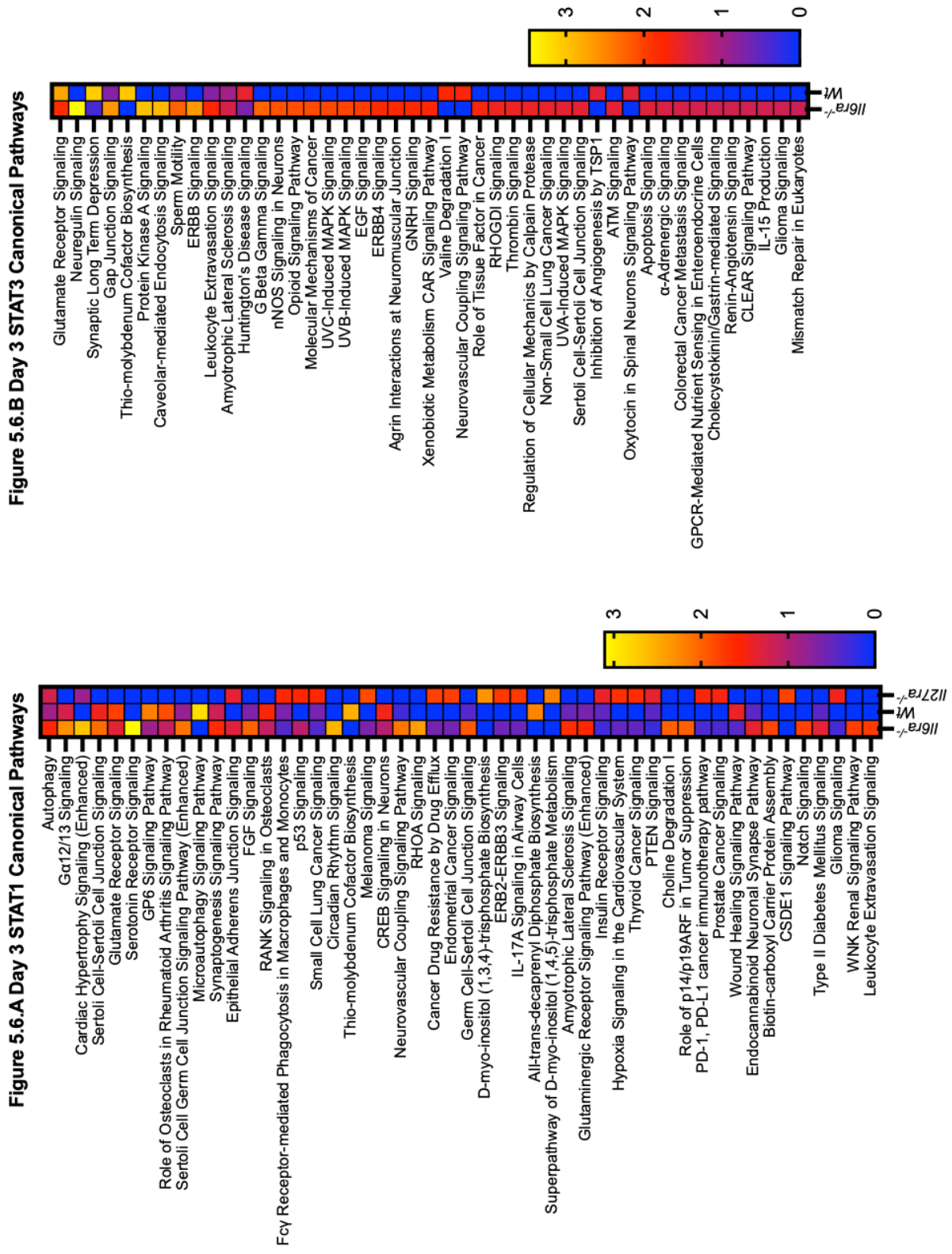


Figure 5.6.A Day 3 STAT1 Canonical Pathways
Autophagy
Gα12/13 Signaling
Cardiac Hypertrophy Signaling (Enhanced)
Sertoli Cell-Sertoli Cell Junction Signaling
Glutamate Receptor Signaling
Serotonin Receptor Signaling
GP6 Signaling Pathway
Role of Osteoclasts in Rheumatoid Arthritis Signaling Pathway
Sertoli Cell Germ Cell Junction Signaling Pathway (Enhanced)
Microautophagy Signaling Pathway
Synaptogenesis Signaling Pathway
Epithelial Adherens Junction Signaling
FGF Signaling
RANK Signaling in Osteoclasts
Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes
p53 Signaling
Small Cell Lung Cancer Signaling
Circadian Rhythm Signaling
Thio-molybdenum Cofactor Biosynthesis
Melanoma Signaling
CREB Signaling in Neurons
Neurovascular Coupling Signaling Pathway
RHOA Signaling
Cancer Drug Resistance by Drug Efflux
Endometrial Cancer Signaling
Germ Cell-Sertoli Cell Junction Signaling
D-myo-inositol (1,3,4)-trisphosphate Biosynthesis
ERB2-ERBB3 Signaling
IL-17A Signaling in Airway Cells
All-trans-decaprenyl Diphosphate Biosynthesis
Superpathway of D-myo-inositol (1,4,5)-trisphosphate Metabolism
Amyotrophic Lateral Sclerosis Signaling
Glutaminergic Receptor Signaling Pathway (Enhanced)
Insulin Receptor Signaling
Hypoxia Signaling in the Cardiovascular System
Thyroid Cancer Signaling
PTEN Signaling
Choline Degradation I
Role of p14/p19ARF in Tumor Suppression
PD-1, PD-L1 cancer immunotherapy pathway
Prostate Cancer Signaling
Wound Healing Signaling Pathway
Endocannabinoid Neuronal Synapse Pathway
Biotin-carboxyl Carrier Protein Assembly
CSDE1 Signaling Pathway
Notch Signaling
Type II Diabetes Mellitus Signaling
Glioma Signaling
WNK Renal Signaling Pathway
Leukocyte Extravasation Signaling
Il27ra-/-
Wt
Il6ra-/-
3
2
1
0
Figure 5.6.B Day 3 STAT3 Canonical Pathways
Glutamate Receptor Signaling
Neuregulin Signaling
Synaptic Long Term Depression
Gap Junction Signaling
Thio-molybdenum Cofactor Biosynthesis
Protein Kinase A Signaling
Caveolar-mediated Endocytosis Signaling
Sperm Motility
ERBB Signaling
Leukocyte Extravasation Signaling
Amyotrophic Lateral Sclerosis Signaling
Huntington's Disease Signaling
G Beta Gamma Signaling
nNOS Signaling in Neurons
Opioid Signaling Pathway
Molecular Mechanisms of Cancer
UVC-Induced MAPK Signaling
UVB-Induced MAPK Signaling
EGF Signaling
ERBB4 Signaling
Agrin Interactions at Neuromuscular Junction
GNRH Signaling
Xenobiotic Metabolism CAR Signaling Pathway
Valine Degradation I
Neurovascular Coupling Signaling Pathway
Role of Tissue Factor in Cancer
RHOGDI Signaling
Thrombin Signaling
Regulation of Cellular Mechanics by Calpain Protease
Non-Small Cell Lung Cancer Signaling
UVA-Induced MAPK Signaling
Sertoli Cell-Sertoli Cell Junction Signaling
Inhibition of Angiogenesis by TSP1
ATM Signaling
Oxytocin in Spinal Neurons Signaling Pathway
Apoptosis Signaling
α-Adrenergic Signaling
Colorectal Cancer Metastasis Signaling
GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells
Cholecystokinin/Gastrin-mediated Signaling
Renin-Angiotensin Signaling
CLEAR Signaling Pathway
IL-15 Production
Glioma Signaling
Mismatch Repair in Eukaryotes
Il6ra-/-
Wt
3
2
1
0

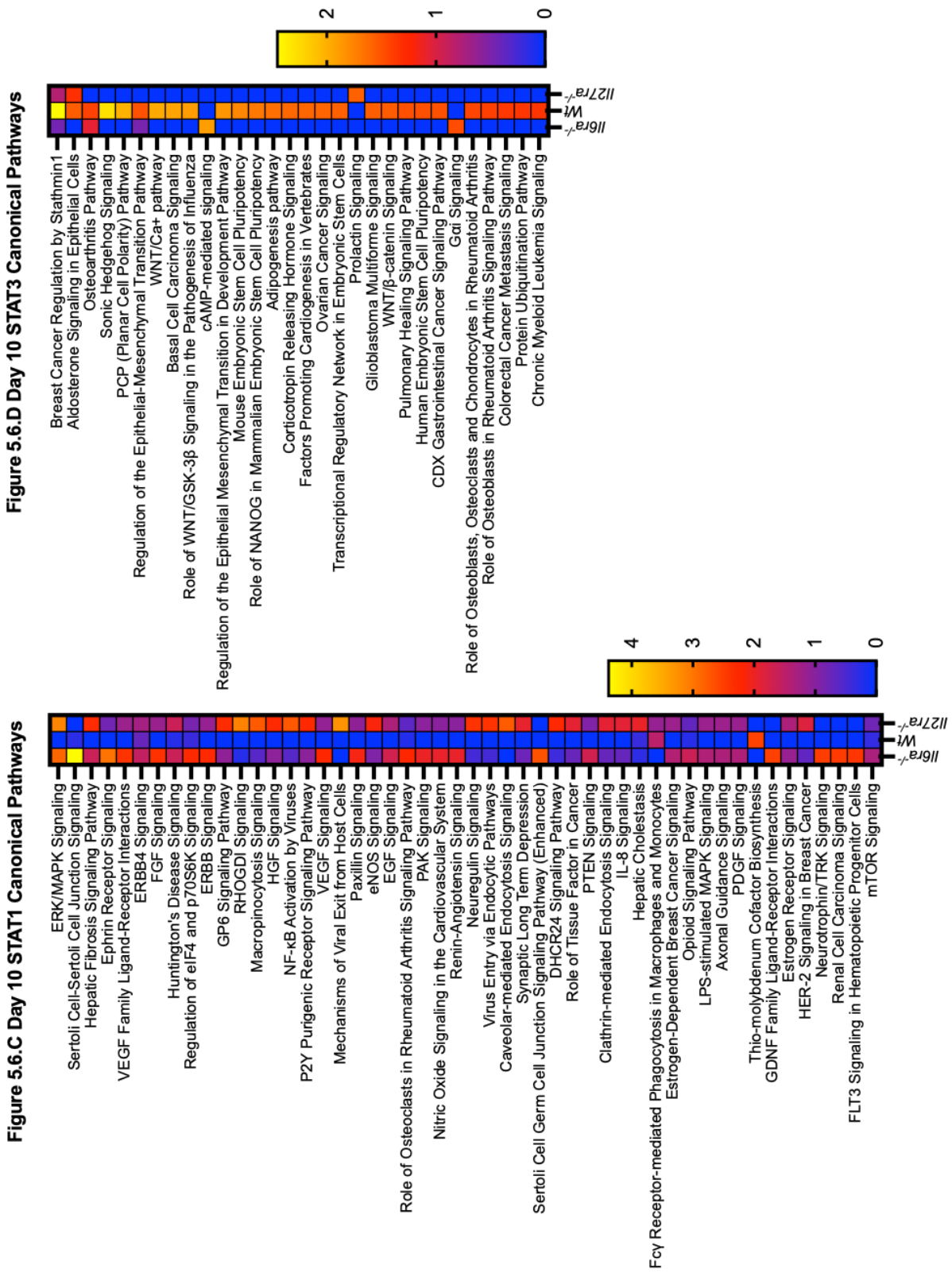


Figure 5.6: Canonical Pathways Identified by IPA Comparative Analysis of ChIP-seq Conditions of STAT1 and STAT3 at Days 3 and 10. The genes identified by ChIP-seq experiments were submitted to IPA for Core then Comparative analysis. The top 50 canonical pathways are exhibited in these heatmaps (note in **B&D** fewer than 50 were identified). They were ranked by their $-\log_{10}$ p-value score which is displayed in the heatmap.

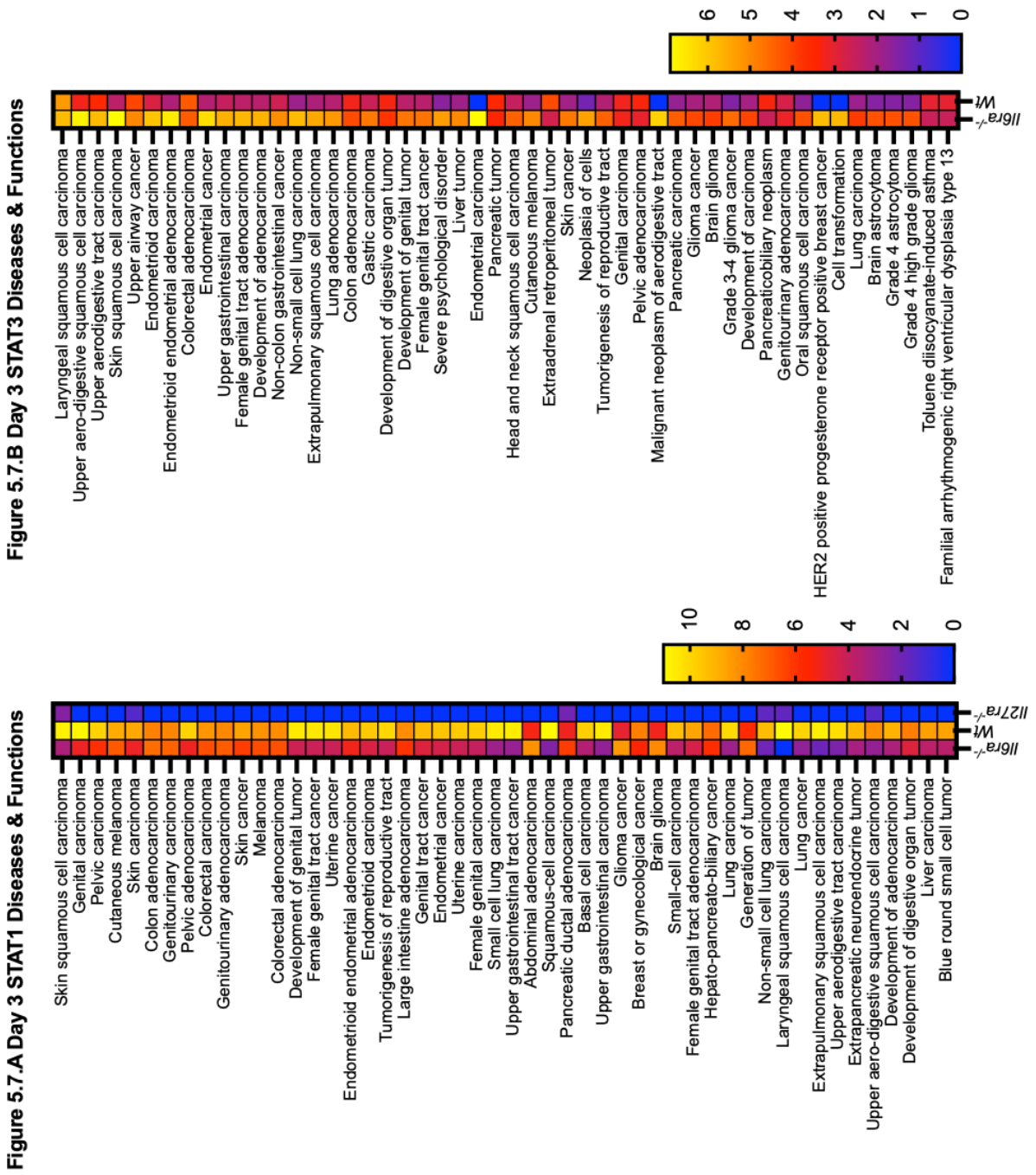


Figure 5.7.A Day 3 STAT1 Diseases & Functions
Skin squamous cell carcinoma
Genital carcinoma
Pelvic carcinoma
Cutaneous melanoma
Skin carcinoma
Colon adenocarcinoma
Genitourinary carcinoma
Pelvic adenocarcinoma
Colorectal carcinoma
Genitourinary adenocarcinoma
Skin cancer
Melanoma
Colorectal adenocarcinoma
Development of genital tumor
Female genital tract cancer
Uterine cancer
Endometrioid adenocarcinoma
Endometrioid carcinoma
Tumorigenesis of reproductive tract
Large intestine adenocarcinoma
Genital tract cancer
Endometrial cancer
Uterine carcinoma
Female genital carcinoma
Small cell lung carcinoma
Upper gastrointestinal tract cancer
Abdominal adenocarcinoma
Squamous-cell carcinoma
Pancreatic ductal adenocarcinoma
Basal cell carcinoma
Upper gastrointestinal carcinoma
Glioma cancer
Breast or gynecological cancer
Brain glioma
Small-cell carcinoma
Female genital tract adenocarcinoma
Hepato-pancreato-biliary cancer
Lung carcinoma
Generation of tumor
Non-small cell lung carcinoma
Laryngeal squamous cell carcinoma
Lung cancer
Extrapulmonary squamous cell carcinoma
Upper aerodigestive tract carcinoma
Extrapancreatic neuroendocrine tumor
Upper aero-digestive squamous cell carcinoma
Development of adenocarcinoma
Development of digestive organ tumor
Liver carcinoma
Blue round small cell tumor
Il27ra-/-
Wt
Il6ra-/-
0
2
4
6
8
10
Figure 5.7.B Day 3 STAT3 Diseases & Functions
Laryngeal squamous cell carcinoma
Upper aero-digestive squamous cell carcinoma
Upper aerodigestive tract carcinoma
Skin squamous cell carcinoma
Upper airway cancer
Endometrioid carcinoma
Endometrioid endometrial adenocarcinoma
Colorectal adenocarcinoma
Endometrial cancer
Upper gastrointestinal carcinoma
Female genital tract adenocarcinoma
Development of adenocarcinoma
Non-colon gastrointestinal cancer
Non-small cell lung carcinoma
Extrapulmonary squamous cell carcinoma
Lung adenocarcinoma
Colon adenocarcinoma
Gastric carcinoma
Development of digestive organ tumor
Development of genital tumor
Female genital tract cancer
Severe psychological disorder
Liver tumor
Endometrial carcinoma
Pancreatic tumor
Head and neck squamous cell carcinoma
Cutaneous melanoma
Extraadrenal retroperitoneal tumor
Skin cancer
Neoplasia of cells
Tumorigenesis of reproductive tract
Genital carcinoma
Pelvic adenocarcinoma
Malignant neoplasm of aerodigestive tract
Pancreatic carcinoma
Glioma cancer
Brain glioma
Grade 3-4 glioma cancer
Development of carcinoma
Pancreaticobiliary neoplasm
Genitourinary adenocarcinoma
Oral squamous cell carcinoma
HER2 positive progesterone receptor positive breast cancer
Cell transformation
Lung carcinoma
Brain astrocytoma
Grade 4 astrocytoma
Grade 4 high grade glioma
Toluene diisocyanate-induced asthma
Familial arrhythmogenic right ventricular dysplasia type 13
Il6ra-/-
Wt
0
1
2
3
4
5
6

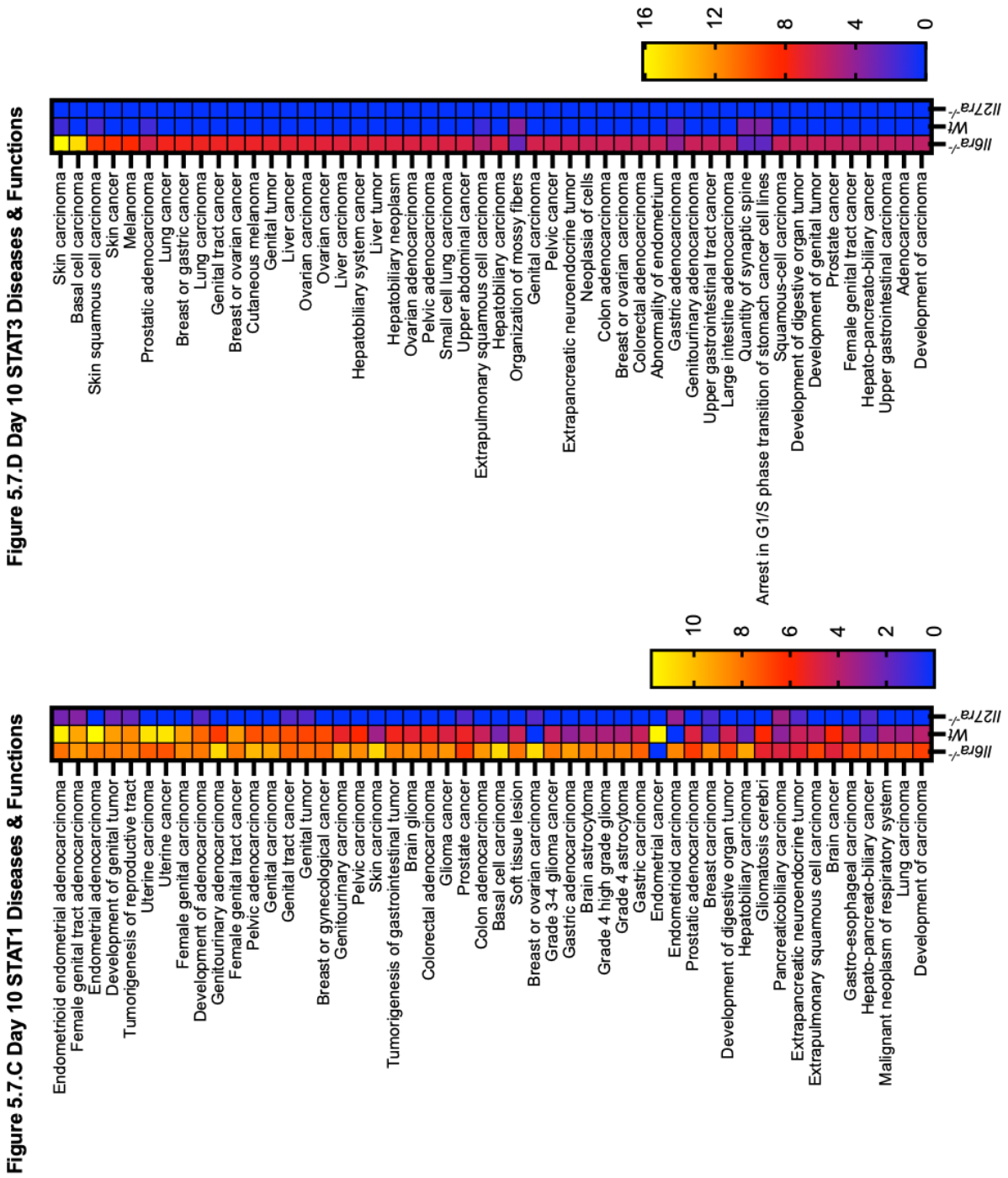


Figure 5.7: Diseases & Functions Identified by IPA Comparative Analysis of ChIP-seq Conditions of STAT1 and STAT3 at Days 3 and 10. The genes identified by ChIP-seq experiments were submitted to IPA for Core then Comparative analysis. The top 50 disease and functions are exhibited in these heatmaps. They were ranked by their $-\log_{10}$ p-value score which is displayed in the heatmap.

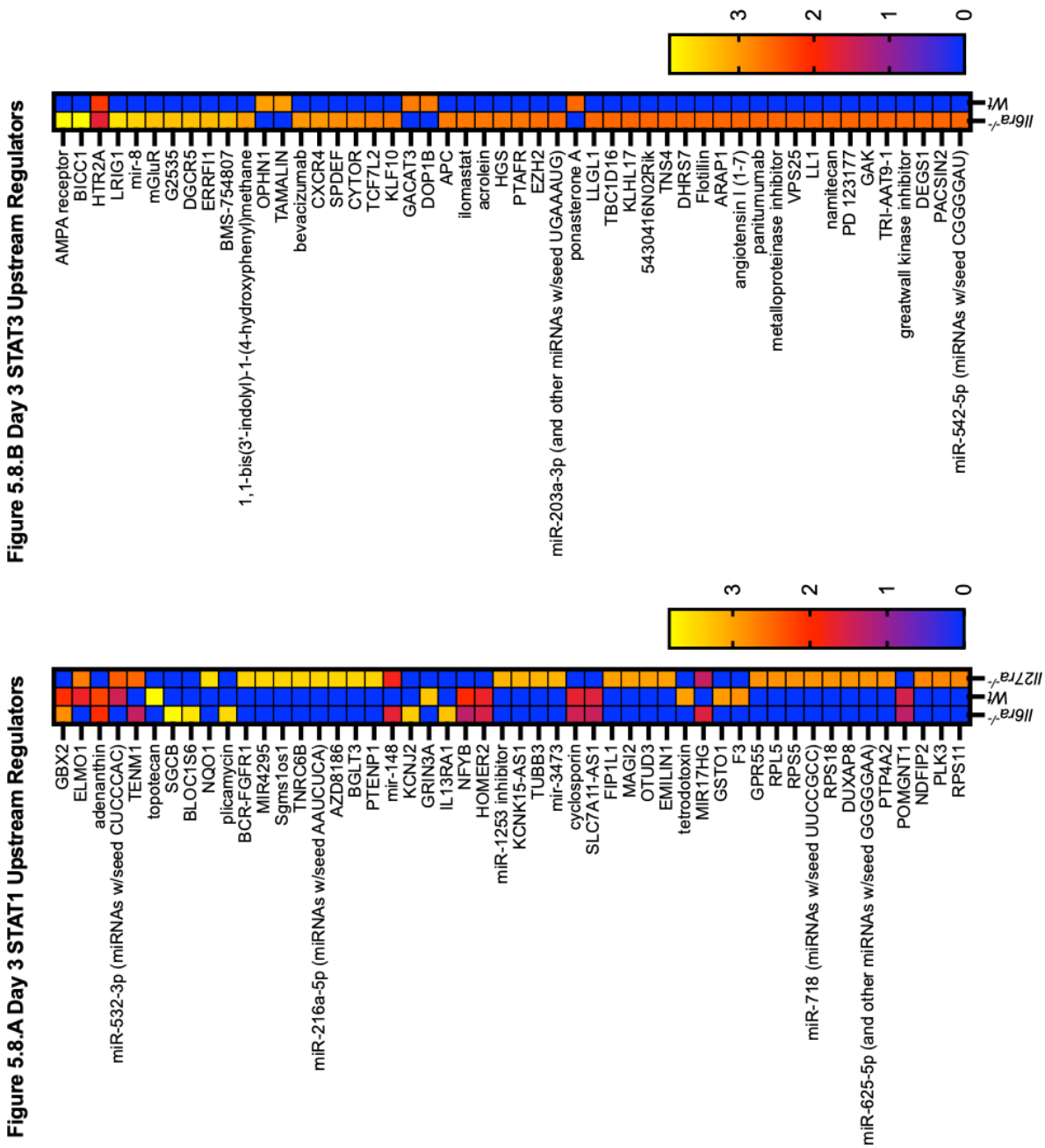


Figure 5.8.B Day 3 STAT3 Upstream Regulators
AMPA receptor
BICC1
HTR2A
LRIG1
mir-8
mGluR
G2535
DGCR5
ERRFI1
BMS-754807
1,1-bis(3'-indolyl)-1-(4-hydroxyphenyl)methane
OPHN1
TAMALIN
bevacizumab
CXCR4
SPDEF
CYTOR
TCF7L2
KLF10
GACAT3
DOP1B
APC
ilomastat
acrolein
HGS
PTAFR
EZH2
miR-203a-3p (and other miRNAs w/seed UGAAAUG)
ponasterone A
LLGL1
TBC1D16
KLHL17
5430416N02Rik
TNS4
DHRS7
Flotillin
ARAP1
angiotensin I (1-7)
panitumumab
metalloproteinase inhibitor
VPS25
LL1
namitecan
PD 123177
GAK
TRI-AAT9-1
greatwall kinase inhibitor
DEGS1
PACSIN2
miR-542-5p (miRNAs w/seed CGGGGAU)
3
2
1
0
Il6ra-/-
Wt
Figure 5.8.A Day 3 STAT1 Upstream Regulators
GBX2
ELMO1
adenanthin
miR-532-3p (miRNAs w/seed CUCCCAC)
TENM1
topotecan
SGCB
BLOC1S6
NQO1
plicamycin
BCR-FGFR1
MIR4295
Sgms1os1
TNRC6B
miR-216a-5p (miRNAs w/seed AAUCUCA)
AZD8186
BGLT3
PTENP1
mir-148
KCNJ2
GRIN3A
IL13RA1
NFYB
HOMER2
miR-1253 inhibitor
KCNK15-AS1
TUBB3
mir-3473
cyclosporin
SLC7A11-AS1
FIP1L1
MAGI2
OTUD3
EMILIN1
tetrodotoxin
MIR17HG
GSTO1
F3
GPR55
RPL5
RPS5
miR-718 (miRNAs w/seed UUCCGCC)
RPS18
DUXAP8
miR-625-5p (and other miRNAs w/seed GGGGGAA)
PTP4A2
POMGNT1
NDFIP2
PLK3
RPS11
3
2
1
0
Il27ra-/-
Wt
Il6ra-/-

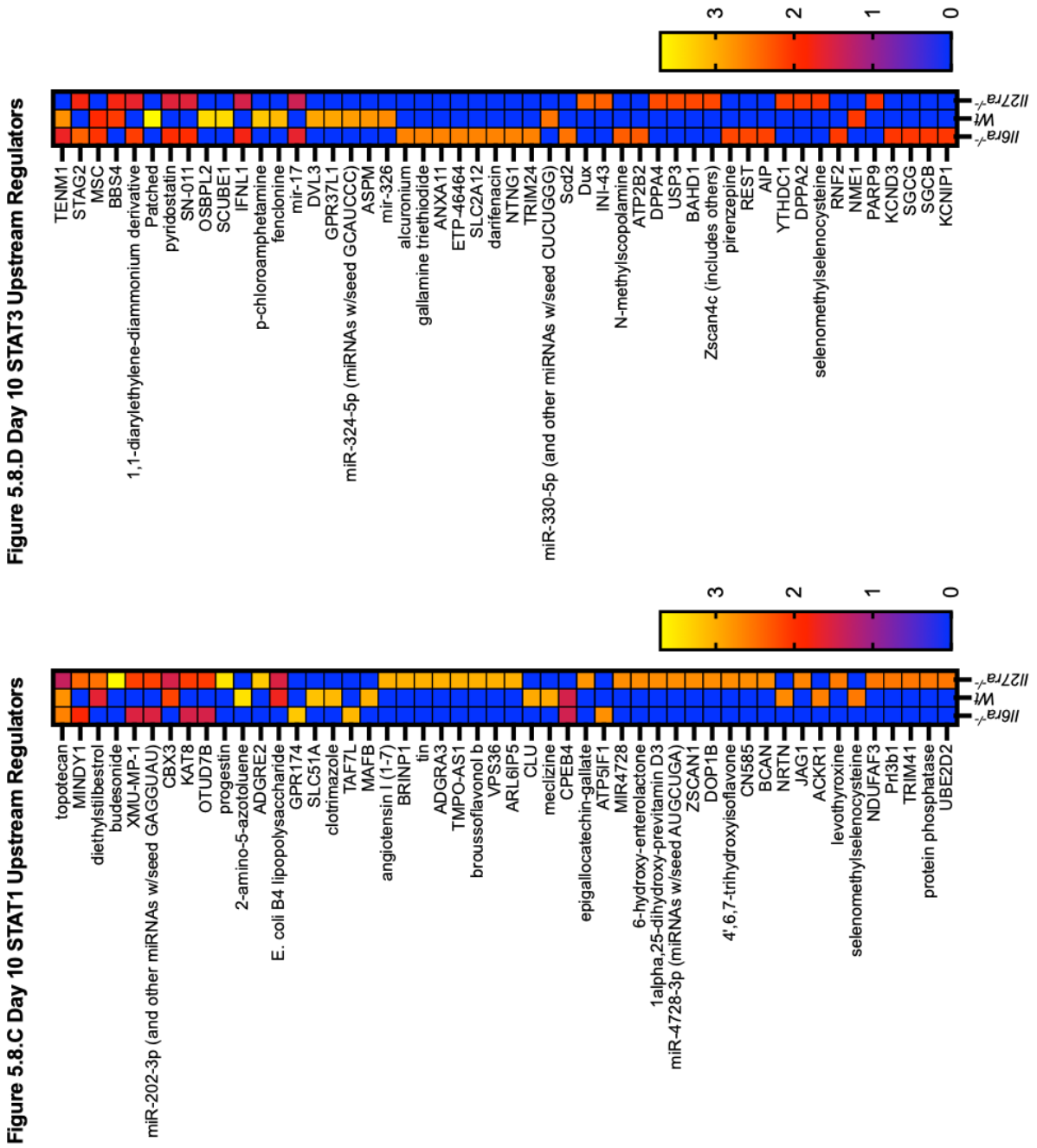


Figure 5.8: Upstream Regulators Identified by IPA Comparative Analysis of ChIP-seq Conditions of STAT1 and STAT3 at Days 3 and 10. The genes identified by ChIP-seq experiments were submitted to IPA for Core then Comparative analysis. The top 50 upstream regulators are exhibited in these heatmaps. They were ranked by their $-\log_{10}$ p-value score which is displayed in the heatmap.

5.3.5 Comparative analysis of genotype-specific and shared binding of STAT1 and STAT3 in mice with AIA

The Venn diagram tool (https://bioinformatics.psb.ugent.be/webtools/Venn/) was used to identify common genomic regions occupied by STAT1 and/or STAT3 in all mouse genotypes. Utilising the same approach described in **Chapter 3, Section 3.3.4** this analysis provides an opportunity to assess the potential signalling dynamics between STAT1 and STAT3, and differences in the biological properties of IL-6 and IL-27.

This analysis focuses on gene-associated peaks only. Both proximally and distally associated genes are included. These are the same regions described in **Section 5.3.1**. duplicated genes (those represented by two or more peaks in one condition) were removed hence it will be noticeable that the sum of Proximal and Distal genes exceeds that of Total genes. This is because often a gene can be represented by multiple peaks which are in differing genomic regions. The pie charts in **Figure 5.9**, **Figure 5.10** and **Figure 5.11** show the proportions of genes identified that display STAT1 or STAT3 binding. These figures provide observations that raise important biological questions about the regulation of synovitis which are discussed later.

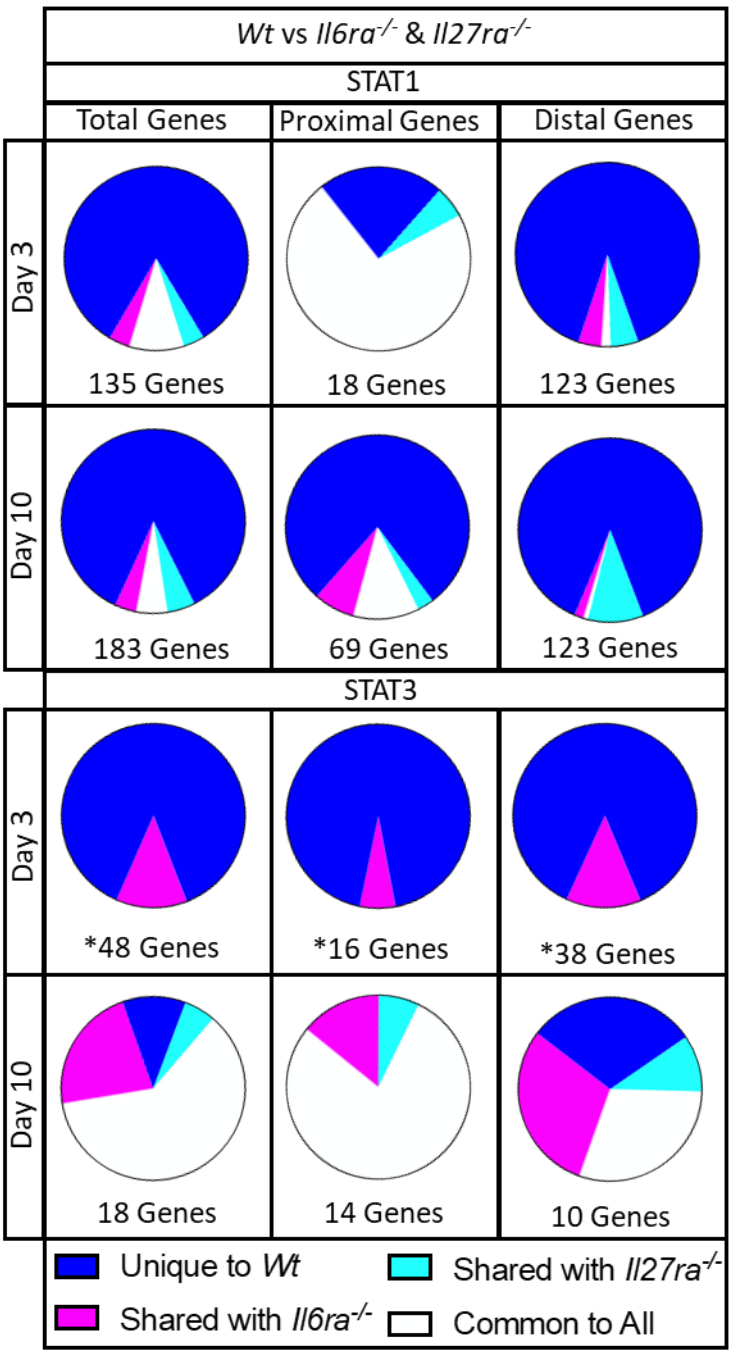


Figure 5.9: The Proportion of Genes Bound to STAT1 or STAT3 in the *Wt* Condition Versus Other Genotypes. The genes bound to STAT1 or STAT3 in the *Wt* genotype at Days 3 and 10 were compared with the *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$ to determine which are uniquely bound in *Wt* genotype, which are bound in both the *Wt* genotype and either *Il6ra*$^{-/-}$ or *Il27ra*$^{-/-}$ and which genes are bound in all genotypes. The proportion of genes in these four categories in relation to the *Wt* genotype are shown in pie charts. This analysis was carried out in all genes (left), proximally associated genes (≤200 bp from a transcription start site; TSS) (middle) and distally associated genes (<200 bp from a TSS and comprising introns, exons and stop codons) (right).

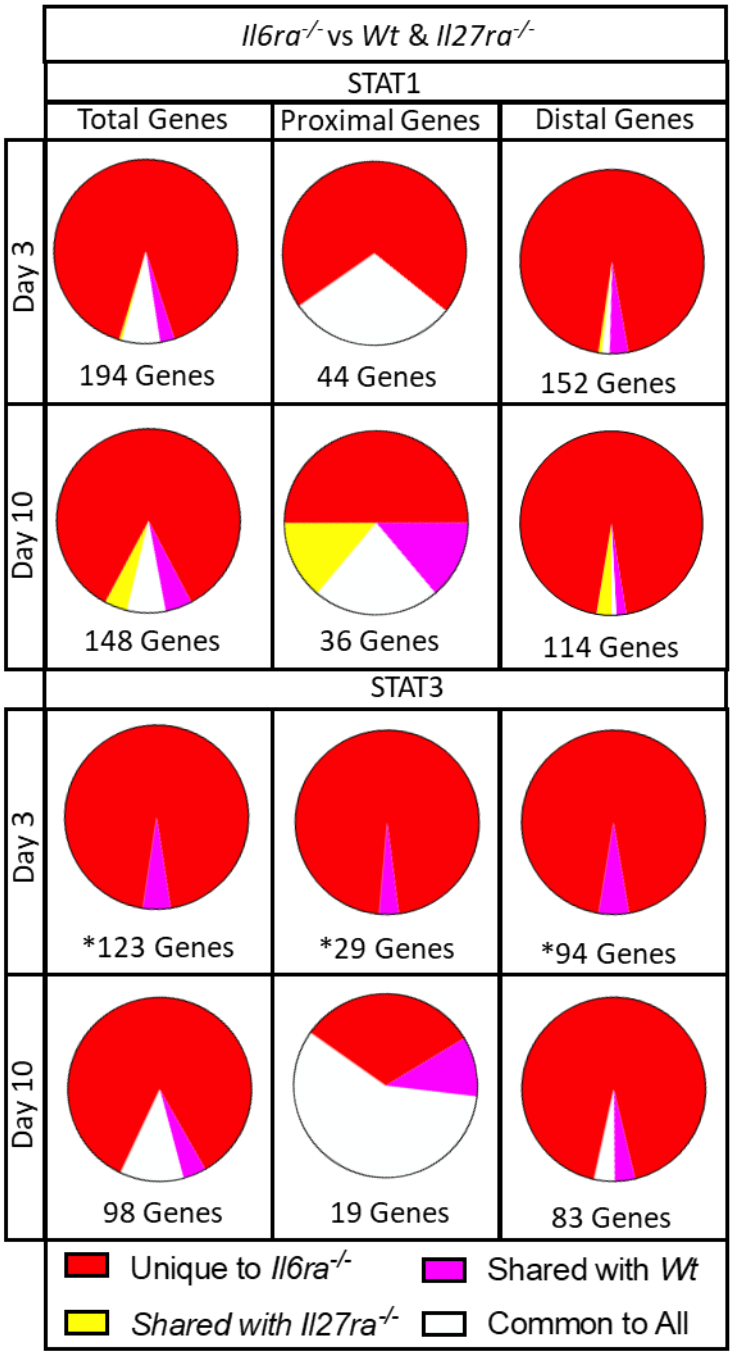


Figure 5.10: The Proportion of Genes Bound to STAT1 or STAT3 in the *Il6ra*$^{-/-}$ Condition Versus Other Genotypes. The genes bound to STAT1 or STAT3 in the Il6ra$^{-/-}$ genotype at Days 3 and 10 were compared with the Wt and Il27ra$^{-/-}$ to determine which are uniquely bound in Il6ra$^{-/-}$ genotype, which are bound in both the Il6ra$^{-/-}$ genotype and either Wt OR Il27ra$^{-/-}$ and which genes are bound in all genotypes. The proportion of genes in these four categories in relation to the Il6ra$^{-/-}$ genotype are shown in pie charts. This analysis was carried out in all genes (left), proximally associated genes (≤200 bp from a transcription start site; TSS) (middle) and distally associated genes (<200 bp from a TSS and comprising introns, exons and stop codons) (right).

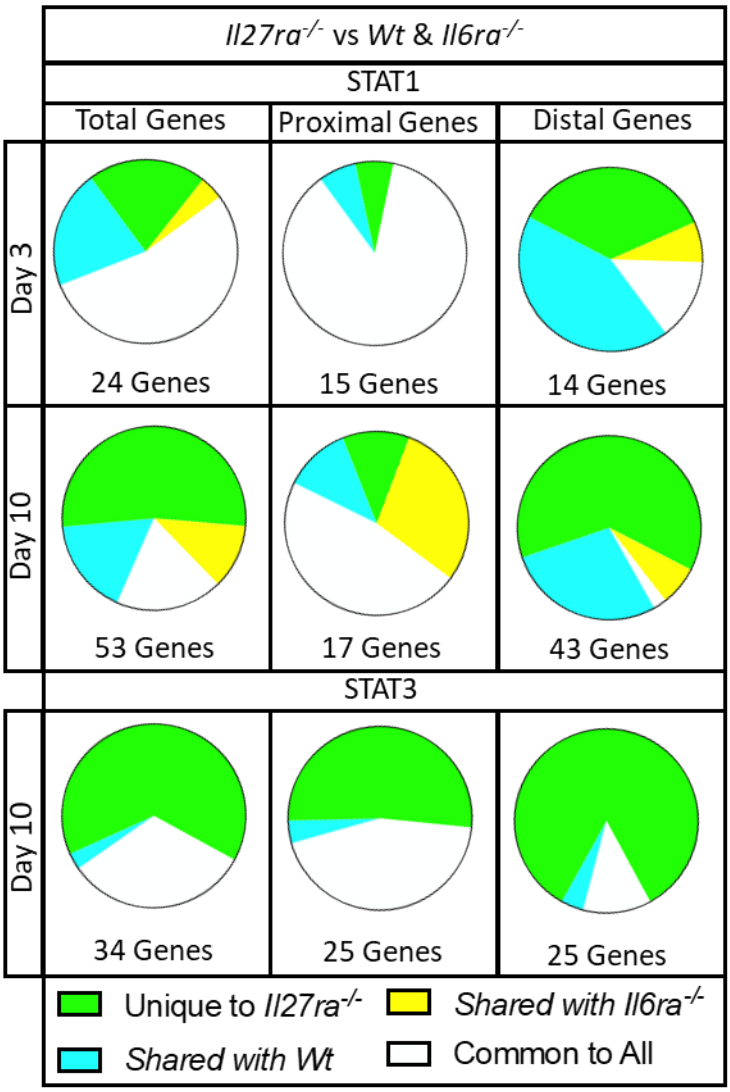


Figure 5.11: The Proportion of Genes Bound to STAT1 or STAT3 in the *Il27ra*$^{-/-}$ Condition Versus Other Genotypes. The genes bound to STAT1 or STAT3 in the Il27ra$^{-/-}$ genotype at Days 3 and 10 were compared with the Wt and Il6ra$^{-/-}$ to determine which are uniquely bound in Il27ra$^{-/-}$ genotype, which are bound in both the Il27ra$^{-/-}$ genotype and either Wt OR Il6ra$^{-/-}$ and which genes are bound in all genotypes. The proportion of genes in these four categories in relation to the Il27ra$^{-/-}$ genotype are shown in pie charts. This analysis was carried out in all genes (left), proximally associated genes (≤200 bp from a transcription start site; TSS) (middle) and distally associated genes (<200 bp from a TSS and comprising introns, exons and stop codons) (right).

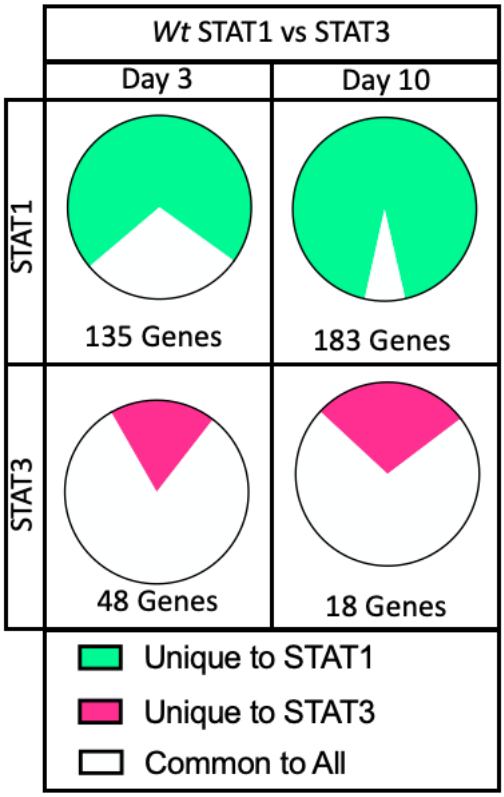


Figure 5.12 The Proportion of Genes Bound to STAT1 vs STAT3 at Day 3 and Day 10 in the Wt Condition. The genes bound to STAT1 were compared to those bound to STAT3 to determine which bind to one transcription factor and which bind to both. The proportion of genes in each category is shown in the pie charts.

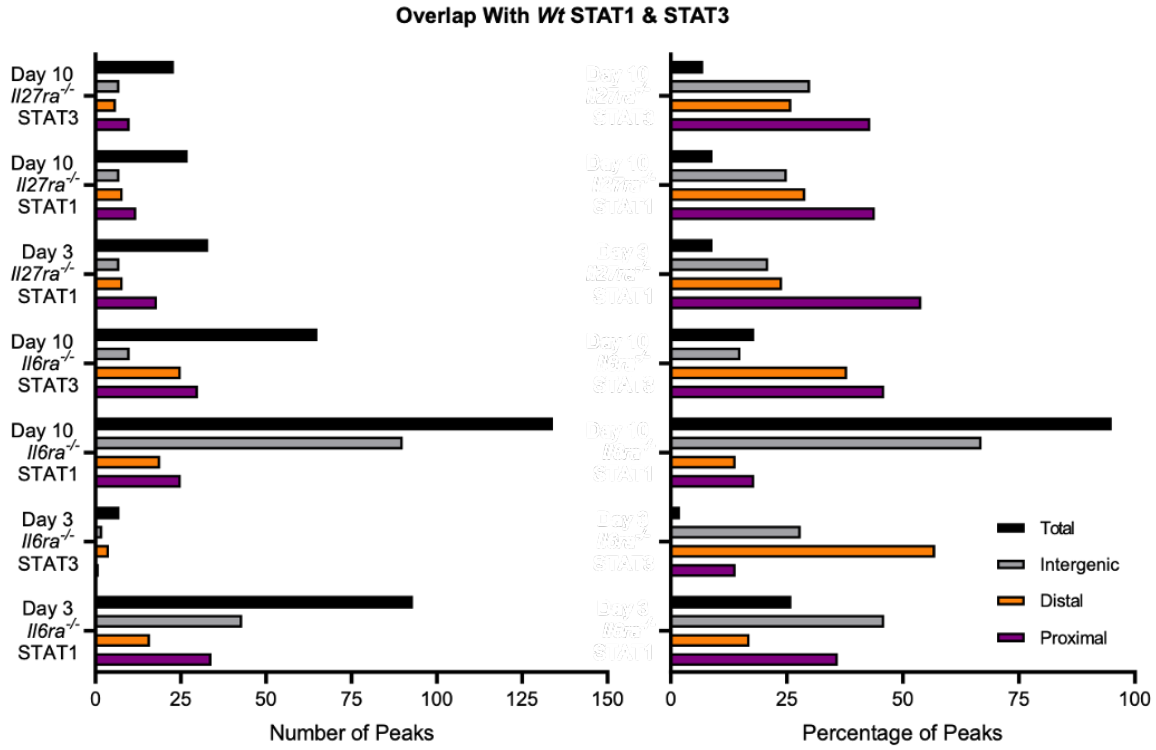


Figure 5.13: The Number and Percentage of Peaks Overlapping with Peaks Representing STAT1 AND STAT3 binding in the *Wt* Genotype. The graph on the left shows the number of peaks in each condition which align with the *Wt* STAT1-STAT3 peaks. The graph on the left indicates what percentage of of these overlapping peaks are associated with each genetic region. In this graph the 'Total' percentage (black) represents what percentage of that condition's peaks overlap with the *Wt* STAT1-STAT3 peaks. Proximal peaks are defined as ≤200 bp distant of a TSS. Distal peaks are defined as >200bp distant of a TSS. They are comprised of introns, exons and stop codons. Intergenic peaks are not associated with any gene.

5.3.6 STAT1 and STAT3 display a regulatory interplay during synovitis

STAT transcription factors share complex mechanisms of regulation, with studies showing that STAT1 often inhibits or shapes the transcriptional output of STAT3, and *vice versa* (Hu and Ivashkiv 2009; Harwardt et al. 2016; Butturini et al. 2020). This process is often termed cross-regulation. To establish whether cross-regulation contributes to the control of gene regulation in synovitis, analysis was focussed on those genes and loci which were common in the *Wt* condition. The white segments of **Figure 5.12** illustrate the genes in question, there are 39 shared genes at Day 3 and 13 at Day 10. These genes represent the genetic loci where both STAT1 and STAT3 exert behaviour and are thus dubbed STAT1-STAT3 genes. To assess the impact of IL-6 or IL-27 signalling inhibition on these loci they were compared with the total gene lists of the cytokine receptor knockout genotypes. The relevance of this is discussed later in this chapter.

Furthermore, to further assess these changes the sequencing peaks themselves were analysed using bed files and bedtools (https://bedtools.readthedocs.io/en/latest/). First, a list of peaks common to the STAT1 and STAT3 conditions in the *Wt* genotype was generated. This peak list was then compared to the other conditions at the same timepoint. The degree of overlap between these peaks is shown in **Figure 5.13.** In all genotypes and timepoints there is greater overlap between STAT1 peaks and the *Wt* STAT1-STAT3 subset. This is suggestive of diminished STAT3 behaviour regardless of IL-6 or IL-27 receptor deficiency at loci shared by STAT1 and STAT3 with normal receptor function. Furthermore, the *Il6ra*$^{-/-}$ genotype shows greater similarity with the *Wt* STAT1-STAT3 subset than their *Il27ra*$^{-/-}$ counterparts.

5.4 Discussion

5.4.1 Interpretation of IPA Analysis

5.4.1.1 Canonical pathways identified in relation to STAT1 and STAT3

A general inspection of the heatmaps in **Figure 5.6** shows a mix of pathways which are relevant to stromal cells or leukocyte immune cells. For instance, the pathways, *Role of Osteoclasts in RA Signalling Pathway, Role of Osteoblasts in RA Signalling Pathway* and *Role of Osteoblasts, Osteoclasts and Chondrocytes in RA*, all of which relate to stromal cell behaviour feature across the heatmaps. Conversely, pathways relating to leucocyte activity such as *Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes* and *Leukocyte Extraversion Signalling* do the same. Our experiments involved the sequencing of whole synovial tissue and thus, it is not surprising to find this mix of cell-type related pathways in the analysis. In the future, it could be interesting to employ magnetic-activated cell sorting (or other cell sorting system) as part of the ChIP-seq preparation protocol to assess the activity of STAT1 and STAT3 in a particular cell subtype. However, this would presumably require an increase in the number of animals required for each condition to collect sufficient cells.

Across the heatmaps a number of canonical pathways stand out as particularly interesting. At Day 3, the genes to which STAT1 binds, (**Figure 5.6.A**) which IPA identified as enriching the *Role of Osteoclasts in RA Signalling Pathway* vary between the *Wt* and *Il6ra*$^{-/-}$ conditions. In the *Wt* condition, the genes are *Col19a1, Col4a4, Map3k5* and *Tnfrsf11A.* In the *Il6ra*$^{-/-}$ condition the genes are *Ctnna3, Map3k5, Pik3cd* and *Tec*. Half of the *Wt* genes relate to connective tissue (*Col19a1* and *Col4a4*) while half of the *Il6ra*$^{-/-}$ genes relate to the immune response (*Pik3cd* and *Tec*). This is suggestive that different signalling pathways in the two different genotypes are responsible for STAT1's involvement with Osteoclasts in RA. The pathway receptor factor κβ *(RANK) Signalling in Osteoclasts* is also enriched in these two conditions by a subset of the aforementioned genes. In the *Wt* condition, the genes are *Map3k5* and *Tnfrsf11A.* In the *Il6ra*$^{-/-}$ condition the genes are *Map3k5* and *Pik3cd*. Given IL-6's complex relationship with osteoclasts both during homeostasis and the inflammatory state of RA (Yoshitake et al. 2008; Takeuchi et al. 2021) and that STAT1 has been implicated in the suppression of bone formation, for which osteoclasts are responsible (Kim et al. 2003; Seeliger et al. 2015) the enrichment of this pathway in these conditions is interesting.

At day 10 in the *Il6ra*$^{-/-}$ STAT1 condition (**Figure 5.6.C**) which indicates the enrichment of the aforementioned pathway are *Creb1, Ctnna3, Il1RAPl1*, *Pik3cd* and *Sos2*; different genes to those present in the day 3 counterpart. This is indicative of temporal changes over the disease course which nevertheless impact the same biological pathways. It is also interesting that at day 10 this pathway is not enriched in the *Wt* STAT1 condition as it was at day 3. However, the similar pathway *Role of Osteoblasts, Osteoclasts and Chondrocytes in RA* and the related pathway *Role of Osteoblasts in RA Signalling Pathway* are enriched in the *Wt* STAT3 condition at day 10. The gene responsible for this enrichment of both pathways is *Smo*.

Apoptosis, a similar self-regulatory process; autophagy and the balanced relationship of the two are crucial in tissue homeostasis. While autophagy involves the regulation and recycling of intercellular components, and typically enhances cell longevity apoptosis is programmed cell death. Disruption of these processes is associated with numerous diseases including RA by promoting the survival of damage-causing immune cells within the synovium. Pathways relating to these processes are enriched in both STAT1 and STAT3 conditions at Day 3 (**Figure 5.6.A&B**). In the STAT1 conditions, different genes are responsible for this enrichment. In descending order of degree of enrichment, they are *Il6ra*$^{-/-}$: *Birc6, Gnb1l, Pil3cd* and *Sesn1*. *Il27ra*$^{-/-}$: *Pten. Wt: Vps33a* and *Vps45*. This difference indicates that although autophagy is an important disease process in all genotypes it is being affected in different ways according to cytokine signalling deficiency.

The *Glutamate Receptor Signalling* Pathway is enriched at Day 3 in *Wt* and *Il6ra*$^{-/-}$ STAT1 and STAT3 conditions. Glutamate receptors are involved in bone remodelling and arthritic pain in RA (Bonnet et al. 2015). Interestingly, in the *Wt,* the same two genes are responsible for this enrichment in both STAT1 and STAT3 conditions: *Grid1* and *Grm1*. However, in the *Il6ra*$^{-/-}$ genotype the genes vary between STAT1 and STAT3 conditions. In the STAT1 they are *Gnb1l* and *Grin2a*. In the STAT3 they are *Grik2* and *Grip1*. IPA indicates that the pathway is enriched to a greater degree in the *Wt* conditions. Research has shown that activators of the glutamate receptors of synovial cells lead to an increase in the expression of IL-6 and thus increasing joint damage (Flood et al. 2007).

5.4.1.2 Diseases and functions identified in relation to STAT1 and STAT3

Upon first analysis of the diseases and functions enriched by the genes identified in ChIP-seq experiments, the top 50+ related to various cancers. However, cancer and the inflammation which characterises RA do share a number of biological processes. For instance, cell proliferation and survival. For this reason, these cancer diseases and functions shown in **Figure 5.12** were not dismissed out of hand. At day 3 a highly ranked disease identified by IPA was *Skin squamous cell carcinoma.* The various conditions share some genes which enrich the aforementioned disease, but there are differences which relate to cancer AND inflammation. For instance, Dock2, which has been associated with RA (Arandjelovic et al. 2019), appears in the *Wt* STAT1 condition. In the *Il6ra*$^{-/-}$ STAT3 condition the genes *Egfr* and *Erbb4* appear. Epidermal growth factor receptor (EGFR) is a therapeutic target in the treatment of RA (Yuan et al. 2013) and *Erbb* gene family members are involved in synovial cell development and signalling in RA (Satoh et al. 2001).

At day 10 the same disease, *Skin carcinoma,* is identified as enriched by IPA in both STAT1 and STAT3 conditions. Of the several genes which enrich the disease in the STAT3 condition, two are of particular interest. The first, *Ddr2*, is involved with the proliferation of synovial fibroblasts in RA (Su et al. 2009). The second gene of particular interest is *Oasl* because of its involvement in the IL-27 mediated signalling pathway (Kwock et al. 2020). Expression of OASL has also been used to predict the efficacy of tocilizumab, an IL-6 signalling inhibitor, in treating RA patients (Sanayama et al. 2014).

5.4.1.3 Upstream regulators identified in relation to STAT1 and STAT3

Analysis at day 3 in the STAT1 conditions showed Engulfment and Cell Motility (ELMO)1 to be greatly enriched in all but the *Il6ra*$^{-/-}$ genotype. In the remaining conditions, the regulator is enriched by the genes *Dock2* and *Elmo1*. Mice deficient in *Elmo1* have been shown to exhibit a lower inflammatory state in arthritis models (not AIA) (Arandjelovic et al. 2019). Interestingly the mice exhibiting the lowest degree of inflammation in the AIA model were the *Il6ra*$^{-/-}$ animals.

5.4.2 Interpretation of Shared Binding Analysis and the questions it raises

The pie charts in **Figures 5.9**, **10** and **11** display the commonality or uniqueness of the genes identified in each ChIP-seq experiment. White segments show genes regulated by STATs but not affected by IL-6 or IL-27 receptor deficiency, likely controlled by other STAT-activating cytokines. Blue, red, and green segments represent genotype-specific genes. In the Wt genotype, both IL-6 and IL-27 are necessary for STAT1 or STAT3 binding to associated genes, while in knockout genotypes, the drivers of STAT1 and STAT3 behaviour are less clear. We question whether STAT activity is influenced by the absence of the missing cytokine signalling or a proportionally greater impact from the remaining cytokine signalling.

Magenta segments (**Figure 5.9** and **Figure 5.10**) represent genes shared between *Wt* and *Il6ra*$^{-/-}$ conditions. Cyan segments (**Figure 5.9** and **Figure 5.11**) represent genes shared between *Wt* and *Il27ra*$^{-/-}$ conditions. Both sets require either cytokine, but STAT behaviour is unaffected by the presence or absence of the other cytokine.

Yellow segments (**Figure 5.10** and **Figure 5.11**) are genes common to BOTH cytokine receptor knockout genotypes but not *Wt*. STAT1 binding seems to require either cytokine but not both, suggesting that inhibition of IL-6 and/or IL-27 activity in the *Wt* mice prevents STAT1 binding to these genes.

5.4.3 Genes which bind both STAT1 and STAT3

As stated previously, the white segment in **Figure 5.12** illustrates the genes which are the subject of this discussion. At Day 3, 21 of the *Wt* STAT1-STAT3 genes share no activity with the other conditions. Conversely, there are 2 genes (*Gm10717* and *Gm10801*) which are common to all genotypes. However, both are only predicted genes and their functions are unassigned. This left 16 of the *Wt* STAT1-STAT3 genes. Of these 11 show STAT1 binding activity in both knockout genotypes but the STAT3 binding activity exhibited in the *Wt* condition is lost in the *Il6ra*$^{-/-}$. Unfortunately, without the *Il27ra*$^{-/-}$ STAT3 counterpart, only a limited conclusion; loss of IL-6 receptor signalling stops STAT3 binding activity in these genes, can be drawn. Of these 11 genes, 10 are predicted genes, the remaining gene is *Mir101c*. Of the *Wt* STAT1-STAT3 genes, there are 3 (*Dock2*, *Mdga2* and *Gm26870*) which only show shared binding activity with the *Il27ra*$^{-/-}$ STAT1 condition, this suggests a total loss of STAT1 and STAT3 activity relating to these genes with the absence of IL-6 receptor

signalling. There is another gene, *Cldn34d,* which in addition to exhibiting binding in the *Wt* STAT1-STAT3 conditions and the *Il27ra*$^{-/-}$ STAT1 condition shows binding in the *Il6ra*$^{-/-}$ STAT3 condition, this suggests a loss of STAT1 activity relating to this gene in the absence of IL-6 signalling. Finally, *Anks1b* is associated with the *Wt* STAT1-STAT3 and *Il6ra*$^{-/-}$ STAT1 conditions. This suggests a loss of STAT3 activity and STAT1 activity in the absence of IL-6R and IL-27R signalling, respectively.

The story at Day 10 is less dynamic. Of the 13 *Wt* STAT1-STAT3 associated genes 2 (*Tcte2* and *Smo*) are unique to the *Wt* genotype. There are 10 shared by all conditions, however, they are all predicted genes. Finally, there is one other predicted gene (*Gm26870*) which loses both STAT1 and STAT3 binding activity as a result of IL-6 signalling absence, though both transcription factors remain active in the *Il27ra*$^{-/-}$ conditions.

5.4.4 Evidence of STAT1 and STAT3 cross-regulation

Again, using the Venn diagram tool introduced in **Section 5.3.7,** the two knock-out genotypes were compared to see if the change in cytokine signalling elicited a switch from STAT1 to STAT3 (or STAT3 to STAT1) on particular genes. Unfortunately, with the missing condition at Day 3, which has been mentioned, it was not possible to make these observations at that time point. However, Day 10 was assessed.

Only two genes underwent a swap in their transcription factor binding. In the *Il6ra*$^{-/-}$ genotype *Vps45* and *Psd3* bind STAT3, while in the *Il27ra*$^{-/-}$ genotype they bind STAT1. Neither of these genes bind either STAT1 or STAT3 in the *Wt* genotype.

There are several genes which do not switch in the manner of the two previously mentioned. In the *Il6ra*$^{-/-}$ genotype STAT1 and STAT3 bind the genes *Cldn34d, Grid1, Mdga2* and the predicted genes *Gm10720* and *Gm10801*. However, in the *Il27ra*$^{-/-}$ genotype their binding with STAT3 is lost while STAT1 binding remains.

5.4.5 Limitations of this Data

As stated in the introduction to this chapter, STAT1 and STAT3 ChIP-seq have never before been performed utilising inflamed synovial tissue. The experiments discussed above present novel and exciting insights into the inflammatory landscape of synovitis and the role played by IL-6, IL-27, STAT1 and STAT3 in its establishment and maintenance.

However, there are two key problems with this ChIP-seq data, both of which have been briefly mentioned earlier in this chapter but will be expanded upon here. They are the lack of Day 3 *Il27ra*$^{-/-}$ STAT3 data and the asynchronous sequencing of the *Il6ra*$^{-/-}$ genotype conditions resulting in their data requiring rationalisation.

Two attempts were made to sequence the Day 3 *Il27ra*$^{-/-}$ STAT3 however, on both occasions there was insufficient material following the preparation of the libraries to do so. It is difficult to determine if this was due to an error made during the preparation of the sample (though other samples were prepared at the same time on both occasions and were sequenced successfully) or if STAT3 activity at Day 3 in the *Il27ra*$^{-/-}$ genotype is so low as to make performing ChIP-seq exceptionally difficult or impossible. It is unfortunate that due to time constraints and the fluctuant availability of suitable mice from the *Il27ra*$^{-/-}$ colony, that a third attempt to perform STAT3 ChIP-seq could not be made. The lack of this condition compromises comparison analysis and our ability to fully elucidate the role of STAT1 and STAT3 in the face of cytokine receptor deficiently and how they influence the heterogeneity of RA.

During this project the *Il6ra*$^{-/-}$ colony showed low breeding rates. Due to the low inflammatory phenotype of these mice, any AIA experiment required more animals than the other two genotypes in order to pool enough synovial cells to perform sequencing experiments. This low breeding rate and high animal requirement meant that it was not possible to carry out the AIA and subsequent sequencing of all genotypes at the same time. Thus, when the *Il6ra*$^{-/-}$ conditions were sequenced, there were fewer samples in the library pool causing a degree of over-sequencing and the calling of low-value peaks.

The discrepancy in the number of peaks between the *Wt* and *Il27ra*$^{-/-}$ conditions as well as STAT1 and STAT3 indicates that there is variance not only in where each transcription factor binds but also in the degree of activity. Unfortunately, this information is lost in the *Il6ra*$^{-/-}$ genotype by the over-sequencing and rationalisation utilising the mean number of peaks from all other conditions.

Chapter 6:

General Discussion

6.1 The work Presented in this Thesis.

The heterogeneity of synovitis in RA is well documented in recent literature (Pitzalis et al. 2013; Rivellese et al. 2022). However, the biological mechanisms driving differences in synovial pathology remain unclear. In this regard, it is currently unknown whether fibroblast-rich, myeloid-rich, and lymphoid-rich synovitis originate from three alternate inflammatory mechanisms or simply reflect different stages within the disease trajectory. Despite these uncertainties, evidence from the study of human synovial biopsies shows that the histological features and transcriptional profile of synovitis differ from patient to patient (Dennis et al. 2014; Lewis et al. 2019). These features often correlate with the severity of the disease and response to commonly prescribed biological medicines (Katchamart et al. 2010; Choy et al. 2013); Rivellese et al., 2022; Humby et al., 2021. Thus, an increased understanding of the variables and molecular mechanisms steering arthritis progression offers exciting opportunities to improve the diagnosis, stratification, and treatment of patients with RA.

The work presented in this thesis has sought to gain a better understanding of the epigenetic landscape of synovitis using available RNA-seq and the generation of ATAC-seq and ChIP-seq datasets from the inflamed synovium of mice with AIA. Attention was specifically given to mice lacking *Il6ra* or *Il27ra*. Prior studies from our laboratory show that IL-6 and IL-27 play pivotal roles in the development of synovitis, with the analysis of mice with AIA showing that the histological features of synovitis in *Il6ra* or *Il27ra* resemble those seen in human forms of fibroblast-rich and lymphoid-rich synovitis. In contrast, *Wt* mice with AIA are more closely related to myeloid-rich synovitis. Access to these mouse strains, therefore, provides existing opportunities to investigate the underpinning mechanisms driving disease heterogeneity in RA (Nowell et al. 2009; Jones et al. 2013; Dennis et al. 2014; Jones et al. 2015; Lewis et al. 2019; Twohig et al. 2019).

In this thesis, I have investigated the regulation of gene expression by evaluating changes in chromatin accessibility and the activation of transcription factors in synovial tissues following AIA onset (see **Figure 6.1**). The results provide new insights into the role played by the cytokines IL-6 and IL-27 in shaping chromatin accessibility and STAT1 and STAT3 transcription factor activation as orchestrators of disease heterogeneity.

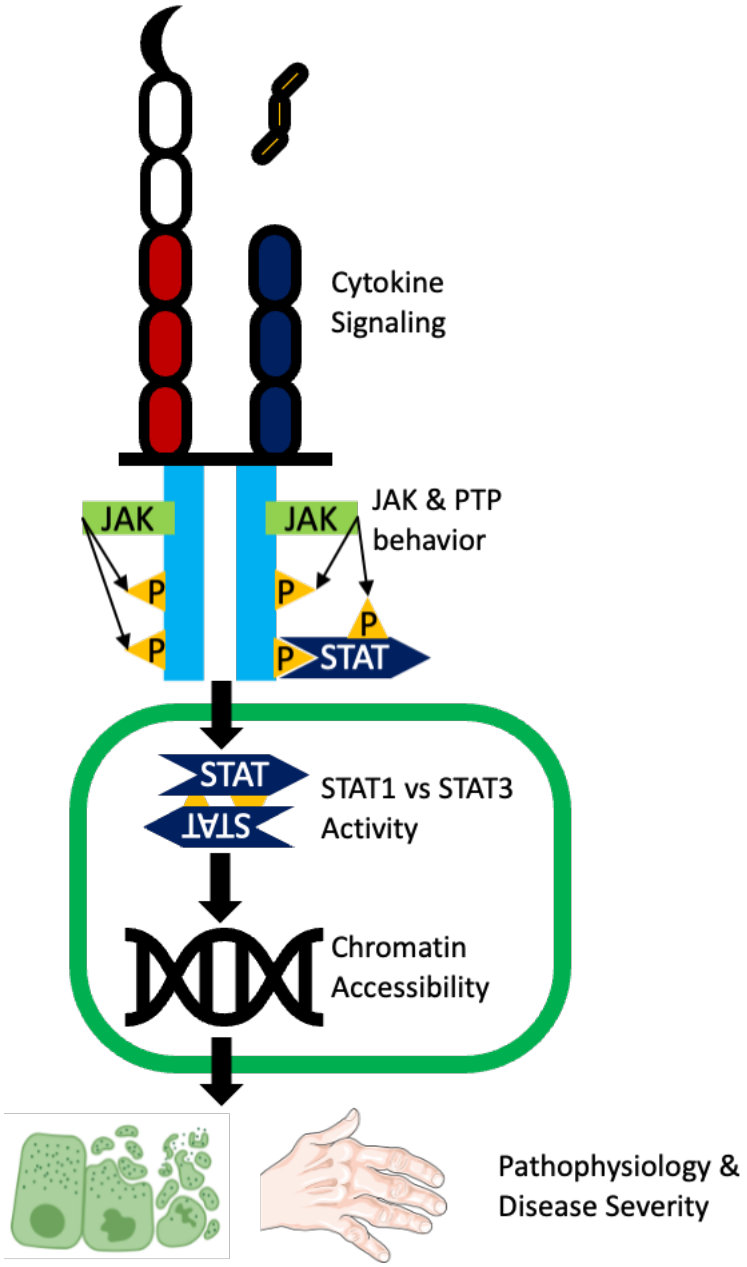


Figure 6.1: The multiple levels of control which fine tune cytokine signals to affect disease. When inflammatory cytokines bind with a receptor there are multiple steps in a pathway which ultimately leads to gene expression and disease presentation. These steps in the cascade provide an opportunity for modulation of the cytokine signaling pathway. This theses has explored two of these steps (shown in the green box); STAT1 and STAT3 activity and chromatin accessibility. (Adapted from Jones et al. 2015; Villarino et al. 2017a; Morris et al. 2018; Klemm et al. 2019; Twohig et al. 2019; Wiede et al. 2019; Jenkins et al. 2021).

6.2 Chromatin Accessibility Remains Broadly Consistent

The commonality of open chromatin regions and the genes within them was demonstrated in **Chapter 4.** However, previous studies of synovitis in *Wt, Il6ra*$^{-/-}$and *Il27ra*$^{-/-}$ mice with AIA have shown clear differences in histological features and transcriptional profiles (Hill, 2019). The synovitis displayed by *Wt, Il6ra*$^{-/-}$and *Il27ra*$^{-/-}$ mice with AIA share many common features of the disease. However, they also showed marked differences. For example, synovial pathology in *Il6ra*$^{-/-}$ mice evolves in the absence of inflammatory infiltrating leukocytes. Whereas the synovitis associated with *Il27ra*$^{-/-}$ mice develops to include evidence of synovial ectopic lymphoid aggregates comprising T-cells and B-cells. Yet, the majority of genes are assessable across all three genotypes at a given timepoint and only a relative few are uniquely accessible in only one genotype. I suggest that much of this similarity is simply due to the samples sequenced belonging to the same tissue type during an inflammatory state. Furthermore, the relatively subtle changes in chromatin accessibility resulting from cytokine receptor inhibition have a profound influence on the disease state. However, as chromatin accessibility is not the only mechanism by which gene expression is regulated two different approaches were exploited. Investigating the shared open regions showed the role of alternative mechanisms of gene regulation on disease outcomes. Conversely, investigating the variable regions showed the direct role of gene accessibility.

6.3 How Subtle Changes in Chromatin Accessibility Shapes Disease

Biological medicines and small molecule inhibitors have revolutionised the treatment of patients with rheumatoid arthritis (Choy et al. 2013). It is now rare for patients with rheumatoid arthritis to display signs of joint deformities including boutonniere and swan new deformities on the hand and hallux valgus in the feet (Ishikawa 2017). Despite these improvements in patient outcomes, many patients still fail to adequately respond to this particular class of biological inhibitors (Buch 2018). Indeed, evidence from clinical trials shows that the ratio of ACR20, ACR50, and ACR70 scores of clinical improvement remain comparable across all studies testing the clinical efficacy of cytokine blockers and inhibitors of lymphocyte activation(Castro et al. 2022). In the UK, guidance on the best course of therapy is often guided by NICE recommendations, clinician awareness of the action of a specific drug and considerations including intolerance to a DMARD intervention (NICE

2021). A trope of clinical rheumatology is a 'trial and error' approach to treatment. However, advances in the use of small needle ultrasound-guided biopsy sampling in clinical trials of RA patients have illustrated the diversity of disease displayed by patients and thus these unsuccessful trials could be avoided if inflamed synovial tissue were to be assessed to establish the best course of treatment (Pitzalis et al. 2020). While headway is being made in this regard to identify synovial biomarkers, much of this work is based on RNA-seq and immunohistochemistry (Rivellese et al. 2022). My data suggests the possibility of utilising chromatin accessibility in these endeavours.

The identification of open regions which are accessible only during the disease state was the first baby step in linking specific accessible chromatin regions with inflammatory and other disease processes. Molecular Pathway analysis and enriched cell type analysis shown in **Chapter 5** demonstrated that relatively minor changes to the chromatin accessibility landscape, resulting from differences in cytokine signalling, had a significant impact as shown by the differences in the enrichment of pathways and cell types. For instance, this was seen in the pathways '*Role of JAK family kinases in IL-6-type Cytokine Signalling'*, Role *of Chondrocytes in RA Signalling Pathway'* and '*Role of Osteoclasts in RA Signalling Pathway.* Additionally, the enrichment of a number of RA and other inflammatory pathways affirms the role played by disease-accessible genes downstream. Furthermore, the cell type analysis showed T-cells and B-cells to be enriched in the *Il27ra*$^{-/-}$ genotype compared with the other two and the *Il6ra*$^{-/-}$ genotype showed relatively mild enrichment of immune type cells reflecting its low inflammatory nature.

When mapping ATAC-seq to RNA-seq it was found that differentially accessible genes (between genotypes) can exhibit the same pattern of expression as other genotypes. While this not only hints at the dynamic transcriptome and accessible chromatin landscape, the discrepancy between RNA-seq and ATAC-seq also shows that RNA-seq alone may be insufficient to fully explore epigenetic aspects of the synovium in RA. The use of both technologies has permitted an additional level of understanding of the disease process.

Some possible explanations for these findings are put forward here. ATAC-seq is a 'snapshot' of chromatin accessibility and the rate at which a promoter region or TSS opens and closes may vary. To understand these subtleties would require a more detailed temporal analysis of chromatin accessibility. Though there are differences in accessibility between genotypes, these are differences in chromatin's permissiveness of mutant Tn5 and

there may not be a difference between genotypes in terms of their chromatin accessibility for transcription factors, co-regulators, and co-receptors.

6.4 STAT1 and STAT3 as drivers of disease

The problems of the of *Il6ra*$^{-/-}$ sequencing over-representation
and the unsuccessful DNA library preparation of the Day 3 *Il27ra*$^{-/-}$ STAT3 sample have been discussed previously. While these unfortunate issues did somewhat impede our analysis, they did not make it impossible.

As shown in **Chapter 5**, pathways such as *Role of Osteoclasts in RA Signalling Pathway, Role of Osteoblasts in RA Signalling Pathway* and *Role of Osteoblasts, Osteoclasts and Chondrocytes in RA* demonstrated STAT1 and STAT3's role in disease outcome via their binding to and activation of genes which propagate inflammation within the synovium. These findings are similar to those of others investigating the role of STAT1 and STAT3 in microglia and renal cells (Przanowski et al. 2014; O'Brown et al. 2015). This highlights the cytokine-STAT-inflammation axis as a conserved disease mechanism warranting greater investigation in other IMIDs.

The central dogma provides geneticists with the reassuring mantra DNA makes RNA; RNA makes protein. Another core tenet of genetics is that for a particular gene to be expressed its chromatin must be open and accessible to a binding transcription factor which then transcribes the gene (Crick 1970; Thurman et al. 2012). When analysing ATAC-seq, STAT1 and STAT3 ChIP-seq and Dr Hill's RNA-seq data together I have shown that deficiency in IL-6 and IL-27 signalling does not impact one stage of gene expression. Rather, the differences in chromatin accessibility AND STAT1-STAT3 behaviour observed between genotypes demonstrate that changes in cytokine signalling impact numerous stages leading to expression and ultimately phenotypic variation.

6.5 Future Perspectives

6.5.1 Animal Models

The number of cells required for ATAC-seq and ChIP-seq experiments exceeded that which could be obtained from the knee joints of a single mouse. This necessitated the use of

multiple animals and the pooling of recovered synovial tissue. However, this brought with it several complexities. One of these was the number of mice required to pool for sufficient cells. The *Il27ra*$^{-/-}$ disease phenotype displayed the highest degree of inflammation, which in addition to meaning recovery of synovial tissue was much easier than in other genotypes, meant that fewer animals were required to reach the required cell number. *Wt* mice displayed a middling degree of inflammation, but this did not cause an issue. However, the *Il6ra*$^{-/-}$ genotype displayed a low degree of inflammation, requiring more animals which was complicated by breeding difficulties. The poor breading rates in the *Il6ra*$^{-/-}$ colony could suggest a link between IL-6 and fertility (Prins et al. 2012). This relationship warrants further exploration, though it is well beyond the scope of this work.

In an ideal research world, all AIA experiments and subsequent sequencing would have been conducted at the same time. There would have been one cell pool of each genotype, sufficient for all ATAC-seq and ChIP-seq experiments (as explained in **Chapter 2,** the recovery and initial processing of tissue is identical in both sequencing types). However, the number of mice required for this large experiment would have exceeded the lab's limited manpower. Additionally, the aforementioned colony problems with *Il6ra*$^{-/-}$ genotype meant that they were not available when the *Wt* and *Il27ra*$^{-/-}$ experiments were conducted.

6.5.1.1 Asynchronous Sequencing

As explained in the relevant chapters, the ATAC-seq and ChIP-seq of the *Il6ra*$^{-/-}$ genotype were not performed at the same time as the other two genotypes. This meant that the *Il6ra*$^{-/-}$ library pool had half the number of samples as the *Wt-Il27ra*$^{-/-}$ pool. This was the cause of the over-reading (as compared with other genotypes) as described in **Chapters 3** and **5** which necessitated the rationalisation of the *Il6ra*$^{-/-}$ data which is described in those chapters. Though not ideal, these adjustments have allowed us to compare and interpret the data from all three genotypes.

6.5.1.2 Additional Murine Genotypes

When analysing the sequencing data gathered during the work of this thesis each of the cytokine receptor knockout genotypes was compared to the *Wt.* In doing so we could observe how inhibiting IL-6 or IL-27 signalling deviated chromatin accessibility or STAT1/3

from the presumed 'normal' progression of synovitis modelled in the *Wt* mice. However, IL-6 and IL-27 are not the only molecules to play a role in the RA disease state, others include TNF, GM-CSF and IL-17 (Gaffen 2009; Lee et al. 2020; Jang et al. 2021).

Even within the IL-6 signalling pathway, there are other avenues to be explored, particularly within the JAK-STAT pathway. A key feature of the JAK-STAT pathway is that it can be controlled at multiple levels (Villarino et al. 2017). One such example is that of protein tyrosine phosphatases (PTPs), which serve as negative regulatory elements to control transcription factors (Seif et al. 2017). Two PTPs in particular have demonstrated involvement with autoimmune conditions and have also been identified in GWAS studies; these are PTPN2 and PTPN22 (Barrett et al. 2008; McAllister et al. 2011; Sharp et al. 2015). Furthermore, various experiments involving the PTPN2 and PTPN22 enzymes have demonstrated their involvement with IMIDs. For instance; $CD4^+$ T-cells obtained from *Lck*-Cre:*Ptpn2*$^{fl/fl}$ (T-cell restricted deletion) and *Ptpn22*$^{-/-}$ mice revealed that PTPN2 and PTPN22 act to inhibit the tyrosine-phosphorylation of STAT1 (pY-STAT1) in IL-6 treated effector memory $CD4^+$ T-cells (Twohig et al. 2019). This regulatory mechanism of STAT1 was observed to alter the gene profile under STAT3 control in memory $CD4^+$ T-cells when compared to IL-6 responses in naïve $CD4^+$ T-cells. These included genes associated with the development of ectopic lymphoid-rich synovitis (Twohig et al. 2019; Wiede et al. 2019).

Examination of PTPN2 in synovial biopsies of patients with RA demonstrated its highest expression in lymphoid-rich synovitis. This suggests that PTPN2 plays a role in regulating Jak-STAT signaling specifically in lymphocyte-driven RA pathology, impacting the composition, organization, and activities of cells involved in the development of ELSs. PTPN22 on the other hand exhibits a more widespread expression pattern across fibroblast-rich, myeloid-rich, and lymphoid-rich synovitis. This suggests the involvement of PTPN22 in various hematopoietic populations, such as myeloid cells, and stromal tissues. Its functions include the inhibition of STAT1 signals in T-cells, macrophages, and fibroblasts. Consequently, the characteristics of PTPN22 may influence Jak-STAT signals across disease pathologies (Twohig et al. 2019).

In murine models, the deficiency of PTPN2 has been shown to exacerbate disease conditions, affecting follicular T-helper cells, regulatory T cells, and B cells. In *Ptpn2*$^{+/-}$ mice, synovitis manifested with an increased expression of synovial Th17 cells and ectopic lymphoid structures in the SKG model of autoimmune arthritis, a genetic model in which

mice spontaneously develop chronic arthritis. This pattern of inflammation is similar to the development of synovitis observed in *Il27ra*$^{-/-}$ mice used in our own AIA experiments (Sakaguchi et al. 2003; Jones et al. 2015; Svensson et al. 2019).

At one time early in the work of this thesis, there existed a plan to perform ATAC-seq and STAT1 and STAT3 ChIP-seq on *Ptpn22*$^{-/-}$ and *Ptpn2*$^{+/-}$ mice with AIA. However, logistical, and other problems brought about by the COVID-19 pandemic made these experiments impossible at the time, which was disappointing. However, given the importance of PTPs in mediating inflammation through the JAK-STAT signalling pathway, these models and experiments are certainly worth pursuing.

A key achievement of the work of this thesis has been the refinement of the ATAC-seq and ChIP-seq methods, in conjunction with the AIA model followed by raw data normalisation and subsequent analysis. Thus, while this section has focused on *Lck*-Cre:*Ptpn2*$^{fl/fl}$, *Ptpn22*$^{-/-}$ and *Ptpn2*$^{+/-}$ mice, other genotypes relevant to synovitis could be utilised and compared with the data presented here.

6.5.2 Additional ChIP-seq experiments

In **Chapter 5** the lack of Day 3 *Il27ra*$^{-/-}$ STAT3 condition was discussed. Though it was attempted twice, unfortunately, due to time constraints and the availability of appropriate *Il27ra*$^{-/-}$ genotype mice it was not possible to reattempt sequencing in this condition. It would be worth revisiting this condition in the future. It would be particularly interesting to see if the failure of the sample preparation was due to experimental error, or if there is such a lack of STAT3 activity in this genotype and timepoint which makes its sequencing difficult. This later point would seem unlikely given the data generated at Day 3 from Wt and *Il6ra*$^{-/-}$ mice.

ChIP-seq experiments in this thesis were limited to STAT1 and STAT3 to explore their relationship and cross-regulation of genes. While the findings of this work are of great interest, there are other transcription factors which play a role in RA's pathogenesis, for instance, STAT4 and P300 are but two possible examples (Remmers et al. 2007; Twohig et al. 2019). STAT4 has been implicated in the development of Th1-type T-cell responses and the differentiation of Th17 cells which play a key role in the IMIDs, including RA (Remmers et al. 2007). P300 is involved in the regulation of synovial fibroblasts and their role in

inflammation and the pathogenesis of RA (Krošel et al. 2023) and it is thought to be related to the expression control of genes regulated by STAT1 (Twohig et al. 2019).

6.5.3 Single Cell ATAC-seq

The bulk ATAC-seq presented in this thesis has enabled the characterisation of the open chromatin regions within the synovium during inflammation. It provides a snapshot of synovitis' epigenetic landscape. We have established that IL-6 and IL-27 receptor deficiency does indeed change chromatin accessibility and we have explored the biological relevance of these differences to see how it affects RA's heterogeneity. However, we have not explored heterogeneity within the synovium. Single-cell ATAC-seq (scATAC-seq) is an adaptation of traditional ATAC-seq. By capturing the unique open chromatin profiles of individual cells scATAC-seq enables the identification of distinct subpopulations and rare cell types. The method provides a high-resolution view of the epigenetic landscape at a single-cell level (Buenrostro et al. 2015; Mezger et al. 2018). In the context of synovitis, the differences in the type and behaviour of inflammatory cells would be identified. The ATAC-seq data gathered in the course of the work of this thesis in combination with new scATAC-seq would provide a comprehensive view of synovitis and the impact of IL-6 and IL-27 receptor deficiency.

6.6 Translational Research

An accurate diagnosis of disease is of the utmost importance when physicians and surgeons come to treat immune-mediated inflammatory diseases. A patient presenting with joint pain could have a plethora of conditions, including gout, osteoarthritis, septic arthritis, psoriatic arthritis, or of course rheumatoid arthritis. With the exception of relieving pain, treatments for these and other joint conditions can vary considerably (Senthelal et al. 2023). Even after an initial diagnosis of RA is made the journey does not end. RA can be further categorised into two groups defined as seropositive or seronegative for rheumatoid factor and/or anti-citrullinated peptide antibodies (Sokolova et al. 2021). Each subtype exhibits a different disease process, treatability, and outcome (Malmström et al. 2016; Martin-Mola et al. 2016; Reed et al. 2020). A patient's response, or lack thereof, to treatment is another way by which RA can be categorised. This is particularly true in the case of treatment with biologic

drugs where a drug's mechanism of action can elucidate the biological process driving that patient's disease. For example, patients responding to tocilizumab or infliximab suggest that IL-6 or TNF respectively is a core component of their disease. Likewise, elements downstream, such as JAK-STAT in the case of IL-6 and NF-κB for TNFα, may also be critical to that patient's disease state (Choy et al 2013). RA can be grouped via histopathological assessment. As described in **Chapter 1, Section 1.5.1** and **Figure 1.3**, these groups are Follicular (lymphoid-rich), Diffuse (myeloid-rich) and Pauci immune (Fibroblast-rich) (Pitzalis et al. 2013). Here we postulate an alternative method of RA disease classification based on its "epigenetic fingerprint".

The work of Dr. Hill's PhD thesis and the RNA-seq data he generated demonstrated differences in the synovial transcriptome of *Wt*, *Il6ra*$^{-/-}$ and *Il27ra*$^{-/-}$. In this thesis, we have presented the differences in the accessible genome of the inflamed synovium in the same mouse genotypes. Though in its infancy, this work paves the way by which RA could be classified by the accessibility of key genomic regions, the activity of disease-related transcription factors and the expression of genes in said regions, similar to other work involving the classification of human cancers (Corces et al. 2018).

Naturally, this would require applying these sequencing technologies to synovial tissue of people with arthritis or by constructing clinical algorithms that predict synovial disease progression based on blood biomarkers that extend beyond current measures of CRP and ESR. The IPA analysis presented in this thesis contextualised the manner in which epigenetic differences; chromatin accessibility and STAT1 and STAT3 binding, influence disease. By stratifying patients by an 'epigenetic fingerprint', be it chromatin accessibility, a transcription profile, transcription factor activity or a combination of these, the mechanisms and pathways driving disease would be revealed as a matter of course. This would allow the most appropriate treatment to be selected and offered. Likewise, new, previously unexplored, potential treatment targets could be identified.

While the ATAC-seq data presented in this thesis stems from three distinct genotypes, humans are rarely so consistent. We, therefore, speculate that were similar ATAC-seq experiments to be performed on RA synovial tissue, the patients to whom the tissue belonged would fall on a spectrum of chromatin accessibility relating to their disease state. However, we recognise that current human trials of this nature are based on highly stratified patient groups (Pitzalis et al. 2020; Rivellese et al. 2022). It is at this juncture that

our own highly stratified cytokine receptor knockout data would enable patents to be placed in accordance with their accessible synovial genomes' similarity with the open regions of their murine counterparts.

Establishing a 'roadmap' of the chromatin accessibility of synovitis AND the healthy synovium has potential therapeutic benefits. Although admittedly a very distant prospect, reshaping open chromatin regions from a disease state to one of healthy homeostasis could be a treatment of RA. The ATAC-seq data presented in **Chapter 3**, showed genes, and by extension chromatin regions, uniquely accessible in the disease state versus a naïve Day 0 control. By blocking, closing or otherwise inhibiting these disease-specific regions, synovitis could be offset. Though lacking the level of precision and specificity which would be required, there are drugs in existence which can influence chromatin accessibility (Arrowsmith and Schapira 2019; Husmann and Gozani 2019; Zaware and Zhou 2019; Yao et al. 2020). In the field of cancer treatment, similar efforts are already underway in the form of DNA methyltransferase inhibitors which prevent DNA hypermethylation to reactivate silenced genes. These have shown some potential as anticancer agents (Zhang et al. 2022). Similarly, Histone deacetylase inhibitors have demonstrated the ability to reverse the silencing of tumour suppressor genes (Li et al. 2020). Taking these drug types into account in combination with a better understanding of the epigenetic landscape of RA may be of great benefit in treating patients with the disease.

Bibliography

7.1 Bibliography

Aaronson, D.S. and Horvath, C.M. 2002. A road map for those who don't know JAK-STAT. *Science* 296(5573), pp. 1653–1655. Available at: https://pubmed.ncbi.nlm.nih.gov/12040185/ [Accessed: 14 April 2021].

Abdallah, F. et al. 2021. Leukocyte Immunoglobulin-Like Receptors in Regulating the Immune Response in Infectious Diseases: A Window of Opportunity to Pathogen Persistence and a Sound Target in Therapeutics. *Frontiers in Immunology* 12, p. 717998. doi: 10.3389/FIMMU.2021.717998/BIBTEX.

Adey, A. et al. 2010. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biology* 11(12), pp. 1–17. Available at: https://genomebiology.biomedcentral.com/articles/10.1186/gb-2010-11-12-r119 [Accessed: 9 May 2022].

Ainola, M., Li, T.F., Mandelin, J., Hukkanen, M., Choi, S.J., Salo, J. and Konttinen, Y.T. 2009. Involvement of a disintegrin and a metalloproteinase 8 (ADAM8) in osteoclastogenesis and pathological bone destruction. *Annals of the rheumatic diseases* 68(3), pp. 427–434. Available at: https://pubmed.ncbi.nlm.nih.gov/18397961/ [Accessed: 28 November 2023].

Akahoshi, T. and Watabe, H. 2009. Infliximab in the treatment of rheumatoid arthritis. *Biologics : Targets & Therapy* 3(12), p. 183. Available at: /pmc/articles/PMC2726073/ [Accessed: 17 September 2023].

Allis, C.D. and Jenuwein, T. 2016. The molecular hallmarks of epigenetic control. *Nature Reviews Genetics 2016 17:8* 17(8), pp. 487–500. Available at: https://www.nature.com/articles/nrg.2016.59 [Accessed: 8 May 2022].

Alpízar-Rodríguez, D., Pluchino, N., Canny, G., Gabay, C. and Finckh, A. 2017. The role of female hormonal factors in the development of rheumatoid arthritis. *Rheumatology (Oxford, England)* 56(8), pp. 1254–1263. Available at: https://pubmed.ncbi.nlm.nih.gov/27686101/ [Accessed: 1 August 2022].

Alqarni, A.M. and Zeidler, M.P. 2020. How does methotrexate work? *Biochemical Society Transactions* 48(2), pp. 559–567. Available at: /biochemsoctrans/article/48/2/559/222539/How-does-methotrexate-work [Accessed: 18 August 2022].

Andreozzi, V. et al. 2022. In-Office Needle Arthroscopic Synovial Biopsy Is an Effective Diagnostic Tool in Patients With Inflammatory Arthritis. *Arthroscopy, Sports Medicine, and Rehabilitation* 4(6), pp. e2099–e2106. doi: 10.1016/J.ASMR.2022.10.003.

Araki, Y. and Mimura, T. 2017. Matrix Metalloproteinase Gene Activation Resulting from Disordred Epigenetic Mechanisms in Rheumatoid Arthritis. *International Journal of Molecular Sciences* 18(5). Available at: /pmc/articles/PMC5454818/ [Accessed: 17 December 2022].

Arandjelovic, S. et al. 2019. A non-canonical role for the engulfment gene ELMO1 in neutrophils that promotes inflammatory arthritis. *Nature immunology* 20(2), p. 141. Available at: /pmc/articles/PMC6402828/ [Accessed: 2 September 2023].

Arrowsmith, C.H. and Schapira, M. 2019. Targeting non-bromodomain chromatin readers. *Nature Structural & Molecular Biology 2019 26:10* 26(10), pp. 863–869. Available at: https://www.nature.com/articles/s41594-019-0290-2 [Accessed: 2 November 2023].

Asquith, D.L., Miller, A.M., McInnes, I.B. and Liew, F.Y. 2009. Animal models of rheumatoid arthritis. *European Journal of Immunology* 39(8), pp. 2040–2044. Available at:

https://onlinelibrary.wiley.com/doi/full/10.1002/eji.200939578 [Accessed: 12 October 2023].
Barrett, J.C. et al. 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature genetics* 40(8), pp. 955–962. Available at: https://pubmed.ncbi.nlm.nih.gov/18587394/ [Accessed: 9 December 2023].
Bedoui, Y. et al. 2019. Methotrexate an Old Drug with New Tricks. *International Journal of Molecular Sciences* 20(20). Available at: /pmc/articles/PMC6834162/ [Accessed: 23 September 2023].
Benjamin, O., Bansal, P., Goyal, A. and Lappin, S.L. 2023. Disease-Modifying Antirheumatic Drugs (DMARD). *StatPearls*. Available at: https://www.ncbi.nlm.nih.gov/books/NBK507863/ [Accessed: 7 October 2023].
Berardi, S., Corrado, A., Maruotti, N., Cici, D. and Cantatore, F.P. 2021. Osteoblast role in the pathogenesis of rheumatoid arthritis. *Molecular Biology Reports* 48(3), p. 2843. Available at: /pmc/articles/PMC8060181/ [Accessed: 5 September 2023].
Biggioggero, M., Crotti, C., Becciolini, A. and Favalli, E.G. 2019. Tocilizumab in the treatment of rheumatoid arthritis: an evidence-based review and patient selection. *Drug Design, Development and Therapy* 13, p. 57. Available at: /pmc/articles/PMC6304084/ [Accessed: 17 September 2023].
Blanco, F.J. et al. 2017. Secukinumab in Active Rheumatoid Arthritis: A Phase III Randomized, Double-Blind, Active Comparator- and Placebo-Controlled Study. *Arthritis & rheumatology (Hoboken, N.J.)* 69(6), pp. 1144–1153. Available at: https://pubmed.ncbi.nlm.nih.gov/28217871/ [Accessed: 17 September 2023].
Bonnet, C.S., Williams, A.S., Gilbert, S.J., Harvey, A.K., Evans, B.A. and Mason, D.J. 2015. Ampa/kainate glutamate receptors contribute to inflammation, degeneration and pain related behaviour in inflammatory stages of arthritis. *Annals of the Rheumatic Diseases* 74(1), pp. 242–251. Available at: https://ard.bmj.com/content/74/1/242 [Accessed: 5 October 2023].
Braicu, C. et al. 2019. A Comprehensive Review on MAPK: A Promising Therapeutic Target in Cancer. *Cancers* 11(10), p. 1618. Available at: /pmc/articles/PMC6827047/ [Accessed: 14 October 2023].
Brand, D.D., Latham, K.A. and Rosloniec, E.F. 2007. Collagen-induced arthritis. *Nature Protocols 2007 2:5* 2(5), pp. 1269–1275. Available at: https://www.nature.com/articles/nprot.2007.173 [Accessed: 12 October 2023].
Van Den Broek, M., Huizinga, T.W.J., Dijkmans, B.A.C. and Allaart, C.F. 2011. Drug-free remission: is it already possible? *Current opinion in rheumatology* 23(3), pp. 266–272. Available at: https://pubmed.ncbi.nlm.nih.gov/21427578/ [Accessed: 7 October 2023].
Buch, M.H. et al. 2009. Mode of action of abatacept in rheumatoid arthritis patients having failed tumour necrosis factor blockade: a histological, gene expression and dynamic magnetic resonance imaging pilot study. *Annals of the rheumatic diseases* 68(7), pp. 1220–1227. Available at: https://pubmed.ncbi.nlm.nih.gov/18772191/ [Accessed: 25 August 2022].
Buch, M.H. 2018. Defining refractory rheumatoid arthritis. *Annals of the Rheumatic Diseases* 77(7), pp. 966–969. Available at: https://ard.bmj.com/content/77/7/966 [Accessed: 9 December 2023].
Buenrostro, J.D. et al. 2015. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523(7561), p. 486. Available at: /pmc/articles/PMC4685948/ [Accessed: 29 October 2023].

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. and Greenleaf, W.J. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods 2013 10:12* 10(12), pp. 1213–1218. Available at: https://www.nature.com/articles/nmeth.2688 [Accessed: 9 May 2022].
Butturini, E., de Prati, A.C. and Mariotto, S. 2020. Redox Regulation of STAT1 and STAT3 Signaling. *International Journal of Molecular Sciences* 21(19), pp. 1–18. Available at: /pmc/articles/PMC7582491/ [Accessed: 30 September 2023].
Casamassimi, A. and Ciccodicola, A. 2019. Transcriptional Regulation: Molecules, Involved Mechanisms, and Misregulation. *International Journal of Molecular Sciences* 20(6). Available at: /pmc/articles/PMC6471904/ [Accessed: 10 October 2023].
Castor, C.W. 1960. The microscopic structure of normal human synovial tissue. *Arthritis and rheumatism* 3(2), pp. 140–151. Available at: https://pubmed.ncbi.nlm.nih.gov/13808324/ [Accessed: 6 August 2022].
Castro, C.T. de et al. 2022. Real-world effectiveness of biological therapy in patients with rheumatoid arthritis: Systematic review and meta-analysis. *Frontiers in Pharmacology* 13, p. 927179. doi: 10.3389/FPHAR.2022.927179/BIBTEX.
Chemin, K., Gerstner, C. and Malmström, V. 2019. Effector functions of CD4+ T cells at the site of local autoimmune inflammation-lessons from rheumatoid arthritis. *Frontiers in Immunology* 10, p. 353. Available at: /pmc/articles/PMC6422991/ [Accessed: 9 September 2023].
Chen, J. et al. 2016. An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome medicine* 8(1). Available at: https://pubmed.ncbi.nlm.nih.gov/27102666/ [Accessed: 11 August 2022].
Choudhary, N., Bhatt, L.K. and Prabhavalkar, K.S. 2018. Experimental animal models for rheumatoid arthritis. *Immunopharmacology and immunotoxicology* 40(3), pp. 193–200. Available at: https://pubmed.ncbi.nlm.nih.gov/29433367/ [Accessed: 12 October 2023].
Choy, E.H., Kavanaugh, A.F. and Jones, S.A. 2013. The problem of choice: current biologic agents and future prospects in RA. *Nature Reviews Rheumatology 2013 9:3* 9(3), pp. 154–163. Available at: https://www.nature.com/articles/nrrheum.2013.8 [Accessed: 17 August 2022].
Coffey, C.M., Crowson, C.S., Myasoedova, E., Matteson, E.L. and Davis, J.M. 2019. Evidence of Diagnostic and Treatment Delay in Seronegative Rheumatoid Arthritis: Missing the Window of Opportunity. *Mayo Clinic proceedings* 94(11), p. 2241. Available at: /pmc/articles/PMC6947665/ [Accessed: 12 August 2023].
Cohen, S.B. et al. 2009. Evaluation of the efficacy and safety of pamapimod, a p38 MAP kinase inhibitor, in a double-blind, methotrexate-controlled study of patients with active rheumatoid arthritis. *Arthritis and rheumatism* 60(2), pp. 335–344. Available at: https://pubmed.ncbi.nlm.nih.gov/19180516/ [Accessed: 14 October 2023].
Collins, A.S., McCoy, C.E., Lloyd, A.T., O'Farrelly, C. and Stevenson, N.J. 2013. miR-19a: An Effective Regulator of SOCS3 and Enhancer of JAK-STAT Signalling. *PLOS ONE* 8(7), p. e69090. Available at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0069090 [Accessed: 2 April 2024].
Corces, M.R. et al. 2018. The chromatin accessibility landscape of primary human cancers. *Science* 362(6413). Available at: https://www.science.org/doi/10.1126/science.aav1898 [Accessed: 16 February 2023].

Corsiero, E., Nerviani, A., Bombardieri, M. and Pitzalis, C. 2016. Ectopic Lymphoid Structures: Powerhouse of Autoimmunity. *Frontiers in Immunology* 7(OCT), p. 17. Available at: /pmc/articles/PMC5066320/ [Accessed: 28 November 2023].
Crick, F. 1970. Central Dogma of Molecular Biology. *Nature 1970 227:5258* 227(5258), pp. 561–563. Available at: https://www.nature.com/articles/227561a0 [Accessed: 16 November 2023].
Croft, A.P. et al. 2019. Distinct fibroblast subsets drive inflammation and damage in arthritis. *Nature 2019 570:7760* 570(7760), pp. 246–251. Available at: https://www.nature.com/articles/s41586-019-1263-7 [Accessed: 5 September 2023].
Crowson, C.S. et al. 2011. The lifetime risk of adult-onset rheumatoid arthritis and other inflammatory autoimmune rheumatic diseases. *Arthritis and rheumatism* 63(3), pp. 633–639. Available at: https://pubmed.ncbi.nlm.nih.gov/21360492/ [Accessed: 1 August 2022].
Damjanov, N., Kauffman, R.S. and Spencer-Green, G.T. 2009. Efficacy, pharmacodynamics, and safety of VX-702, a novel p38 MAPK inhibitor, in rheumatoid arthritis: results of two randomized, double-blind, placebo-controlled clinical studies. *Arthritis and rheumatism* 60(5), pp. 1232–1241. Available at: https://pubmed.ncbi.nlm.nih.gov/19404957/ [Accessed: 14 October 2023].
Dann, G.P. et al. 2017. ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature 2017 548:7669* 548(7669), pp. 607–611. Available at: https://www.nature.com/articles/nature23671 [Accessed: 8 May 2022].
Dennis, G. et al. 2014. Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics. *Arthritis Research and Therapy* 16(2), pp. 1–18. Available at: https://arthritis-research.biomedcentral.com/articles/10.1186/ar4555 [Accessed: 14 October 2023].
Desai, R.J., Pawar, A., Khosrow-Khavar, F., Weinblatt, M.E. and Kim, S.C. 2022. Risk of venous thromboembolism associated with tofacitinib in patients with rheumatoid arthritis: a population-based cohort study. *Rheumatology (United Kingdom)* 61(1), pp. 121–130. doi: 10.1093/rheumatology/keab294.
Desai, R.J., Pawar, A., Weinblatt, M.E. and Kim, S.C. 2019. Comparative Risk of Venous Thromboembolism in Rheumatoid Arthritis Patients Receiving Tofacitinib Versus Those Receiving Tumor Necrosis Factor Inhibitors: An Observational Cohort Study. *Arthritis and Rheumatology* 71(6), pp. 892–900. doi: 10.1002/art.40798.
Edwards, J.C.W., Leandro, M.J. and Cambridge, G. 2005. B lymphocyte depletion in rheumatoid arthritis: targeting of CD20. *Current directions in autoimmunity* 8, pp. 175–192. Available at: https://pubmed.ncbi.nlm.nih.gov/15564721/ [Accessed: 24 August 2022].
Emery, P., Breedveld, F.C., Dougados, M., Kalden, J.R., Schiff, M.H. and Smolen, J.S. 2002. Early referral recommendation for newly diagnosed rheumatoid arthritis: evidence based development of a clinical guide. *Annals of the rheumatic diseases* 61(4), pp. 290–297. Available at: https://pubmed.ncbi.nlm.nih.gov/11874828/ [Accessed: 17 August 2022].
ENCODE Guidelines for Experiments Generating ChIP-seq Data. 2017. Available at: https://www.encodeproject.org/data-standards/chip-seq/ [Accessed: 1 October 2023].
Ernst, J. and Kellis, M. 2010. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature biotechnology* 28(8), p. 817. Available at: /pmc/articles/PMC2919626/ [Accessed: 12 November 2023].
Felsenfeld, G. 1978. Chromatin. *Nature* 271(5641), pp. 115–122. doi: 10.1038/271115a0.

Ferguson, F.M. and Gray, N.S. 2018. Kinase inhibitors: the road ahead. *Nature reviews. Drug discovery* 17(5), pp. 353–376. Available at: https://pubmed.ncbi.nlm.nih.gov/29545548/ [Accessed: 14 October 2023].
Flood, S., Parri, R., Williams, A., Duance, V. and Mason, D. 2007. Modulation of interleukin-6 and matrix metalloproteinase 2 expression in human fibroblast-like synoviocytes by functional ionotropic glutamate receptors. *Arthritis and rheumatism* 56(8), pp. 2523–2534. Available at: https://pubmed.ncbi.nlm.nih.gov/17665433/ [Accessed: 5 October 2023].
Frank-Bertoncelj, M. et al. 2017. Epigenetically-driven anatomical diversity of synovial fibroblasts guides joint-specific fibroblast functions. *Nature communications* 8. Available at: https://pubmed.ncbi.nlm.nih.gov/28332497/ [Accessed: 16 August 2022].
Fransen, J. and van Riel, P.L.C.M. 2009. The Disease Activity Score and the EULAR Response Criteria. *Rheumatic Disease Clinics of North America* 35(4), pp. 745–757. doi: 10.1016/j.rdc.2009.10.001.
Frisell, T., Saevarsdottir, S. and Askling, J. 2016. Family history of rheumatoid arthritis: an old concept with new developments. *Nature reviews. Rheumatology* 12(6), pp. 335–343. Available at: https://pubmed.ncbi.nlm.nih.gov/27098907/ [Accessed: 7 October 2023].
Fullerton, J.N. and Gilroy, D.W. 2016. Resolution of inflammation: a new therapeutic frontier. *Nature reviews. Drug discovery* 15(8), pp. 551–567. Available at: https://pubmed.ncbi.nlm.nih.gov/27020098/ [Accessed: 13 December 2022].
Furman, D. et al. 2019. Chronic inflammation in the etiology of disease across the life span. *Nature Medicine 2019 25:12* 25(12), pp. 1822–1832. Available at: https://www.nature.com/articles/s41591-019-0675-0 [Accessed: 13 December 2022].
Gabay, C. et al. 2013. Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (ADACTA): a randomised, double-blind, controlled phase 4 trial. *Lancet (London, England)* 381(9877), pp. 1541–1550. Available at: https://pubmed.ncbi.nlm.nih.gov/23515142/ [Accessed: 7 October 2023].
Gaffen, S.L. 2009. Role of IL-17 in the Pathogenesis of Rheumatoid Arthritis. *Current rheumatology reports* 11(5), p. 365. Available at: /pmc/articles/PMC2811488/ [Accessed: 9 December 2023].
Gangadharan, S., Mularoni, L., Fain-Thornton, J., Wheelan, S.J. and Craig, N.L. 2010. DNA transposon Hermes inserts into DNA in nucleosome-free regions in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 107(51), pp. 21966–21972. Available at: www.pnas.org/cgi/doi/10.1073/pnas.1016382107 [Accessed: 9 May 2022].
Genovese, M.C. et al. 2013. Efficacy and safety of secukinumab in patients with rheumatoid arthritis: a phase II, dose-finding, double-blind, randomised, placebo controlled study. *Annals of the rheumatic diseases* 72(6), pp. 863–869. Available at: https://pubmed.ncbi.nlm.nih.gov/22730366/ [Accessed: 15 October 2023].
Genovese, M.C. et al. 2014. A phase II randomized study of subcutaneous ixekizumab, an anti-interleukin-17 monoclonal antibody, in rheumatoid arthritis patients who were naive to biologic agents or had an inadequate response to tumor necrosis factor inhibitors. *Arthritis & rheumatology (Hoboken, N.J.)* 66(7), pp. 1693–1704. Available at: https://pubmed.ncbi.nlm.nih.gov/24623718/ [Accessed: 17 September 2023].
Gerosa, M., De Angelis, V., Riboldi, P. and Meroni, P.L. 2008. Rheumatoid Arthritis: A Female Challenge. *http://dx.doi.org/10.2217/17455057.4.2.195* 4(2), pp. 195–201. Available at: https://journals.sagepub.com/doi/10.2217/17455057.4.2.195 [Accessed: 9 September 2023].

Grandi, F.C., Modi, H., Kampman, L. and Corces, M.R. 2022. Chromatin accessibility profiling by ATAC-seq. *Nature Protocols 2022*, pp. 1–35. Available at: https://www.nature.com/articles/s41596-022-00692-9 [Accessed: 9 May 2022].
Gregersen, P.K., Silver, J. and Winchester, R.J. 1987. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis and rheumatism* 30(11), pp. 1205–1213. Available at: https://pubmed.ncbi.nlm.nih.gov/2446635/ [Accessed: 16 August 2022].
Gujral, P., Mahajan, V., Lissaman, A.C. and Ponnampalam, A.P. 2020. Histone acetylation and the role of histone deacetylases in normal cyclic endometrium. *Reproductive Biology and Endocrinology* 18(1), pp. 1–11. Available at: https://rbej.biomedcentral.com/articles/10.1186/s12958-020-00637-5 [Accessed: 10 September 2023].
Hansel, T.T., Kropshofer, H., Singer, T., Mitchell, J.A. and George, A.J.T. 2010. The safety and side effects of monoclonal antibodies. *Nature Reviews Drug Discovery 2010 9:4* 9(4), pp. 325–338. Available at: https://www.nature.com/articles/nrd3003 [Accessed: 17 September 2023].
Hara, M. et al. 2014. Safety and efficacy of combination therapy of iguratimod with methotrexate for patients with active rheumatoid arthritis with an inadequate response to methotrexate: an open-label extension of a randomized, double-blind, placebo-controlled trial. *Modern rheumatology* 24(3), pp. 410–418. Available at: https://pubmed.ncbi.nlm.nih.gov/24252050/ [Accessed: 14 October 2023].
Haraoui, B. and Bykerk, V. 2007. Etanercept in the treatment of rheumatoid arthritis. *Therapeutics and Clinical Risk Management* 3(1), p. 99. Available at: /pmc/articles/PMC1936291/ [Accessed: 17 September 2023].
Harrington, R., Al Nokhatha, S.A. and Conway, R. 2020. JAK Inhibitors in Rheumatoid Arthritis: An Evidence-Based Review on the Emerging Clinical Data. *Journal of Inflammation Research* 13, p. 519. Available at: /pmc/articles/PMC7500842/ [Accessed: 25 August 2022].
Harwardt, T. et al. 2016. Human Cytomegalovirus Immediate-Early 1 Protein Rewires Upstream STAT3 to Downstream STAT1 Signaling Switching an IL6-Type to an IFNγ-Like Response. *PLoS Pathogens* 12(7). Available at: /pmc/articles/PMC4936752/ [Accessed: 30 September 2023].
Hill, R.J. et al. 2008. Pamapimod, a novel p38 mitogen-activated protein kinase inhibitor: preclinical analysis of efficacy and selectivity. *The Journal of pharmacology and experimental therapeutics* 327(3), pp. 610–619. Available at: https://pubmed.ncbi.nlm.nih.gov/18776065/ [Accessed: 14 October 2023].
Hirahara, K. et al. 2015. Asymmetric Action of STAT Transcription Factors Drives Transcriptional Outputs and Cytokine Specificity. *Immunity* 42(5), pp. 877–889. Available at: https://pubmed.ncbi.nlm.nih.gov/25992861/ [Accessed: 29 April 2021].
Holers, V.M. 2013. Autoimmunity to citrullinated proteins and the initiation of rheumatoid arthritis. *Current Opinion in Immunology* 25(6), pp. 728–735. doi: 10.1016/J.COI.2013.09.018.
Holoshitz, J. 2010. The Rheumatoid Arthritis HLA-DRB1 Shared Epitope. *Current opinion in rheumatology* 22(3), p. 293. Available at: /pmc/articles/PMC2921962/ [Accessed: 7 October 2023].
Hu, X. and Ivashkiv, L.B. 2009. Cross-regulation of Signaling and Immune Responses by IFN-γ and STAT1. *Immunity* 31(4), p. 539. Available at: /pmc/articles/PMC2774226/ [Accessed: 30 September 2023].

Humby, F. et al. 2009. Ectopic lymphoid structures support ongoing production of class-switched autoantibodies in rheumatoid synovium. *PLoS medicine* 6(1), pp. 0059–0075. Available at: https://pubmed.ncbi.nlm.nih.gov/19143467/ [Accessed: 9 August 2022].
Hurst, S.M. et al. 2001. IL-6 and Its Soluble Receptor Orchestrate a Temporal Switch in the Pattern of Leukocyte Recruitment Seen during Acute Inflammation. *Immunity* 14(6), pp. 705–714. Available at: http://www.cell.com/article/S1074761301001510/fulltext [Accessed: 8 October 2023].
Husmann, D. and Gozani, O. 2019. Histone lysine methyltransferases in biology and disease. *Nature Structural & Molecular Biology 2019 26:10* 26(10), pp. 880–889. Available at: https://www.nature.com/articles/s41594-019-0298-7 [Accessed: 2 November 2023].
Ingegnoli, F., Coletto, L.A., Scotti, I., Compagnoni, R., Randelli, P.S. and Caporali, R. 2021. The Crucial Questions on Synovial Biopsy: When, Why, Who, What, Where, and How? *Frontiers in Medicine* 8, p. 705382. Available at: /pmc/articles/PMC8377390/ [Accessed: 14 October 2023].
Ishikawa, H. 2017. The latest treatment strategy for the rheumatoid hand deformity. *Journal of Orthopaedic Science* 22(4), pp. 583–592. doi: 10.1016/J.JOS.2017.02.007.
Jang, D.I. et al. 2021. The Role of Tumor Necrosis Factor Alpha (TNF-α) in Autoimmune Disease and Current TNF-α Inhibitors in Therapeutics. *International Journal of Molecular Sciences* 22(5), pp. 1–16. Available at: /pmc/articles/PMC7962638/ [Accessed: 9 December 2023].
Jenkins, R.H., Hughes, S.T.O., Figueras, A.C. and Jones, S.A. 2021. Unravelling the broader complexity of IL-6 involvement in health and disease. *Cytokine* 148, p. 155684. doi: 10.1016/J.CYTO.2021.155684.
Jiang, X., Alfredsson, L., Klareskog, L. and Bengtsson, C. 2014. Smokeless tobacco (moist snuff) use and the risk of developing rheumatoid arthritis: Results from a case-control study. *Arthritis Care and Research* 66(10), pp. 1582–1586. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.22325 [Accessed: 11 August 2022].
Jin, B., Li, Y. and Robertson, K.D. 2011. DNA Methylation: Superior or Subordinate in the Epigenetic Hierarchy? *Genes & Cancer* 2(6), p. 607. Available at: /pmc/articles/PMC3174260/ [Accessed: 10 September 2023].
Johnsson, H. and Najm, A. 2021. Synovial biopsies in clinical practice and research: current developments and perspectives. *Clinical Rheumatology* 40(7), p. 2593. Available at: /pmc/articles/PMC8189968/ [Accessed: 15 September 2023].
Jones, G.W. et al. 2010. Loss of CD4+ T cell IL-6R expression during inflammation underlines a role for IL-6 trans signaling in the local maintenance of Th17 cells. *Journal of immunology (Baltimore, Md. : 1950)* 184(4), pp. 2130–2139. Available at: https://pubmed.ncbi.nlm.nih.gov/20083667/ [Accessed: 10 July 2022].
Jones, G.W. et al. 2015. Interleukin-27 inhibits ectopic lymphoid-like structure development in early inflammatory arthritis. *Journal of Experimental Medicine* 212(11), pp. 1793–1802. Available at: www.jem.org/cgi/doi/10.1084/jem.20132307 [Accessed: 15 September 2022].
Jones, G.W., Greenhill, C.J. and Williams, J.O. 2013. Exacerbated inflammatory arthritis in response to hyperactive gp130 signalling is independent of IL-17A. *Ann Rheum Dis* 72, pp. 1738–1742. Available at: http://ard.bmj.com/ [Accessed: 14 October 2023].
Jones, G.W., Hill, D.G., Sime, K. and Williams, A.S. 2018. In vivo models for inflammatory arthritis. *Methods in Molecular Biology* 1725, pp. 101–118. Available at: https://link.springer.com/protocol/10.1007/978-1-4939-7568-6_9 [Accessed: 6 September 2022].

Jung, S.M., Kim, K.W., Yang, C.W., Park, S.H., Ju, J.H. and Mamura, M. 2014. Cytokine-Mediated Bone Destruction in Rheumatoid Arthritis. *Journal of Immunology Research* 2014. Available at: /pmc/articles/PMC4176903/ [Accessed: 7 October 2023].
Just, S.A. et al. 2018. Patient-reported outcomes and safety in patients undergoing synovial biopsy: comparison of ultrasound-guided needle biopsy, ultrasound-guided portal and forceps and arthroscopic-guided synovial biopsy techniques in five centres across Europe. *RMD Open* 4(2), p. e000799. Available at: https://rmdopen.bmj.com/content/4/2/e000799 [Accessed: 14 October 2023].
Källberg, H. et al. 2011. Smoking is a major preventable risk factor for rheumatoid arthritis: estimations of risks after various exposures to cigarette smoke. *Annals of the rheumatic diseases* 70(3), pp. 508–511. Available at: https://pubmed.ncbi.nlm.nih.gov/21149499/ [Accessed: 11 August 2022].
Kaneko, A. 2013. Tocilizumab in rheumatoid arthritis: efficacy, safety and its place in therapy. *Therapeutic Advances in Chronic Disease* 4(1), p. 15. Available at: /pmc/articles/PMC3539265/ [Accessed: 17 September 2023].
Katchamart, W., Johnson, S., Lin, H.J.L., Phumethum, V., Salliot, C. and Bombardier, C. 2010. Predictors for remission in rheumatoid arthritis patients: A systematic review. *Arthritis Care and Research* 62(8), pp. 1128–1143. doi: 10.1002/ACR.20188.
Kay, J. and Upchurch, K.S. 2012. ACR/EULAR 2010 rheumatoid arthritis classification criteria. *Rheumatology* 51(suppl_6), pp. vi5–vi9. Available at: https://dx.doi.org/10.1093/rheumatology/kes279 [Accessed: 21 September 2023].
Keffer, J., Probert, L., Cazlaris, H., Georgopoulos, S., Kaslaris, E., Kioussis, D. and Kollias, G. 1991. Transgenic mice expressing human tumour necrosis factor: a predictive genetic model of arthritis. *The EMBO Journal* 10(13), pp. 4025–4031. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/j.1460-2075.1991.tb04978.x [Accessed: 12 October 2023].
Kershaw, N.J. et al. 2013. SOCS3 binds specific receptor–JAK complexes to control cytokine signaling by direct kinase inhibition. *Nature structural & molecular biology* 20(4), p. 469. Available at: /pmc/articles/PMC3618588/ [Accessed: 2 April 2024].
Khosrow-Khavar, F., Kim, S.C., Lee, H., Lee, S.B. and Desai, R.J. 2022. Tofacitinib and risk of cardiovascular outcomes: results from the Safety of TofAcitinib in Routine care patients with Rheumatoid Arthritis (STAR-RA) study. *Annals of the rheumatic diseases* 81(6), pp. 798–804. doi: 10.1136/annrheumdis-2021-221915.
Kiener, H.P., Niederreiter, B., Lee, D.M., Jimenez-Boj, E., Smolen, J.S. and Brenner, M.B. 2009. Cadherin 11 promotes invasive behavior of fibroblast-like synoviocytes. *Arthritis and rheumatism* 60(5), pp. 1305–1310. Available at: https://pubmed.ncbi.nlm.nih.gov/19404963/ [Accessed: 9 August 2022].
Kim, M., Choe, Y.H. and Lee, S. Il. 2022. Lessons From the Success and Failure of Targeted Drugs for Rheumatoid Arthritis: Perspectives for Effective Basic and Translational Research. *Immune Network* 22(1). Available at: /pmc/articles/PMC8901706/ [Accessed: 15 October 2023].
Kim, S. et al. 2003. Stat1 functions as a cytoplasmic attenuator of Runx2 in the transcriptional program of osteoblast differentiation. *Genes & Development* 17(16), p. 1979. Available at: /pmc/articles/PMC196253/ [Accessed: 31 August 2023].
Kinne, R.W. et al. 1995. Synovial fibroblast-like cells strongly express jun-B and C-fos proto-oncogenes in rheumatoid- and osteoarthritis. *Scandinavian journal of rheumatology.*

Supplement 101(S101), pp. 121–125. Available at: https://pubmed.ncbi.nlm.nih.gov/7747113/ [Accessed: 27 November 2023].
Klemm, S.L., Shipony, Z. and Greenleaf, W.J. 2019. Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics 2018 20:4* 20(4), pp. 207–220. Available at: https://www.nature.com/articles/s41576-018-0089-8 [Accessed: 7 May 2022].
Kotas, M.E. and Medzhitov, R. 2015. Homeostasis, inflammation, and disease susceptibility. *Cell* 160(5), pp. 816–827. Available at: https://pubmed.ncbi.nlm.nih.gov/25723161/ [Accessed: 13 December 2022].
Krabben, A., Huizinga, T.W.J. and Mil, A.H.M. 2015. Biomarkers for radiographic progression in rheumatoid arthritis. *Current pharmaceutical design* 21(2), pp. 147–169. Available at: https://pubmed.ncbi.nlm.nih.gov/25163742/ [Accessed: 16 August 2022].
Krause, M.L. and Makol, A. 2016. Management of rheumatoid arthritis during pregnancy: challenges and solutions. *Open Access Rheumatology : Research and Reviews* 8, p. 23. Available at: /pmc/articles/PMC5098768/ [Accessed: 9 September 2023].
Krošel, M. et al. 2023. The histone acetyl transferases CBP and p300 regulate stress response pathways in synovial fibroblasts at transcriptional and functional levels. *Scientific Reports 2023 13:1* 13(1), pp. 1–10. Available at: https://www.nature.com/articles/s41598-023-44412-z [Accessed: 30 November 2023].
Kwock, J.T. et al. 2020. IL-27 signaling activates skin cells to induce innate antiviral proteins and protects against Zika virus infection. *Science Advances* 6(14). Available at: /pmc/articles/PMC7112749/ [Accessed: 2 September 2023].
Van De Laar, C.J., Voshaar, M.A.H.O., Fakhouri, W.K.H., Zaremba-Pechmann, L., De Leonardis, F., De La Torre, I. and Van De Laar, M.A.F.J. 2020. Cost-Effectiveness of a JAK1/JAK2 Inhibitor vs a Biologic Disease-Modifying Antirheumatic Drug (bDMARD) in a Treat-to-Target Strategy for Rheumatoid Arthritis. *ClinicoEconomics and Outcomes Research: CEOR* 12, p. 213. Available at: /pmc/articles/PMC7167259/ [Accessed: 16 September 2023].
Landewé, R.B.M. et al. 2002. COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. *Arthritis and rheumatism* 46(2), pp. 347–356. Available at: https://pubmed.ncbi.nlm.nih.gov/11840436/ [Accessed: 30 August 2022].
Laragione, T., Brenner, M., Sherry, B. and Gulko, P.S. 2011. CXCL10 and its receptor CXCR3 regulate synovial fibroblast invasion in rheumatoid arthritis. *Arthritis and rheumatism* 63(11), p. 3274. Available at: /pmc/articles/PMC3205193/ [Accessed: 28 November 2023].
Latchman, D.S. 1997. Transcription factors: An overview. *The International Journal of Biochemistry & Cell Biology* 29(12), pp. 1305–1312. doi: 10.1016/S1357-2725(97)00085-X.
Lee, C.K., Shibata, Y., Rao, B., Strahl, B.D. and Lieb, J.D. 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nature Genetics 2004 36:8* 36(8), pp. 900–905. Available at: https://www.nature.com/articles/ng1400 [Accessed: 8 May 2022].
Lee, J.C. et al. 2013. Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell* 155(1), pp. 57–69. Available at: https://pubmed.ncbi.nlm.nih.gov/24035192/ [Accessed: 16 August 2022].
Lee, K.M.C., Achuthan, A.A. and Hamilton, J.A. 2020. GM-CSF: A Promising Target in Inflammation and Autoimmunity. *ImmunoTargets and Therapy* 9, p. 225. Available at: /pmc/articles/PMC7605919/ [Accessed: 9 December 2023].

Lenschow, D.J. and Bluestone, J.A. 1993. T cell co-stimulation and in vivo tolerance. *Current opinion in immunology* 5(5), pp. 747–752. Available at: https://pubmed.ncbi.nlm.nih.gov/7694594/ [Accessed: 25 August 2022].
Lewis, M.J., Barnes, M.R., Blighe, K., Taylor, P.C., Townsend, M.J. and Correspondence, P. 2019. Molecular Portraits of Early Rheumatoid Arthritis Identify Clinical and Treatment Response Phenotypes In Brief. Available at: https://doi.org/10.1016/j.celrep.2019.07.091 [Accessed: 14 October 2023].
Li, G., Tian, Y. and Zhu, W.G. 2020. The Roles of Histone Deacetylases and Their Inhibitors in Cancer Therapy. *Frontiers in Cell and Developmental Biology* 8, p. 576946. doi: 10.3389/FCELL.2020.576946/BIBTEX.
Liu, B. et al. 2004. PIAS1 selectively inhibits interferon-inducible genes and is important in innate immunity. *Nature Immunology 2004 5:9* 5(9), pp. 891–898. Available at: https://www.nature.com/articles/ni1104 [Accessed: 30 September 2023].
Liu, K.D., Gaffen, S.L. and Goldsmith, M.A. 1998. JAK/STAT signaling by cytokine receptors. *Current Opinion in Immunology* 10(3), pp. 271–278. doi: 10.1016/S0952-7915(98)80165-9.
Liu, L. et al. 2019. Folate Supplementation for Methotrexate Therapy in Patients With Rheumatoid Arthritis: A Systematic Review. *Journal of clinical rheumatology : practical reports on rheumatic & musculoskeletal diseases* 25(5), pp. 197–202. Available at: https://pubmed.ncbi.nlm.nih.gov/29975207/ [Accessed: 7 October 2023].
Liu, Y. et al. 2013. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nature Biotechnology 2013 31:2* 31(2), pp. 142–147. Available at: https://www.nature.com/articles/nbt.2487 [Accessed: 16 August 2022].
Malmström, V., Catrina, A.I. and Klareskog, L. 2016. The immunopathogenesis of seropositive rheumatoid arthritis: from triggering to targeting. *Nature Reviews Immunology 2016 17:1* 17(1), pp. 60–75. Available at: https://www.nature.com/articles/nri.2016.124 [Accessed: 7 October 2023].
Martin, P. et al. 2015. Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nature communications* 6. Available at: https://pubmed.ncbi.nlm.nih.gov/26616563/ [Accessed: 16 August 2022].
Martin-Mola, E., Balsa, A., García-Vicuna, R., Gómez-Reino, J., González-Gay, M.A., Sanmartí, R. and Loza, E. 2016. Anti-citrullinated peptide antibodies and their value for predicting responses to biologic agents: a review. *Rheumatology International 2016 36:8* 36(8), pp. 1043–1063. Available at: https://link.springer.com/article/10.1007/s00296-016-3506-3 [Accessed: 7 October 2023].
Masi, A.T. 1983. Articular patterns in the early course of rheumatoid arthritis. *The American journal of medicine* 75(6A), pp. 16–26. Available at: https://pubmed.ncbi.nlm.nih.gov/6660237/ [Accessed: 2 August 2022].
Massalska, M., Maslinski, W. and Ciechomska, M. 2020. Small Molecule Inhibitors in the Treatment of Rheumatoid Arthritis and Beyond: Latest Updates and Potential Strategy for Fighting COVID-19. *Cells* 9(8). Available at: /pmc/articles/PMC7464410/ [Accessed: 16 September 2023].
Matthews, A.G., Li, J., He, C., Ott, J. and Andrade, M. de. 2009. Adjusting for HLA-DRβ1 in a genome-wide association analysis of rheumatoid arthritis and related biomarkers. *BMC Proceedings* 3(Suppl 7), p. S12. Available at: /pmc/articles/PMC2795892/ [Accessed: 9 September 2023].

McAllister, K., Eyre, S. and Orozco, G. 2011. Genetics of rheumatoid arthritis: GWAS and beyond. *Open Access Rheumatology : Research and Reviews* 3, p. 31. Available at: /pmc/articles/PMC5074784/ [Accessed: 9 December 2023].
McInnes, I.B. and Gravallese, E.M. 2021. Immune-mediated inflammatory disease therapeutics: past, present and future. *Nature Reviews Immunology 2021 21:10* 21(10), pp. 680–686. Available at: https://www.nature.com/articles/s41577-021-00603-1 [Accessed: 2 August 2023].
McInnes, I.B. and Schett, G. 2007. Cytokines in the pathogenesis of rheumatoid arthritis. *Nature Reviews Immunology 2007 7:6* 7(6), pp. 429–442. Available at: https://www.nature.com/articles/nri2094 [Accessed: 22 August 2022].
McInnes, I.B. and Schett, G. 2011. The Pathogenesis of Rheumatoid Arthritis. *https://doi.org/10.1056/NEJMra1004965* 365(23), pp. 2205–2219. Available at: https://www.nejm.org/doi/full/10.1056/nejmra1004965 [Accessed: 1 August 2022].
McInnes, I.B. and Schett, G. 2017. Pathogenetic insights from the treatment of rheumatoid arthritis. *The Lancet* 389(10086), pp. 2328–2337. doi: 10.1016/S0140-6736(17)31472-1.
Mease, P.J. 2007. Adalimumab in the treatment of arthritis. *Therapeutics and Clinical Risk Management* 3(1), p. 133. Available at: /pmc/articles/PMC1936294/ [Accessed: 8 October 2023].
Meng, W. et al. 2017. DNA methylation mediates genotype and smoking interaction in the development of anti-citrullinated peptide antibody-positive rheumatoid arthritis. *Arthritis research & therapy* 19(1). Available at: https://pubmed.ncbi.nlm.nih.gov/28356135/ [Accessed: 16 August 2022].
Metzker, M.L. 2009. Sequencing technologies — the next generation. *Nature Reviews Genetics 2010 11:1* 11(1), pp. 31–46. Available at: https://www.nature.com/articles/nrg2626 [Accessed: 13 September 2022].
Mezger, A. et al. 2018. High-throughput chromatin accessibility profiling at single-cell resolution. *Nature Communications* 9(1). Available at: /pmc/articles/PMC6128862/ [Accessed: 29 October 2023].
Millrine, D. et al. 2023. Th1 cells alter the inflammatory signature of IL-6 by channeling STAT transcription factors to Alu-like retroelements. *bioRxiv*, p. 2022.07.18.499157. Available at: https://www.biorxiv.org/content/10.1101/2022.07.18.499157v3 [Accessed: 11 July 2023].
Morris, R., Kershaw, N.J. and Babon, J.J. 2018. The molecular details of cytokine signaling via the JAK/STAT pathway. *Protein Science : A Publication of the Protein Society* 27(12), p. 1984. Available at: /pmc/articles/PMC6237706/ [Accessed: 24 September 2023].
Na, H.S. et al. 2020. The establishment of a rheumatoid arthritis primate model in Macaca fascicularis. *Journal of Translational Medicine* 18(1), pp. 1–10. Available at: https://translational-medicine.biomedcentral.com/articles/10.1186/s12967-020-02402-z [Accessed: 12 October 2023].
Najm, A. and McInnes, I.B. 2021. IL-23 orchestrating immune cell activation in arthritis. *Rheumatology (Oxford, England)* 60(Suppl 4), p. iv4. Available at: /pmc/articles/PMC8527242/ [Accessed: 12 August 2023].
NICE. 2021. Adalimumab, etanercept, infliximab and abatacept for treating moderate rheumatoid arthritis after conventional DMARDs have failed. Available at: www.nice.org.uk/guidance/ta715 [Accessed: 14 October 2023].
Nowell, M.A. et al. 2009. Therapeutic targeting of IL-6 trans signaling counteracts STAT3 control of experimental inflammatory arthritis. *Journal of immunology (Baltimore, Md. :*

1950) 182(1), pp. 613–622. Available at: https://pubmed.ncbi.nlm.nih.gov/19109195/ [Accessed: 2 March 2023].
O'Brown, Z.K., Van Nostrand, E.L., Higgins, J.P. and Kim, S.K. 2015. The Inflammatory Transcription Factors NFκB, STAT1 and STAT3 Drive Age-Associated Transcriptional Changes in the Human Kidney. *PLOS Genetics* 11(12), p. e1005734. Available at: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1005734 [Accessed: 3 December 2023].
Okada, Y. et al. 2013. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature 2013 506:7488* 506(7488), pp. 376–381. Available at: https://www.nature.com/articles/nature12873 [Accessed: 9 September 2023].
Okada, Y. et al. 2014. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 506(7488), pp. 376–381. Available at: https://pubmed.ncbi.nlm.nih.gov/24390342/ [Accessed: 16 August 2022].
Olins, D.E. and Olins, A.L. 2003. Chromatin history: our view from the bridge. *Nature Reviews Molecular Cell Biology 2003 4:10* 4(10), pp. 809–814. Available at: https://www.nature.com/articles/nrm1225 [Accessed: 8 May 2022].
Oliver, J., Plant, D., Webster, A.P. and Barton, A. 2015. Genetic and genomic markers of anti-TNF treatment response in rheumatoid arthritis. *Biomarkers in medicine* 9(6), pp. 499–512. Available at: https://pubmed.ncbi.nlm.nih.gov/26079957/ [Accessed: 16 August 2022].
Onuora, S. 2014. Methotrexate and bridging glucocorticoids in early RA. *Nature Reviews Rheumatology 2014 10:12* 10(12), pp. 698–698. Available at: https://www.nature.com/articles/nrrheum.2014.197 [Accessed: 7 October 2023].
O'Shea, J.J., Lahesmaa, R., Vahedi, G., Laurence, A. and Kanno, Y. 2011. Genomic views of STAT function in CD4+ T helper cell differentiation. *Nature Reviews Immunology 2011 11:4* 11(4), pp. 239–250. Available at: https://www.nature.com/articles/nri2958 [Accessed: 30 September 2023].
O'Shea, J.J. and Plenge, R. 2012. JAK and STAT Signaling Molecules in Immunoregulation and Immune-Mediated Disease. *Immunity* 36(4), pp. 542–550. doi: 10.1016/J.IMMUNI.2012.03.014.
Ostrowska, M., Maśliński, W., Prochorec-Sobieszek, M., Nieciecki, M. and Sudoł-Szopińska, I. 2018. Cartilage and bone damage in rheumatoid arthritis. *Reumatologia* 56(2), p. 111. Available at: /pmc/articles/PMC5974634/ [Accessed: 10 September 2023].
Panigrahi, A. and O'Malley, B.W. 2021. Mechanisms of enhancer action: the known and the unknown. *Genome Biology 2021 22:1* 22(1), pp. 1–30. Available at: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02322-1 [Accessed: 13 September 2022].
Papadaki, M., Rinotas, V., Violitzi, F., Thireou, T., Panayotou, G., Samiotaki, M. and Douni, E. 2019. New insights for RANKL as a proinflammatory modulator in modeled inflammatory arthritis. *Frontiers in Immunology* 10(FEB), p. 439489. doi: 10.3389/FIMMU.2019.00097/BIBTEX.
Park, P.J. 2009. ChIP–seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics 2009 10:10* 10(10), pp. 669–680. Available at: https://www.nature.com/articles/nrg2641 [Accessed: 15 September 2022].
Peng, L. et al. 2020. Expression levels of CXCR4 and CXCL12 in patients with rheumatoid arthritis and its correlation with disease activity. *Experimental and Therapeutic Medicine* 20(3), p. 1925. Available at: /pmc/articles/PMC7401245/ [Accessed: 27 November 2023].

Perkins, D.J., William, E., Clair, S.T., Misukonis, M.A. and Weinberg, J.B. 1998. REDUCTION OF NOS2 OVEREXPRESSION IN RHEUMATOID ARTHRITIS PATIENTS TREATED WITH ANTI-TUMOR NECROSIS FACTOR a! MONOCLONAL ANTIBODY (cA2). *ARTHRITIS & RHEUMATISM* 41(12), pp. 2205–2210. Available at: https://onlinelibrary.wiley.com/doi/10.1002/1529-0131 [Accessed: 27 November 2023].
Pesu, M., Laurence, A., Kishore, N., Zwillich, S.H., Chan, G. and O'Shea, J.J. 2008. Therapeutic targeting of Janus kinases. *Immunological reviews* 223(1), pp. 132–142. Available at: https://pubmed.ncbi.nlm.nih.gov/18613833/ [Accessed: 25 August 2022].
Pfeifle, R. et al. 2016. Regulation of autoantibody activity by the IL-23–TH17 axis determines the onset of autoimmune disease. *Nature Immunology 2016 18:1* 18(1), pp. 104–113. Available at: https://www.nature.com/articles/ni.3579 [Accessed: 7 October 2023].
Pitzalis, C., Choy, E.H.S. and Buch, M.H. 2020. Transforming clinical trials in rheumatology: towards patient-centric precision medicine. *Nature Reviews Rheumatology 2020 16:10* 16(10), pp. 590–599. Available at: https://www.nature.com/articles/s41584-020-0491-4 [Accessed: 2 December 2023].
Pitzalis, C., Jones, G.W., Bombardieri, M. and Jones, S.A. 2014. Ectopic lymphoid-like structures in infection, cancer and autoimmunity. *Nature Reviews Immunology 2014 14:7* 14(7), pp. 447–462. Available at: https://www.nature.com/articles/nri3700 [Accessed: 8 September 2023].
Pitzalis, C., Kelly, S. and Humby, F. 2013. New learnings on the pathophysiology of RA from synovial biopsies. *Current Opinion in Rheumatology* 25(3), pp. 334–344. doi: 10.1097/BOR.0B013E32835FD8EB.
Prins, J.R., Gomez-Lopez, N. and Robertson, S.A. 2012. Interleukin-6 in pregnancy and gestational disorders. *Journal of Reproductive Immunology* 95(1–2), pp. 1–14. doi: 10.1016/J.JRI.2012.05.004.
Przanowski, P. et al. 2014. The signal transducers Stat1 and Stat3 and their novel target Jmjd3 drive the expression of inflammatory genes in microglia. *Journal of Molecular Medicine* 92(3), pp. 239–254. Available at: https://link.springer.com/article/10.1007/s00109-013-1090-5 [Accessed: 3 December 2023].
Radner, H., Lesperance, T., Accortt, N.A. and Solomon, D.H. 2017. Incidence and Prevalence of Cardiovascular Risk Factors Among Patients With Rheumatoid Arthritis, Psoriasis, or Psoriatic Arthritis. *Arthritis Care & Research* 69(10), pp. 1510–1518. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.23171 [Accessed: 8 September 2023].
Raison, C.L., Capuron, L. and Miller, A.H. 2006. Cytokines sing the blues: inflammation and the pathogenesis of depression. *Trends in Immunology* 27(1), pp. 24–31. doi: 10.1016/J.IT.2005.11.006.
Randen, I. et al. 1992. Clonally related IgM rheumatoid factors undergo affinity maturation in the rheumatoid synovial tissue. *The Journal of Immunology* 148(10).
Raza, K. and Filer, A. 2015. The therapeutic window of opportunity in rheumatoid arthritis: does it ever close? *Annals of the Rheumatic Diseases* 74(5), pp. 793–794. Available at: https://ard.bmj.com/content/74/5/793 [Accessed: 15 September 2023].
Raza, K. and Gerlag, D.M. 2014. Preclinical Inflammatory Rheumatic Diseases: An Overview and Relevant Nomenclature. *Rheumatic Disease Clinics of North America* 40(4), pp. 569–580. doi: 10.1016/J.RDC.2014.07.001.
Redlich, K. and Smolen, J.S. 2012. Inflammatory bone loss: pathogenesis and therapeutic intervention. *Nature reviews. Drug discovery* 11(3), pp. 234–250. Available at: https://pubmed.ncbi.nlm.nih.gov/22378270/ [Accessed: 9 August 2022].

Reed, E. et al. 2020. Presence of autoantibodies in “seronegative” rheumatoid arthritis associates with classical risk factors and high disease activity. *Arthritis Research and Therapy* 22(1), pp. 1–11. Available at: https://arthritis-research.biomedcentral.com/articles/10.1186/s13075-020-02191-2 [Accessed: 2 November 2023].
Remmers, E.F. et al. 2007. STAT4 and the Risk of Rheumatoid Arthritis and Systemic Lupus Erythematosus. *The New England journal of medicine* 357(10), p. 977. Available at: /pmc/articles/PMC2630215/ [Accessed: 15 October 2023].
Reske, J.J., Wilson, M.R. and Chandler, R.L. 2020. ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. *Epigenetics & Chromatin* 13(1), p. 22. doi: 10.1186/s13072-020-00342-y.
Rider, P., Carmi, Y. and Cohen, I. 2016. Biologics for Targeting Inflammatory Cytokines, Clinical Uses, and Limitations. *International Journal of Cell Biology* 2016. Available at: /pmc/articles/PMC5204077/ [Accessed: 8 October 2023].
Rivellese, F. et al. 2022. Rituximab versus tocilizumab in rheumatoid arthritis: synovial biopsy-based biomarker analysis of the phase 4 R4RA randomized trial. *Nature Medicine 2022 28:6* 28(6), pp. 1256–1268. Available at: https://www.nature.com/articles/s41591-022-01789-0 [Accessed: 14 October 2023].
Rose-John, S., Jenkins, B.J., Garbers, C., Moll, J.M. and Scheller, J. 2023. Targeting IL-6 trans-signalling: past, present and future prospects. *Nature Reviews Immunology 2023*, pp. 1–16. Available at: https://www.nature.com/articles/s41577-023-00856-y [Accessed: 17 September 2023].
Sakaguchi, S., Takahashi, T., Hata, H., Nomura, T. and Sakaguchi, N. 2003. SKG mice, a new genetic model of rheumatoid arthritis. *Arthritis Research & Therapy* 5(Suppl 3), p. 10. Available at: /pmc/articles/PMC2833801/ [Accessed: 19 November 2023].
Sanayama, Y. et al. 2014. Prediction of therapeutic responses to tocilizumab in patients with rheumatoid arthritis: biomarkers identified by analysis of gene expression in peripheral blood mononuclear cells using genome-wide DNA microarray. *Arthritis & rheumatology (Hoboken, N.J.)* 66(6), pp. 1421–1431. Available at: https://pubmed.ncbi.nlm.nih.gov/24591094/ [Accessed: 2 September 2023].
Saraiva, F. 2021. Ultrasound-Guided Synovial Biopsy: A Review. *Frontiers in Medicine* 8, p. 632224. Available at: /pmc/articles/PMC8100029/ [Accessed: 15 September 2023].
Satoh, K., Kikuchi, S., Sekimata, M., Kabuyama, Y., Homma, M.K. and Homma, Y. 2001. Involvement of ErbB-2 in rheumatoid synovial cell growth. *Arthritis & Rheumatism* 44(2), pp. 260–265. doi: 10.1002/1529-0131(200102)44:2<260::AID-ANR42>3.0.CO;2-P.
Scheinman, R. 2013. NF-κB and Rheumatoid Arthritis: Will Understanding Genetic Risk Lead to a Therapeutic Reward? *Forum on immunopathological diseases and therapeutics* 4(2), p. 93. Available at: /pmc/articles/PMC3963006/ [Accessed: 28 November 2023].
Schurich, A., Raine, C., Morris, V. and Ciurtin, C. 2018. The role of IL-12/23 in T cell–related chronic inflammation: implications of immunodeficiency and therapeutic blockade. *Rheumatology* 57(2), pp. 246–254. Available at: https://dx.doi.org/10.1093/rheumatology/kex186 [Accessed: 15 October 2023].
Seeliger, C., Schyschka, L., Kronbach, Z., Wottge, A., van Griensven, M., Wildemann, B. and Vester, H. 2015. Signaling pathway STAT1 is strongly activated by IFN-β in the pathogenesis of osteoporosis. *European Journal of Medical Research* 20(1), p. 1. Available at: /pmc/articles/PMC4300729/ [Accessed: 31 August 2023].

Seif, F., Khoshmirsafa, M., Aazami, H., Mohsenzadegan, M., Sedighi, G. and Bahar, M. 2017. The role of JAK-STAT signaling pathway and its regulators in the fate of T helper cells. *Cell Communication and Signaling 2017 15:1* 15(1), pp. 1–13. Available at: https://biosignaling.biomedcentral.com/articles/10.1186/s12964-017-0177-y [Accessed: 19 November 2023].
Senthelal, S., Li, J., Ardeshirzadeh, S. and Thomas, M.A. 2023. Arthritis. *StatPearls*, p. Available from: https://www.ncbi.nlm.nih.gov/books. Available at: https://www.ncbi.nlm.nih.gov/books/NBK518992/ [Accessed: 2 November 2023].
Sharp, J.T., Lidsky, M.D., Collins, L.C. and Moreland, J. 1971. Methods of scoring the progression of radiologic changes in rheumatoid arthritis. Correlation of radiologic, clinical and laboratory abnormalities. *Arthritis and rheumatism* 14(6), pp. 706–720. Available at: https://pubmed.ncbi.nlm.nih.gov/5135791/ [Accessed: 21 September 2023].
Sharp, R.C., Abdulrahim, M., Naser, E.S. and Naser, S.A. 2015. Genetic variations of PTPN2 and PTPN22: Role in the pathogenesis of Type 1 diabetes and Crohn's disease. *Frontiers in Cellular and Infection Microbiology* 5(DEC), p. 160103. doi: 10.3389/FCIMB.2015.00095/BIBTEX.
Shevach, E.M. 2008. Immunology. Regulating suppression. *Science (New York, N.Y.)* 322(5899), pp. 202–203. Available at: https://pubmed.ncbi.nlm.nih.gov/18845735/ [Accessed: 25 August 2022].
Sitt, J.C.M., Griffith, J.F. and Wong, P. 2016. Ultrasound-guided synovial biopsy. *The British Journal of Radiology* 89(1057). Available at: /pmc/articles/PMC4985949/ [Accessed: 14 October 2023].
Skoczyńska, M. and Swierkot, J. 2018. The role of diet in rheumatoid arthritis. *Reumatologia* 56(4), p. 259. Available at: /pmc/articles/PMC6142028/ [Accessed: 8 September 2023].
Smolen, J.S. et al. 2006. Predictors of joint damage in patients with early rheumatoid arthritis treated with high-dose methotrexate with or without concomitant infliximab: results from the ASPIRE trial. *Arthritis and rheumatism* 54(3), pp. 702–710. Available at: https://pubmed.ncbi.nlm.nih.gov/16508926/ [Accessed: 18 August 2022].
Smolen, J.S. et al. 2017. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Annals of the rheumatic diseases* 76(6), pp. 960–977. Available at: https://pubmed.ncbi.nlm.nih.gov/28264816/ [Accessed: 18 August 2022].
Smolen, J.S. et al. 2018. Rheumatoid arthritis. *Nature Reviews Disease Primers 2018 4:1* 4(1), pp. 1–23. Available at: https://www.nature.com/articles/nrdp20181 [Accessed: 1 August 2022].
Sokolova, M. V., Schett, G. and Steffen, U. 2021. Autoantibodies in Rheumatoid Arthritis: Historical Background and Novel Findings. *Clinical Reviews in Allergy & Immunology 2021 63:2* 63(2), pp. 138–151. Available at: https://link.springer.com/article/10.1007/s12016-021-08890-1 [Accessed: 12 August 2023].
Stahl, E.A. et al. 2010. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nature Genetics 2010 42:6* 42(6), pp. 508–514. Available at: https://www.nature.com/articles/ng.582 [Accessed: 15 October 2023].
Stanford, S.M. and Bottini, N. 2014. PTPN22: the archetypal non-HLA autoimmunity gene. *Nature Reviews Rheumatology 2014 10:10* 10(10), pp. 602–611. Available at: https://www.nature.com/articles/nrrheum.2014.109 [Accessed: 13 September 2023].
Stolt, P., Källberg, H., Lundberg, I., Sjögren, B., Klareskog, L. and Alfredsson, L. 2005. Silica exposure is associated with increased risk of developing rheumatoid arthritis: results from

the Swedish EIRA study. *Annals of the rheumatic diseases* 64(4), pp. 582–586. Available at: https://pubmed.ncbi.nlm.nih.gov/15319232/ [Accessed: 11 August 2022].
Su, J. et al. 2009. Discoidin domain receptor 2 is associated with the increased expression of matrix metalloproteinase-13 in synovial fibroblasts of rheumatoid arthritis. *Molecular and cellular biochemistry* 330(1–2), pp. 141–152. Available at: https://pubmed.ncbi.nlm.nih.gov/19415460/ [Accessed: 2 September 2023].
Sugiyama, D., Nishimura, K., Tamaki, K., Tsuji, G., Nakazawa, T., Morinobu, A. and Kumagai, S. 2010. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. *Annals of the rheumatic diseases* 69(1), pp. 70–81. Available at: https://pubmed.ncbi.nlm.nih.gov/19174392/ [Accessed: 11 August 2022].
Svensson, M.N.D. et al. 2019. Reduced expression of phosphatase PTPN2 promotes pathogenic conversion of Tregs in autoimmunity. *The Journal of Clinical Investigation* 129(3), pp. 1193–1210. Available at: https://doi.org/10.1172/JCI123267DS1 [Accessed: 19 November 2023].
Tahir, H. et al. 2017. Secukinumab in Active Rheumatoid Arthritis after Anti-TNFα Therapy: A Randomized, Double-Blind Placebo-Controlled Phase 3 Study. *Rheumatology and Therapy* 4(2), p. 475. Available at: /pmc/articles/PMC5696298/ [Accessed: 15 October 2023].
Takeuchi, T., Yoshida, H. and Tanaka, S. 2021. Role of interleukin-6 in bone destruction and bone repair in rheumatoid arthritis. *Autoimmunity Reviews* 20(9), p. 102884. doi: 10.1016/J.AUTREV.2021.102884.
Tanaka, Y. 2021. Recent progress in treatments of rheumatoid arthritis: an overview of developments in biologics and small molecules, and remaining unmet needs. *Rheumatology* 60(Supplement_6), pp. vi12–vi20. Available at: https://dx.doi.org/10.1093/rheumatology/keab609 [Accessed: 16 September 2023].
Taylor, P.C. et al. 2023. Anti-GM-CSF otilimab versus sarilumab or placebo in patients with rheumatoid arthritis and inadequate response to targeted therapies: a phase III randomised trial (contRAst 3). *Annals of the Rheumatic Diseases* 0, p. ard-2023-224449. Available at: https://ard.bmj.com/content/early/2023/09/11/ard-2023-224449 [Accessed: 15 October 2023].
Theeuwes, W.F. et al. 2023. CD64 as novel molecular imaging marker for the characterization of synovitis in rheumatoid arthritis. *Arthritis Research and Therapy* 25(1), pp. 1–15. Available at: https://arthritis-research.biomedcentral.com/articles/10.1186/s13075-023-03147-y [Accessed: 28 November 2023].
Thurman, R.E. et al. 2012. The accessible chromatin landscape of the human genome. *Nature 2012 489:7414* 489(7414), pp. 75–82. Available at: https://www.nature.com/articles/nature11232 [Accessed: 8 May 2022].
Tlustochowicz, W., Rahman, P., Seriolo, B., Krammer, G., Porter, B., Widmer, A. and Richards, H.B. 2016. Efficacy and Safety of Subcutaneous and Intravenous Loading Dose Regimens of Secukinumab in Patients with Active Rheumatoid Arthritis: Results from a Randomized Phase II Study. *The Journal of rheumatology* 43(3), pp. 495–503. Available at: https://pubmed.ncbi.nlm.nih.gov/26834211/ [Accessed: 15 October 2023].
Tobón, G.J., Youinou, P. and Saraux, A. 2010. The environment, geo-epidemiology, and autoimmune disease: Rheumatoid arthritis. *Journal of autoimmunity* 35(1), pp. 10–14. Available at: https://pubmed.ncbi.nlm.nih.gov/20080387/ [Accessed: 1 August 2022].
Trynka, G., Sandor, C., Han, B., Xu, H., Stranger, B.E., Liu, X.S. and Raychaudhuri, S. 2013. Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nature*

genetics 45(2), pp. 124–130. Available at: https://pubmed.ncbi.nlm.nih.gov/23263488/ [Accessed: 16 August 2022].
Tsompana, M. and Buck, M.J. 2014. Chromatin accessibility: a window into the genome. *Epigenetics & Chromatin 2014 7:1* 7(1), pp. 1–16. Available at: https://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/1756-8935-7-33 [Accessed: 8 September 2022].
Twohig, J.P. et al. 2019. Activation of naïve CD4 + T cells re-tunes STAT1 signaling to deliver unique cytokine responses in memory CD4 + T cells. *Nature Immunology* 20(4), pp. 458–470. Available at: https://www.nature.com/articles/s41590-019-0350-0 [Accessed: 25 April 2021].
Verschueren, P. et al. 2015. Methotrexate in combination with other DMARDs is not superior to methotrexate alone for remission induction with moderate-to-high-dose glucocorticoid bridging in early rheumatoid arthritis after 16 weeks of treatment: the CareRA trial. *Annals of the rheumatic diseases* 74(1), pp. 27–34. Available at: https://pubmed.ncbi.nlm.nih.gov/25359382/ [Accessed: 18 August 2022].
Viatte, S. et al. 2016. Association Between Genetic Variation in FOXO3 and Reductions in Inflammation and Disease Activity in Inflammatory Polyarthritis. *Arthritis & rheumatology (Hoboken, N.J.)* 68(11), pp. 2629–2636. Available at: https://pubmed.ncbi.nlm.nih.gov/27214848/ [Accessed: 16 August 2022].
Villarino, A. V., Kanno, Y. and O'Shea, J.J. 2017. Mechanisms and consequences of Jak–STAT signaling in the immune system. *Nature Immunology 2017 18:4* 18(4), pp. 374–384. Available at: https://www.nature.com/articles/ni.3691 [Accessed: 15 July 2023].
Vital, E.M. et al. 2010. Management of nonresponse to rituximab in rheumatoid arthritis: predictors and outcome of re-treatment. *Arthritis and rheumatism* 62(5), pp. 1273–1279. Available at: https://pubmed.ncbi.nlm.nih.gov/20131284/ [Accessed: 24 August 2022].
Vos, K., Thurlings, R.M., Wijbrandts, C.A., van Schaardenburg, D., Gerlag, D.M. and Tak, P.P. 2007. Early effects of rituximab on the synovial cell infiltrate in patients with rheumatoid arthritis. *Arthritis and rheumatism* 56(3), pp. 772–778. Available at: https://pubmed.ncbi.nlm.nih.gov/17328049/ [Accessed: 24 August 2022].
WHO EML 23rd List (2023). 2023. Available at: http://apps.who.int/bookorders. [Accessed: 7 October 2023].
Wicks, I.P. and Roberts, A.W. 2015. Targeting GM-CSF in inflammatory diseases. *Nature Reviews Rheumatology 2015 12:1* 12(1), pp. 37–48. Available at: https://www.nature.com/articles/nrrheum.2015.161 [Accessed: 15 October 2023].
Wiede, F. et al. 2019. T-cell–specific PTPN2 deficiency in NOD mice accelerates the development of type 1 diabetes and autoimmune comorbidities. *Diabetes* 68(6), pp. 1251–1266. Available at: https://research.monash.edu/en/publications/t-cellspecific-ptpn2-deficiency-in-nod-mice-accelerates-the-devel [Accessed: 19 November 2023].
Xie, S., Li, S., Tian, J. and Li, F. 2020. Iguratimod as a New Drug for Rheumatoid Arthritis: Current Landscape. *Frontiers in Pharmacology* 11. Available at: /pmc/articles/PMC7054862/ [Accessed: 14 October 2023].
Xu, F., Cong, P., Lu, Z., Shi, L., Xiong, L. and Zhao, G. 2023. Integration of ATAC-Seq and RNA-Seq identifies key genes and pathways involved in the neuroprotection of S-adenosylmethionine against perioperative neurocognitive disorder. *Computational and Structural Biotechnology Journal* 21, pp. 1942–1954. doi: 10.1016/J.CSBJ.2023.03.001.

Xu, Y. et al. 2015. Regulatory Effect of Iguratimod on the Balance of Th Subsets and Inhibition of Inflammatory Cytokines in Patients with Rheumatoid Arthritis. *Mediators of Inflammation* 2015. Available at: /pmc/articles/PMC4680115/ [Accessed: 14 October 2023].
Yao, Z., Chen, Y., Cao, W. and Shyh-Chang, N. 2020. Chromatin-modifying drugs and metabolites in cell fate control. *Cell Proliferation* 53(11), p. e12898. Available at: https://onlinelibrary.wiley.com/doi/full/10.1111/cpr.12898 [Accessed: 2 November 2023].
Ye, Y. et al. 2019. Iguratimod represses B cell terminal differentiation linked with the inhibition of PKC/EGR1 axis. *Arthritis Research & Therapy* 21(1). Available at: /pmc/articles/PMC6458835/ [Accessed: 14 October 2023].
Yin, H., Liu, N., Sigdel, K.R. and Duan, L. 2022. Role of NLRP3 Inflammasome in Rheumatoid Arthritis. *Frontiers in Immunology* 13, p. 931690. doi: 10.3389/FIMMU.2022.931690/BIBTEX.
Yoshida, H. et al. 2001. WSX-1 is required for the initiation of Th1 responses and resistance to L. major infection. *Immunity* 15(4), pp. 569–578. Available at: https://pubmed.ncbi.nlm.nih.gov/11672539/ [Accessed: 10 July 2022].
Yoshitake, F., Itoh, S., Narita, H., Ishihara, K. and Ebisu, S. 2008. Interleukin-6 Directly Inhibits Osteoclast Differentiation by Suppressing Receptor Activator of NF-κB Signaling Pathways *. *Journal of Biological Chemistry* 283(17), pp. 11535–11540. Available at: http://www.jbc.org/article/S0021925820619482/fulltext [Accessed: 31 August 2023].
Yuan, F.L., Li, X., Lu, W.G., Sun, J.M., Jiang, D.L. and Xu, R.S. 2013. Epidermal growth factor receptor (EGFR) as a therapeutic target in rheumatoid arthritis. *Clinical rheumatology* 32(3), pp. 289–292. Available at: https://pubmed.ncbi.nlm.nih.gov/23179003/ [Accessed: 2 September 2023].
Yuan, J. et al. 2023. Integration of ATAC-Seq and RNA-Seq identifies the key genes in myocardial ischemia. *Genes & Diseases* 10(1), p. 62. Available at: /pmc/articles/PMC10066257/ [Accessed: 10 October 2023].
Zaware, N. and Zhou, M.M. 2019. Bromodomain biology and drug discovery. *Nature Structural & Molecular Biology 2019 26:10* 26(10), pp. 870–879. Available at: https://www.nature.com/articles/s41594-019-0309-8 [Accessed: 2 November 2023].
Zhang, F. et al. 2019. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nature Immunology 2019 20:7* 20(7), pp. 928–942. Available at: https://www.nature.com/articles/s41590-019-0378-1 [Accessed: 26 November 2023].
Zhang, Z. et al. 2022. Recent progress in DNA methyltransferase inhibitors as anticancer agents. *Frontiers in Pharmacology* 13. Available at: /pmc/articles/PMC10107375/ [Accessed: 9 December 2023].
Zhao, J. et al. 2022. DNA Methylation of T Lymphocytes as a Therapeutic Target: Implications for Rheumatoid Arthritis Etiology. *Frontiers in Immunology* 13. Available at: /pmc/articles/PMC8927780/ [Accessed: 10 August 2023].

Appendix

8.1 Appendix

8.1.1 Code for ATAC-seq normalisation

```
trim_galore --cores 4 -o ${sampleID} --paired
${sampleID}/${sampleID}_1.fastq. gz
${sampleID}/${sampleID}_2.fastq.gz

fastqc -o ${sampleID} ${sampleID}/${sampleID}_1.fastq.gz
fastqc -o ${sampleID} ${sampleID}/${sampleID}_2.fastq.gz
fastqc -o ${sampleID}
${sampleID}/${sampleID}.trimmed_1.fastq.gz fastqc -o
${sampleID} ${sampleID}/${sampleID}.trimmed_2.fastq.gz

multiqc ${params.projectDir}/${params.outputDir}/* -n multiQC

bwa mem -t 4 ${genome}
${sampleID}/${sampleID}.trimmed_1.fastq.gz ${sampleID}/
${sampleID}.trimmed_2.fastq.gz | samtools view -bS - |
samtools sort - -o ${sample ID}/${sampleID}.bam samtools index
${sampleID}/${sampleID}.bam

sort -k1,1 -k2,2n ${blacklist} | bedtools merge -d 1000 >
${sampleID}/${params.blacklistName}.sorted.bed bedtools
intersect -a ${sampleID}/${sampleID}.bam -b $
{sampleID}/${params.blacklistName}.sorted.bed -v > ${sampl
eID}/${sampleID}.blacklist.bam

samtools view -h ${sampleID}/${sampleID}.bam | python3
${params.projectDir} /src/removeChrom.py - -
${params.mtChromsomeName} | samtools view -bh - | samtools
sort - -o ${sampleID}//${sampleID}.noMT.bam samtools index
${sampleID}/${sampleID}.noMT.bam

samtools view -h ${sampleID}/${sampleID}.bam | python3
${params.projectDir} /src/removeChrom.py - -
${params.mtChromsomeName} | samtools view -bh - | samtools
sort - -o ${sampleID}//${sampleID}.noMT.bam samtools index
${sampleID}/${sampleID}.noMT.bam

dups <- readsDupFreq(bam, index=bai) complexity <-
estimateLibComplexity(dups, times=100,
interpolate.sample.sizes=seq(0.1, 1, by=0.01))

samtools view -h -b -s \$a ${sampleID}/${sampleID}.proper.bam
> ${sampleID}/${sampleID}.subsample.bam

java -jar -Xmx40G ${params.picardExecutable} MarkDuplicates
I=${sampleID}/${sampleID}.subsample.bam O=${sampl
eID}/${sampleID}.markdup.bam
```

```
M=${sampleID}/${sampleID}.markdup.txt REMOVE_DUPLICATES=false
VALIDATION_STRINGENCY= SILENT java -jar -Xmx40G
${params.picardExecutable} MarkDuplicates
I=${sampleID}/${sampleID}.subsample.bam O=${sampl
eID}/${sampleID}.rmdup.bam M=${sampleID}/${sampleID}.rmdup.txt
REMOVE_DUPLICATES=true VALIDATION_STRINGENCY=SILEN T samtools
index ${sampleID}/${sampleID}.markdup.bam samtools index
${sampleID}/${sampleID}.rmdup.bam

samtools sort -n ${sampleID}/${sampleID}.rmdup.bam | samtools
fixmate - - | samtools view -bf 0x2 - | bedtools bamtobed -i
stdin -bedpe > ${sampleID}/${sample ID}.rmdup.bedpe

bash ${params.projectDir}/src/bedpeTn5shift.sh
${sampleID}/${sampleID}.rmdup.bedpe >
${sampleID}/${sampleID}.rmdup.tn5.bedpe

bash ${params.projectDir}/src/bedpeMinimalConvert.sh
${sampleID}/${sampleID}.rmdup.tn5.bedpe >
${sampleID}/${sampleID}.rmdup.tn5.minimal.bedpe

$string .= "macs2 callpeak -t ".$sampleID."/".$sampleID.".rm
dup.tn5.minimal.bedpe -c
".$wdirs[0]."/".$control.".rmdup.tn5.minimal.bedpe -f BAMPE -g
".$genomeSize." --outdir ".$sampleID." -n
".$sampleID.".".$dash_pvalue." --broad --broad-cutoff
".$pvalue." --keep-dup all";
```

8.1.2 Code for ChIP-seq normalisation

```
trim_galore --cores 4 -o ${sampleID} --paired
${sampleID}/${sampleID}_1.fastq.gz
${sampleID}/${sampleID}_2.fastq.gz

fastqc -o ${sampleID} ${sampleID}/${sampleID}_1.fastq.gz
fastqc -o ${sampleID} ${sampleID}/${sampleID}_2.fastq.gz

bwa index -a bwtsw ${genome}
bwa mem -t 4 ${genome}
${sampleID}/${sampleID}.trimmed_1.fastq.gz
${sampleID}/${sampleID}.trimmed_2.fastq.gz | samtools view -bS
- | samtools sort - -o ${sampleID}/${samp
leID}.bam

java -jar -Xmx40G ${params.picardExecutable} MarkDuplicates
I=${sampleID}/${sampleI
D}.bam O=${sampleID}/${sampleID}.markdup.bam
M=${sampleID}/${sampleID}.markdup.txt REMO
VE_DUPLICATES=false VALIDATION_STRINGENCY=SILENT

bamtools stats -in ${sampleID}/${sampleID}.markdup.bam >
${sampleID}/${sampleID}.markdup.stats.txt

macs2 callpeak -t sampleID.rmdup.bam -c control.rmdup.bam -f
BAMPE -g ".$genomeSize." --outdir ".$sampleID." -n ".$sampleID
-B -q ".$pvalue

qvalue = 0.05

bedtools getfasta -bed
".$sampleID."/".$sampleID.".$pvalue."_peak
s.narrowPeak -fi ".$genome." -fo
".$sampleID."/".$sampleID.".$pvalue.".peaks.fasta

#TSS
bedtools intersect -wa -wb -a $sampleID/$sampleFile -b
$tssFile > $sampleID/tss.$pvalue.txt

#intron
bedtools intersect -wa -wb -a $sampleID/$sampleFile -b
$intronFile > $sampleID/intron.
$pvalue.txt

#exon
bedtools intersect -wa -wb -a $sampleID/$sampleFile -b
$exonFile > $sampleID/exon.$pva
lue.txt
```

```
#stop
bedtools intersect -wa -wb -a $sampleID/$sampleFile -b
$stopFile > $sampleID/stop.$pva
lue.txt
```