



Functional genomics of psoriasis

Alicia Lledo Lara
Hertford College
University of Oxford

*A thesis submitted in partial
fulfilment of the requirements for the degree of
Doctor of Philosophy
Trinity Term, 2018*

Abstract

Functional genomics of psoriasis

Alicia Lledo Lara, Hertford College, Trinity Term 2018

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy of the University of Oxford

This is my abstract...

Acknowledgements

Thank you, thank you, thank you.

Declarations

I declare that unless otherwise stated, all work presented in this thesis is my own. Several aspects of each project relied upon collaboration where part of the work was conducted by others.

Submitted Abstracts

Title	Year
Authors	

Associated Publications

Title

Journal

Authors

Other Publications

Title

Journal

Authors

Contents

Abstract	i
Acknowledgements	ii
Declarations	iii
Submitted Abstracts	iv
Associated Publications	v
Contents	vi
List of Figures	ix
List of Tables	x
Abbreviations	xi
1 Introduction	1
1.1 Psoriasis and psoriatic arthritis	1
1.1.1 Epidemiology and global impact	1
1.1.2 Psoriasis and inflammatory dermatoses	2
1.1.3 PsA and spondyloarthropathies	3
1.2 Pathophysiology of psoriasis and psoriatic arthritis	4
1.2.1 Clinical presentation and diagnosis	4
1.2.2 Aetiology of psoriasis and PsA	8
1.2.3 Cell types involved in psoriasis and PsA pathogenesis	11
1.2.4 Therapeutic intervention and prognosis	17
1.3 Genetics of psoriasis and psoriatic arthritis	18
1.3.1 Heritability	18
1.3.2 Non-GWAS and linkage studies	19
1.3.3 Genome-wide association studies	20
1.3.4 Relevance of non-coding versus coding variants in disease susceptibility	23
1.3.5 The role of GWAS studies in highlighting immune-relevant cell types and pathways	25
1.3.6 Limitations and future of GWAS studies	34
1.4 Functional interpretation of genome-wide association studies in complex diseases	36
1.4.1 Overcoming the limitations of GWAS: post-GWAS studies	36
1.4.2 The use of fine-mapping to prioritise causal variants	37
1.4.3 Understanding the epigenetic landscape in complex diseases	39
1.4.4 The chromatin landscape	42

CONTENTS

1.4.5	Transcriptional profiles in disease	49
1.4.6	Transcriptional regulation in complex diseases	51
1.4.7	Integration and interpretation of data	53
1.4.8	Approaches to establish disease mechanisms and causality of genetic variant	53
2	Material and Methods	54
2.1	Ethical approval and recruitment of study participants	54
2.1.1	Psoriasis patient recruitment	54
2.1.2	PsA patient recruitment	55
2.1.3	Healthy volunteer recruitment	56
2.2	Sample processing	56
2.2.1	PBMC and synovial fluid cells isolation	56
2.2.2	Primary cell isolation using magnetic-activated cell sorting	57
2.2.3	Primary cell isolation using fluorescence-activated cell sorting	57
2.2.4	Skin biopsies processing and adherent assay	58
2.3	Experimental protocols	59
2.3.1	Cryopreservation and cell culture	59
2.3.2	ATAC-seq, Fast-ATAC and Omni-ATAC	60
2.3.3	Chromatin Immunoprecipitation with sequencing library preparation by Tn5 transposase	64
2.3.4	RNA extraction and RNA-seq	65
2.3.5	Single-cell analysis of the V(D)J T cell receptor repertoire .	68
2.3.6	DNA extraction and SNP genotyping	69
2.3.7	Mass cytometry analysis	69
2.4	Computational and statistical analysis	70
2.4.1	ATAC-seq data analysis	70
2.4.2	ChIPm data analysis	73
2.4.3	RNA-seq data analysis	74
2.4.4	Genomic region annotation, pathway enrichment analysis and visualisation	74
2.4.5	Enrichment analysis for genomic annotation features . . .	75
2.4.6	Statistical fine-mapping	75
3	Establishment of laboratory methods and analytical tools to assess genome-wide chromatin accessibility in clinical samples	76
3.1	Introduction	76
3.2	Results	77
3.2.1	Establishment of an ATAC-seq data analysis pipeline based on current knowledge	77
3.2.2	Assessment of ATAC-seq transposition times and comparison with FAST-ATAC protocol in relevant cell types	89
3.2.3	Limitations of ATAC-seq and FAST-ATAC to assess chromatin accessibility in KC	89
3.2.4	Discussion	93

CONTENTS

4 Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis	96
4.1 Introduction	96
4.2 Results	96
4.2.1 PsA patients cohort description and datasets	96
4.2.2 Differences in the chromatin accessibility landscape between circulating and SF immune cells	99
4.2.3 Differential gene expression analysis in paired circulating and synovial immune cells	106
4.3 Discussion	106
Appendices	107
A Establishment of methods to assess genome-wide chromatin accessibility	108
Bibliography	125

List of Figures

1.1	Crosstalk between innate and adaptive immunity in psoriasis	16
3.1	Measurements for quality control assessment in ATAC-seq samples	82
3.2	Peak calling and sequencing depth in ATAC-seq samples	85
3.3	Peak calling filtering using IDR analysis in ATAC-seq samples . .	86
3.4	Differential chromatin accessibility analysis for different background reads cut-offs.	88
3.5	Exploration of the differential chromatin accessibility analysis using 80% as the empirical cut-off.	89
3.6	Assessment of the effect of transposition times on the ATAC-seq QC parameters	90
3.7	Differences in MT DNA abundance and signal specificity between ATAC-seq and FAST-ATAC protocols	91
3.8	QC assessment of ATAC-seq in KC enriched cell suspension derived from a psoriatic lesional skin biopsy	92
3.9	QC assessment of FAST-ATAC and Omni-ATAC in cultured NHEK	94
3.10	QC assessment of Omni-ATAC in NHEK and chromatin accessibility signal for the samples generated with the different ATAC-seq protocols	95
4.1	QC of FAST-ATAC PsA samples in four cell types	101
4.2	Combined PCA analysis of all four cell types isolated from blood and SF.	102
4.3	Enrichment of eQTLs in the combined cell types PsA accessible chromatin master list.	103
4.4	Annotation with genomic regions and chromatin states of the PsA DOCs from the four cell types differential analysis.	105
A.1	FAST-ATAC and Omni-ATAC NHEK tapestation profiles.	109
A.2	Assessment of TSS enrichment from ATAC-seq and FAST-ATAC in healthy and psoriasis skin biopsies samples.	110

List of Tables

1.1	Variables and scoring used in the Psoriasis Area and Severity Index (PASI)	6
1.2	CASPARI criteria for diagnosis of PsA	7
2.1	Antibody panel used for FACS separation of primary cell populations in controls and PsA samples	58
2.2	Modified Illumina Nextera indexing primers	62
2.3	ChIPm buffers modified from Schmidl <i>et. al.</i> , 2015	66
2.4	Antibody panel used for immunoprecipitation of histone marks in ChIPm	67
2.5	Molecules targeted by the phenotyping and cytokine production antibody staining panel in PBMCs and SFMCs.	70
3.1	Summary table of ATAC-seq methodology analysis for peak calling, filtering and differential analysis.	78
3.2	ATAC-seq percentage of MT reads and fraction of reads in called peaks	83
3.3	Description of the most relevant parameter from the ATAC-seq and FAST-ATAC protocols assayed in NHEK and skin biopsies.	93
4.1	Description of PsA patients cohort recruitment and metadata.	98
4.2	Datasets generated for the PsA cohort samples	99
4.3	Summary results of the chromatin accessibility analysis between SF and PB in PsA samples	104
4.4	Summary results of the chromatin accessibility analysis between SF and PB in PsA samples	105

Abbreviations

Abbreviation	Definition
Ab	Antibody
ATAC-seq	
Atopic dermatitis	AD
ChIPm	
CLE	cutaneous lupus erythematosus
DMARDs	disease-modifying antirheumatic drugs
Fast-ATAC	
IDR	
GWAS	Genome-wide association studies
KC	Keratinocytes
NSAID	nonsteroidal antiinflammatory drug
Omni-ATAC	
PCA	
PI	Protein inhibitor
PsA	
QC	
qPCR	quantitative polymerase chain reaction
RA	Rheumatoid arthritis
SDS	Sodium dodecyl sulfate
SF	Synovial fluid

Chapter 1

Introduction

1.1 Psoriasis and psoriatic arthritis

Psoriasis and psoriatic arthritis (PsA) have been progressively identified as two different common complex disease entities. Psoriasis is a chronic inflammatory dermatose disease with episodes of relapse and remittance (Nestle et al. 2009). On the other hand, PsA is a seronegative chronic inflammatory disease within the family of spondyloarthritis (Moll et al. 1973; Coates et al. 2016) that usually develops after the psoriasis skin manifestations(Villanova2016). Psoriasis and PsA have shared and distinct clinical features, which are likely a reflection of the commonalities and differences in genetic loci contributing to disease development. It is important to understand those commonalities and differences at the physiological and genetic level in order to better understand the relevance of the genetic variability in the risk to develop psoriasis and PsA.

1.1.1 Epidemiology and global impact

Psoriasis represents a serious global health problem that currently affects about 100 million people worldwide, including children and adults with no sex bias (Organization 2016). Although there is a very weak correlation with geographic latitude (Jacobson et al. 2011), it has been reported to vary upon ethnicity. For example, psoriasis prevalence in adults is lower among African,

Introduction

African American and Asian (0.4-0.7%) compared to American and Canadian (4.6 and 4.7%, respectively) populations. In the UK, psoriasis prevalence ranges between 2-3% and it affects approximately 1.8 million people (Perera et al. 2012).

PsA prevalence in the general population ranges between 0.04-1.2% (Perera et al. 2012) but it dramatically increases to 10-30% within psoriasis cases (Reich2008; Gelfand et al. 2005) and evidences the association between the two diseases. Particularly, in the UK, 14% of the psoriasis patients develop chronic inflammatory arthritis in the form of PsA at some point of the disease course (Ibrahim et al. 2009).

Although psoriasis can be developed at any age, onset of disease seems to have a bimodal distribution strongly influenced by the Human Leukocyte Antigen (HLA) Cw*06:02 (HLA-Cw6:02), an allele for one of the genes in the Major Histocompatibility Complex (MHC) I, involved in antigen presentation (Henseler and Christophers 1985) and the strongest genetic association with psoriasis and PsA risk ((Ellinghaus2010, Strange2010, Stuart2010; Sun2010). The early-onset or Type I is characterised by development of disease around 16-22 and 30-39 years and a prevalence for HLA-C*06:02 (85.4% of the cases). In contrasts, the late-onset or Type II group manifests disease between 50-60 years old and presents positive HLA-C*06:02 only in 14.6% of the cases.

Psoriasis and PsA also represent an economical burden for the countries' economies due to treatment and associated morbidity. For example, in the UK treatment and management of psoriasis in 2015 ranged between 4,000 to 14,000, before and after requirements of biological therapy, respectively (Burgos-Pol and Dermo 2016) and the costs are even greater for PsA (Poole et al. 2010).

1.1.2 Psoriasis and inflammatory dermatoses

The group of inflammatory dermatoses affects up to 70% of the population, regardless age and geographic location (**ICD-10**), and it represents the 4th leading

Introduction

cause of nonfatal burden (**Roderick2014**). The skin is the biggest organ in the human body constituting an effective barrier between the environment and the internal organs. The most external layer, the epidermis, plays a relevant role in the innate and adaptive immunity (Proksch et al. 2008) and its alterations due to exogenous or endogenous factors can lead to development of inflammatory dermatose conditions, such as psoriasis, atopic dermatitis (AD) or cutaneous lupus erythematosus (CLE) (**Johnson-Huang; 2009**). Lesions in psoriasis can be non-pustular and pustular which reflects the heterogeneity in the type, location and severity of the disease and impairs the clinical classification (Perera et al. 2012). As a result, several phenotypes of psoriasis including vulgaris, guttate, pustular, erythroderma and nail pitting have been defined and it is under debate whether some of those should be considered a different disease entity (Marrakchi et al. 2011).

1.1.3 PsA and spondyloarthropaties

PsA belong to the family known as spondylarthropaties (SpA) which also includes other subtypes such as ankylosing spondylitis (AS), reactive arthritis (ReA), idiopathic inflammatory bowel disease (IBD) and undifferentiated SpA (Baeten et al. 2013). All SpA subtypes are characterised by structural damage (bone formation and erosion) as well as inflammation of joints and extraarticular sites such as eyes, gut and skin. Additional SpA criteria have led to a reduced classification of SpA into axial and peripheral SpA based on the affected joint (spine/sacroiliac or peripheral) and the presence of extraarticular features (**Runwaleit2001; Runwaleit2001**). Studies in human families and rat models with HLA-B27 positive status have shown manifestation of different SpA forms, such as psoriasis and IBD, within a single family or individual (**Said-Nahal2000 “parencite ; Hammer et al. 1990**). These observations support the hypothesis that SpA subtypes may be a single multifaceted condition

Introduction

with shared genetic, immunopathological and structural features and dynamic phenotypes (Baeten et al. 2013). Conversely, some studies suggest that multiple genetic factors may be involved in the determination of the axial and peripheral arthritis and partially explain the immunopathological differences between the two (Porcher et al. 2005; Appel et al. 2011; Noordenbos et al. 2012).

As a phenotype, PsA can be further subdivided in five clinical groups based on Moll and Wright criteria: distal, destructive, symmetric, asymmetric and spinal (Moll et al. 1973). These subclasses mainly differed upon the location, number and distribution of the affected joints. Later studies have questioned this method of classification due overlapping of the different subsets and lack of inclusion of dactylitis (diffuse swelling of a digit) a distinctive feature of PsA (Reich et al. 2009). This phenotypic heterogeneity increases the difficulty in the design and achievement of meaningful outcomes from clinical studies.

1.2 Pathophysiology of psoriasis and psoriatic arthritis

1.2.1 Clinical presentation and diagnosis

Approximately 90% of all psoriasis cases are plaque psoriasis vulgaris that manifests with raising well demarcated plaques, erythema and scaling. The thickening (acanthosis) and vascularisation of the epidermis leads to the plaques formation (Perera et al. 2012) that can vary in size and distribution, being the most common the elbows, knees and scalp (Griffiths and Lancet 2007). The second most common type is psoriasis guttate (10% of all cases) characterised by acute onset of small droplike papules usually in the trunk and proximal extremities (Vence et al. 2015). Type I psoriasis commonly appears in the form of guttate lesions after bacterial infection whilst type II involves spontaneous chronic plaques (Perera et al. 2012).

Introduction

In PsA the most common manifestation is the symmetric/polyarticular (more than 50%) followed by the asymmetric/oligoarticular (around 30%) PsA, that affects single or few distal interphalangeal or phalangeal joints (Reich et al. 2009; McGonagle et al. 2011). The psoriatic lesions precede joint inflammation in approximately 60-70% of the cases(McGonagle; 2011; Gladman et al. 2005). Particularly, nail, scalp and intergluteal lesions constitute a predictive biomarker for development of joint inflammation (Moll1976; McGonagle; 2011; Griffiths and Lancet 2007). This reinforces the need of appropriate coordination between dermatologists and rheumatologists for an early diagnostic and treatment that could prevent functional joint disability.

Several comorbidities have been associated with psoriasis and PsA, with comparatively greater prevalence in PsA. For example, intraocular inflammation known as uveitis affects 8% of PsA patients compared to 2% of the psoriasis ones (Husted et al. 2011; Oliveira et al. 2015). Other comorbidities include inflammatory bowel disease(IBD), cardiovascular disease (CVD) (Gelfand et al. 2006), type II diabetes (T2D) (**Sapiro2007**) and metabolic syndrome (**Cohrn20017**).

The diagnosis of psoriasis and PsA is mainly based in clinical assessment since there is a lack of appropriate biomarkers at early stages of disease (Villanova et al. 2013). Evaluating the severity of psoriasis skin lesions remains challenging and different measures have been implemented. The Psoriasis Area and Severity Index (PASI) (Fredriksson and Dermatology 1978) is the most widely used in research and drug trials (Finlay 2005). This test quantifies lesional burden weighted by body part based on the amount of affected body surface area and the degree of severity of erythema, induration and scale (Table 1.1). Disease is considered mild for PASI \leq 7 and it is classified as moderate-to-severe for PASI \geq 7-12, depending on the study (**Schmitt2005; add ref from cell types**; Finlay 2005).

Introduction

To evaluate PsA, analysis of performance of the previously mentioned Moll and Wright criteria together with additional ones led to the configuration of the Classification Criteria for Psoriatic Arthritis (CASPAR) (**Taylor2006**), the most widely used. It requires the patient displaying inflammatory arthritis, enthesitis, and/or spondylitis and three points from a list of associated elements (Table 1.2). Another composite measure commonly used to evaluate treatment efficacy for PsA is the PsA Response Criteria (PsARC) based on the number of tender joints (TJC) and swollen joints (SJC) over 68 and 66, respectively, as well as a physician global assessment based on a short questionnaire (**Philipp2011; Clegg et al. 1996**).

PASI	description
Body location	Head and neck, upper limbs, trunk and lower limbs
Feature	Redness, thickness and scaling
Severity scale	Absent, mild, moderate, severe or very severe
Affected area (%)	0, 1-9, 10-29, 30-49, 50-69, 70-89 or 90-100

Table 1.1: For each of the four body locations the test quantifies the percentage of affected area and the severity of three intensity features: redness, thickness and scaling.

Introduction

CASPAR:	a patient with	must have three points	inflammatory articular points	disease from	(joint, spine, or five	or enthesial categories
					a. Current skin or scalp disease b. History of psoriasis c. Family history of psoriasis	
					Typical psoriatic nail dystrophy	
					a. Using preferrably by enzyme-linked immunosorbent assay (EMSA) b. Swelling of an entire finger c. History of dactylitis	
					Ossification near joint margins	
						Table 1.2: xxxx

Introduction

1.2.2 Aetiology of psoriasis and PsA

Psoriasis and PsA are complex chronic inflammatory diseases where a dysregulated immune response initiates as result of genetic predisposition and exposure to a particular environmental trigger (Figure??). One of the greater controversies has been characterising the origin of the pathologies as well as the connection between skin and joint inflammation. Particularly, for psoriasis it remains unclear whether disruption of the skin triggers activation of the immune response or viceversa.

Histopathological alterations in skin and joints

The epidermis is the most external structure of the skin and it is formed by approximately 90% keratinocytes (KC) organised in a layer structure that self-renews in a time dependent manner from the bottom to the surface (Wikramanayake et al. 2014). As the KC differentiate they undergo changes in morphology, replication ability and keratin composition of their intracellular matrix. In the context of psoriasis impaired epidermis cell renewal leads histological alterations and development of the psoriatic lesions. KC undergo upregulation in the proliferation rate (hyperplasia) that causes aberrant cell differentiation (parakeratosis) (ref) thickening of the epidermis and the subsequent scale formation (ref). Concomitantly, inflammation causes immune cell infiltration and hypervascularisation of the lesion driven by upregulation in the expression of angiogenic factors and activation of the endothelium (Perera et al. 2012).

In PsA, joint affection usually follows skin lesions and it involves a wide range of histological changes in the joints, particularly bone remodeling (Haddad and Chandran 2013). One of the most common structural changes is the arthritis caused by the swelling and inflammation of the joints (**Schett2011**). As result of this inflammation, alterations in bone remodeling leads to osteolysis with

Introduction

subsequent bone resorption and erosion at the affected joints (**Mensah2017**). This phenomenon is particularly relevant in arthritis mutilans or chronic absorptive arthritis, one of the most severe forms of PsA (Haddad and Chandran 2013). Bone erosion is also the main histopathological process driving dactylitis, where bone lysis resolves in shortening of the digits (Gladman et al. 2005). On the other hand, 35% of the PsA patients undergo inflammation of the connective tissue at the insertion of tendons or ligaments, phenomenon known as enthesitis (McGonagle et al. 2011; Polachek et al. 2017). Overtime, this causes debilitating structural changes due to formation of bony spurs along the insertion sites(**Schett2011**).

Dysregulation of the innate and adaptive immune response

The dysregulated immune response in psoriasis and PSA is the result of the interaction between innate and adaptive immune cells (ref section) resulting in feedback loops involving a complex cytokine milieu. Among the most relevant cytokines of the innate immunity involved in disease initiation are IFN- α and IFN- γ (**Leanne2009**). They are mainly produced by circulating plasmacytoid DC (pDC) and myeloid DC (mDC), respectively, upon activation by KC pro-inflammatory cytokines (Perera et al. 2012). Both are upregulated at the mRNA level in the lesional skin and contribute to lymphocyte recruitment and maintenance of DC activation (**Schmid1994**).

Another key cytokine in this dysregulated inflammatory response is TNF- α which has a prominent role in bone turnover and bone remodeling in PsA (Mensah et al. 2008). It is produced by activated KC, mast cells but also by adaptive immune cells types, including infiltrated T helper(Th) 1 and Th-17 cells infiltrated in the psoriatic lesion and PsA inflamed joints (Perera et al. 2012) and it induces activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-B) signaling pathways (ref). It also activates several kinase signaling pathways as well as cell death programs (ref). In the context of inflammation,

Introduction

NF-B represents a master transcriptional regulator of both, the innate and adaptive immune system that induces expression of proinflammatory cytokines, antiapoptotic genes and genes involved in chronic inflammation maintenance (ref). The importance of this transcription factor (TF) in psoriasis and PsA pathogenesis is reflected by the association with disease of several genetic variants in some of the negative regulators of its proinflammatory activity, including NF-B inhibitor alpha *NFKBIA* and TNF receptor-associated factor 3 interacting protein 2 *TRAF3IP2* (ref).

Interleukin-23 (IL23) and Th17 axis represents a key loop for the maintenance of psoriasis and PsA inflammatory response and a very important link between innate and adaptive immunity. IL-23 is an innate regulatory cytokine, mainly produced by mDC and macrophages homing the inflamed skin and it binds to the IL-23 receptor (IL-23R), which expression is upregulated in the DC and T cells of the lesion and in circulating Th cells (ref). In psoriasis, IL-23 is the mediator for the pathogenic loop between activated KC and T cells (ref). Both IL-23 and IL-23R present protective and pathogenic genetic variants associated with psoriasis and PsA risk (ref). The activation of the IL-23 pathway leads importantly to increased IL-17 production through NF-activation by *TRAF3IP2* (ref). IL17 favors maintenance of the adaptive immune mediated Th17 response through recruitment and activation of neutrophils, induction of proinflammatory cytokines including IL-1 β and IL - 6 and also perpetuation of KC activation (ref) ([https : //www.ncbi.nlm.nih.gov/pmc/articles/PMC3580541/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3580541/)). More recently, interleukin 22 (IL-22) has arisen as another of the key cytokines in mediating the dysregulated crosstalk between the innate cells known as Th22 (ref). It mediates some of the histological changes in skin as well as AMP production by KC (ref).

Introduction

Environmental factors and disease

Several environmental triggers are known to be associated with increased risk and worsening of psoriasis and PsA development. A wide range of drugs including antidepressant, antihypertensive and anticytokine therapies have been clinically associated with initiation, exacerbation and worsening of psoriasis (Kim et al. 2010). Infectious agents such streptococcal throat infection have likewise been associated with development of type I psoriasis (Valdimarsson2009; Gudjonsson and of 2003; Diluvio et al. 2006). Consistently with other chronic inflammatory disease such as IBD and AS, recent studies have also observed perturbation in the composition of the gut and skin microbiota in psoriasis and PsA patients (Eppinga and rheumatology 2014). On the other hand, physical trauma, including tattoos, surgical incisions and mechanical stress can trigger the appearance of lesions in psoriatic uninvolved skin as well as joint inflammation in digits (Weiss2002; Nestle et al. 2009). Lastly, as for most of the complex diseases, behavioral factors including smoking, alcohol and stress have been linked to psoriasis and PsA without a clear conclusion of their involvement in triggering disease (Meglio et al. 2014).

1.2.3 Cell types involved in psoriasis and PsA pathogenesis

Identifying the most relevant cell types contributing to psoriasis and PsA pathogenesis remains challenging. There has been a reinterpretation of both phenotypes that understands them as dynamic and continuous processes where different cell types became predominantly important at different stages of the pathology.

KC are one of the most relevant cell type at early stages of psoriasis pathogenesis, which is reinforced by the genetic association between skin specific genes from the late cornified envelope (LCE) family and psoriasis (Tsoi et al. 2012). Several studies have shown the role of KC as immune sentinels through

Introduction

MHC-II antigen presentation and production of antimicrobial peptides (AMP), cytokines and chemokines (Black and of 2007). There is evidence of complex formation between the cationic AMP LL-37 and self-DNA/RNA released by KC upon damaged triggered by environmental factors (Lande et al. 2007). This complex acts as an antigen for activation of the skin-resident DC (Nestle et al. 2005) and that initiate and perpetuates the skin inflammatory response through secretion of pro-inflammatory cytokines, importantly IL-1, IL-6 and TNF- α (**Feldmeyer2007**; Arend et al. 2008; Nestle et al. 2009). Studies in mouse models have also shown development of psoriatic lesions in immunodeficient mice upon human xenotransplant of psoriasis skin(Boyman et al. 2004). Overall, these findings would support the hypothesis attributing the initiation of the chronic inflammatory response in psoriasis as the consequence of the epidermis dysfunction Proskch2008.

mDC and pDC are professional APC also considered important innate immune cells in disease initiation through T cell activation and subsequent triggering of the adaptive immune response (Mahil et al. 2016). pDC are circulating cells absent in healthy skin that infiltrate into the lesional and uninvolved dermis of psoriasis patients and get activated by the aforementioned KC self-DNA and LL-37 complex through Toll-like Receptor (TLR)-9 (Nestle et al. 2005; Lande et al. 2007). In contrast, quiescent mDC are epidermal resident cells and upon secretion of IFN- α by pDC a 30-fold increase of mature mDC is observed in lesional skin but not in uninvolved or healthy tissue (ref). Different mDC subpopulation mediate the Th1 and Th17 response as well perpetuation of KC activation through IL-23 production (ref). Studies in immunodeficient psoriasis mice models have shown that blockage of downstream IFN- α signaling or its production by pDC failed to induce T cell activation and onset of psoriasis (Nestle et al. 2005).

Introduction

Neutrophils are also thought to be closely involved in disease initiation through their ability to form neutrophil extracellular traps (NET) that contain host DNA and LL-37 (Hu et al. 2016). There is evidence of increased NET formation in peripheral blood and lesional skin of psoriasis patients and they seem to be contributing to pDC and CD4⁺ T activation (Hu et al. 2016). Neutrophils have also been identified in recent studies as one of the main sources of IL-17 production in the skin lesions (Lin et al. 2011) and they also release a wide range of proteases which some induce KC proliferation (**Mahil2006**).

In the context of the innate immunity, the involvement of monocytes and macrophages in psoriasis and PsA has not been extensively studied. Resident macrophages in the healthy dermis undergo a 3-fold increase upon skin lesion and they are involved in disease development through TNF α production (Perera et al. 2012; Mahil et al. 2016). Similarly, mice models for chronic psoriasiform skin inflammation have shown macrophage migration into the affected skin and TNF- α production for maintenance of the skin lesions (**Stratis2006; Wang2006**). Some studies using isolated monocytes from psoriasis patients PBMC have shown greater phagocytic and bactericidal activity compare to those from healthy individuals (**Bar-Eli1979**). Later studies have also shown increased circulating intermediate monocytes (CD14⁺ high CD16⁺ high) and monocyte aggregation in psoriasis patients causing enhanced platelet activation and angiogenesis (**Golden2015**). In PsA, synovial membranes levels of monocytes/macrophage metalloproteinases which mediate bone erosion through differentiation into osteoclasts are comparable to those found in RA joints (Hitchon et al. 2002). Overall, these observations highlight the systemic aspects of both pathologies.

Historically, T lymphocytes have been considered one of the most relevant cell types in initiation and maintenance of psoriasis and PsA and its role is also supported by genetic findings. Report cases in humans have demonstrated that bone marrow transplantation can initiate or terminate

Introduction

psoriasis (**Gardembas1990**; Eedy et al. 1990). The percentage of circulating T cells in psoriasis have shown reduced number of T cells in moderate-to-severe and severe psoriasis patients but increased percentage of the memory populations CD4⁺CD45RO⁺ and CD8⁺CD45RO⁺ when compared to milder phenotypes and healthy controls (**Lecewicz-Toru2001**; Langewouters et al. 2008). There is still controversy regarding the total CD4⁺ and CD8⁺ abundance and CD4⁺/CD8⁺ ratios in PBMC, which may be due to the phenotype heterogeneity of psoriasis patients in the different studies (**Lecewicz-Toru2001**; Cameron and of 2003; Langewouters et al. 2008). In PsA, no differences abundance of circulating T cells have been identified compared to healthy individuals (Costello et al. 1999).

In healthy skin, CD4⁺ and CD8⁺ are found in the dermis and epidermis, respectively (Clark et al. 2006; Perera et al. 2012) and an increase in activated memory CD4⁺CD45RO⁺and CD8⁺CD45RO⁺ can be detected as soon as 3 days after lesion appearance (Clark et al. 2006), highlighting the importance of the memory population. *In vivo* studies conducted in mice by Boyman and colleagues showed that development of psoriasis following engrafted human pre-lesional skin was dependent of local T cell proliferation without injection of additional factors (**Boyle2013**), supporting that recruitment of circulating T cells may be restricted to the priming event and it is minimal afterward (Perera et al. 2012). The relative importance of CD4⁺versus CD8⁺ in psoriasis initiation has been tested immunodeficient mice with pre-lesional skin xenografts suggesting a model where CD4⁺ but not CD8⁺ T cells were required for the progression of uninvolved to lesional skin in mice (Nickoloff and Wrone-Smith 1999). Interestingly, injection of CD4⁺ activated cells was followed by an increase in activated resident CD8⁺ T cells expressing the acute activation marker CD69 which suggests that skin CD4⁺ cells drive activation of resident T cells and that the activated CD8⁺ resident population would act as the main effector cells. In

Introduction

PsA, CD4⁺ are significantly more abundant than CD8⁺ cells in synovial tissues (Diani et al. 2015). However, among the CD8⁺ population the memory cells are prevalent in the SF and they are significantly increased when compared to controls (Costello et al. 1999). The contribution of regulatory T (Treg) cells have also been investigated with controversial outcomes regarding relative abundance and impaired function in both pathologies (Perera et al. 2012).

Based on the T cells cytokine profile, psoriasis and PsA has been demonstrated to be a type 1 Th/Tc disease, where activation of naïve CD4⁺ and CD8⁺ cells is driven by IL-12 and IFN- γ (Austin et al. 1999; Perera et al. 2012). Later studies have also identified additional T cells subsets including Th-17/Tc-17 and Th-22/Tc-22, characterised for the high production of IL-17 and IL-22, respectively, two cytokines very relevant for perpetuation of the inflammatory response (Mahil et al. 2016). The importance of Th17 cells and IL-17 production has been assessed in skin, joints and blood, where elevated IL-17 and also IL-23 mRNA and protein levels have been found in psoriasis and PsA patients compared to controls (Cai2012; Dolcino2015). It has been shown that the predominant CD8⁺ populations in the SF are also IL-17 producers and their abundance correlates with markers of inflammation and structural changes in the joint (Menon et al. 2014). This finding is in line with observations in skin and it suggests a prominent role of CD8⁺ IL-17 producing cells in the different stages of both pathologies. Understanding the importance of IL-17 has also led to the discovery of other immune cells producing this pivotal cytokine, including innate lymphoid (ILC) cells and $\gamma\delta$ T cells which have also started to be investigated in the context of psoriasis and PsA pathophysiology and treatment (Meglio et al. 2014; Leijten et al. 2015). IL-17 producing cell have also been hypothesised to be responsible for the link between skin and joint lesions. Although the precise mechanisms for transition between psoriasis and

Introduction

PsA is unknown, studies using psoriasis and RA mice models have shown that skin lesions facilitate arthritis and joint inflammation

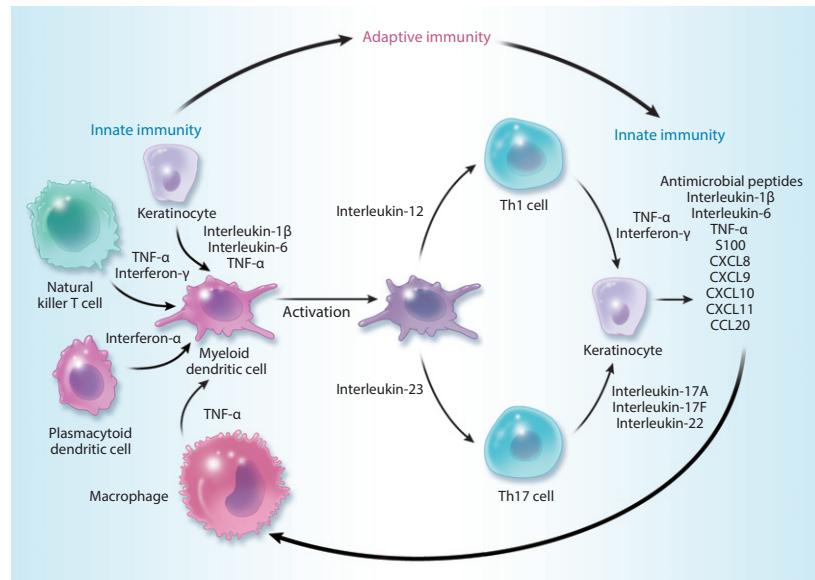


Figure 1.1: Figure adapted from (Nestle et al. 2009)

Introduction

1.2.4 Therapeutic intervention and prognosis

Currently, there is no cure for either psoriasis or PsA and the different treatments available are focused in managing symptoms. The approach to treat them are usually dependent on the disease severity. Cases of mild-to-moderate psoriasis are usually managed with topical therapies for which corticosteroids and emollients are the most commonly used and affordable ones (Menter et al. 2009). In psoriasis, other topical treatments are used in combination with corticosteroids such ultra violet (UV) light therapy and vitamin D analogues, which inhibit KC proliferation, stimulate KC differentiation and inhibits T cell proliferation (Rizova2001). For PsA patients presenting swelling of two or less joints, intra-articular injection of glucocorticosteroids together with joint aspiration have shown to reduce pain and inflammation as a short-time measure (Coates et al. 2016).

Treatment of most forms of PsA and more moderate-to-severe psoriasis require the use of a broad range of systemic therapies. For mild cases of PsA nonsteroidal anti-inflammatory drug (NSAID) are the most commonly used to help controlling the mild inflammatory symptoms (Coates et al. 2016). For more severe forms of PsA, disease-modifying antirheumatic drugs (DMARDs) including the an antagonist of folic acid methotrexate (MTX) and the phosphodiesterase 4 inhibitor apremilast are used with immunosuppressive effects on activated T cells and cytokine production, respectively (Schmitt2014; Keating2017; Gossec et al. 2016; Polachek et al. 2017). Biologic systemic agents are the most specific for the treatment of severe psoriasis and PsA. They are cell-based molecular species that modulate the immune response in a physiological manner (Perera et al. 2012). Among the biologic agents targeting cytokines, TNF α inhibitors (TNFi) have been broadly used for the past five decades to treat both, psoriasis and PsA due to the relevance of TNF α in disease. Three TNFi have been approved for the treatment of psoriasis: etanercept, infliximab

Introduction

and adalimumab (**Ahil2016**). In addition to those, certolizumab pegol and golimumab are also used in the management of PsA and other rheumatoid diseases (**Coates2016b**). All the TNFi are antibody-based agents but etanercept, which is a soluble receptor, and they show differences in the frequency and via of administration (Mease et al. 2000). Although TNF- α blockade is one of the most effective treatments, some patients experience common side effects such as increased risk of infection, reactivation of latent infections, demyelinating disease and induced pustular psoriasis have been identified (Nickoloff and Nestle 2004). Between 20 to 50% of the patients fail to respond to the first TNFi, due to side-effects or tolerance, and they require switching to a second or third one (Abramson and and 2016). New biologic therapies have been developed to target other key cytokines in the pathogenesis of PsA and psoriasis, such as IL-12 and IL-23 (ustekinumab) or IL-17 (secukinumab and ixekizumab) (Mahil et al. 2016). These new biologics represent a substantial benefit for treating patients and they are routinely administered to individuals failing to respond after a switch to a second TNFi (**Coates2016b**).

1.3 Genetics of psoriasis and psoriatic arthritis

As complex diseases, the risk to develop psoriasis and PsA is not only influenced by the surrounding environmental conditions but also by the genetic background of each individual. Determining the magnitude of contribution of the genetic factors in the development of these disease and identifying the exact genes or regions involved in the predisposition to psoriasis and PsA remains challenging.

1.3.1 Heritability

Several studies have shown a trend towards the increase of psoriasis and PsA prevalence over the last 30 years in different countries (Organization 2016).

Introduction

This importantly reflects changes in life style habits and it highlights the need to better understand the genetic factors that predispose to disease upon interaction with environmental stresses.

The contribution of genetics in the development of psoriasis has also been demonstrated in several twins studies. The concordance of psoriasis has been shown to be greater in monozygotic (33-55%) compared to dizygotic (13-21%) twins with some variation between studies and populations, estimating an 80% of heritability in this condition (**Faber1974; Penderson2008; Duffy et al. 1993**). For PsA, similar concordance between mono- and di- zygotic twins has been shown, probably due to lack of statistical power and appropriate diagnosis (**Penderson2008**). In the general population, approximately 40% of the patients with psoriasis or PsA have family history in first degree relatives (Gladman et al. 1986). Interestingly, the recurrence in first-degree relatives ha been shown to be greater in PsA (40) compared to psoriasis (8) in a study in Icelandic population (Chandran et al. 2009). This could suggest differences in the heritability between the two phenotypes and maybe an stronger genetic contribution in PsA.

1.3.2 Non-GWAS and linkage studies

Different approaches have been undertaken to uncover the genetic variability contributing to the predisposition to psoriasis and PsA. The appearance of next generation sequencing (NGS) techniques and the progressive reduction of cost has allowed to move from candidate genes studies looking at the genetic variability at particular locus to a genome-wide approach.

The study of psoriasis and PsA genetics architecture started with linkage analysis in family pedigrees with autosomal dominant condition to try to . The use of this approach yielded nine psoriasis susceptibility loci (PSORS1-9) (Capon 2017) with the strongest association in PSORS1 (**International2003**). PSORS1 lies within chromosome 6p21.3 and its identification by linkage

Introduction

analysis confirmed the association previously identified in serological studies between psoriasis susceptibility and the MHC I (**Rusell1972; Tiilikainen1980**). Importantly, Mendelian forms of disease with rare highly penetrant mutations have been identified in family studies for two genes within PSORS2 (17q25): zinc finger protein 750 (ZNF750) (**Tomfohrde1994**) and caspase domain family member 14 (CARD14) (Jordan et al. 2012). Rare gain of function and *de novo* mutations and also common variants in CARD14 have been identified in psoriasis and PsA patients, suggesting an important role of the genetic variation in this gene for Mendelian and multi-factorial forms of disease (Jordan et al. 2012; Tsoi et al. 2012). In PsA, a region close to the psoriasis PSORS8 was also identified (Karason et al. 2003). Nevertheless, the ability of independent studies to only reproduce PSOR1, 2 and 4, highlighted the limitations of the linkage studies to understand the genetics of complex diseases (Capon 2017). Gene based studies in psoriasis and PsA have also identified the relevance of genetic variability in the activating killer immunoglobulin receptors 2DS1 (KIR2DS1) (**uszczek2004; Williams2005**), similarly to AS and RA (**Yen2001; Carter et al. 2007**). This receptor is expressed on NK and NK T cells and mainly triggered by HLA-Cw*06:02. Similarly, specific association with PsA but not psoriasis was found for microsatellites and promoter polymorphisms in TNF- α (**H”-o”hler2002**).

1.3.3 Genome-wide association studies

The technological advances experienced in sequencing and genotyping has allowed to implement association studies at a genome-wide scale. The genome-wide association studies (GWAS) have benefit from the understanding of common (frequency >1%) single base-pair changes known as single nucleotide polymorphisms (SNPs) in different populations through whole genome sequencing (WGS) projects such as HapMap (**The international HapMaP Consortium**) project and the 1000 Genomes project (**The 1000 Genomes**). GWAS have

Introduction

focused in the association to a particular phenotype of single-nucleotide bi-allelic substitutions with minor allele frequency (MAF) >5% in a case-control design (Ku et al. 2010). This followed the hypothesis driving the field of complex diseases where common diseases are more likely to be caused by common variants (Schork et al. 2009). Due to the organisation of the genome into segments of strong linkage desequilibrium (LD) where genetic variants are strongly correlated with each other (measured as a squared correlation r^2), the genotyped SNPs in GWAS are used a proxy for the disease causative variant. In the context of complex diseases, GWAS have greater power than the previous linkage studies when looking at the influence of many loci with low penetrance and small effects in disease risk and they have identified common alleles conferring effect sizes (Cui et al. 2010). Disease causal variants can be non-genotyped SNPs or other type of genetic variability such as copy number variants (CNVs), also highly frequent in the genome, which are less widely studies by GWAS (Hirschhorn 2005; Ku et al. 2010).

Since the improvement of genotyping technologies, several GWAS studies have been performed in psoriasis and PsA (Table). Since the first psoriasis GWAS analysis in 2007 more the number of risk loci associated with psoriasis and PsA have increased. Currently, there are a total of sixty-three associations to psoriasis and PsA susceptibility at a genome-wide significance ($p\text{val} \leq 5 \times 10^{-8}$) which explain 28% of the heritability (Tsoi2017). Most of the studies have been performed in Caucasian European or North American cohorts but lately few GWAS studies Chinese populations have also been published (Zhang et al. 2009; Sun et al. 2010; Yin et al. 2015). The earlier studies were performed in discrete cohort sizes with moderate power that confirmed association with loci overlapping the PSOR1, PSOR2 and PSOR4 regions from the linkage studies (). HLA-C has consistently been identified in all the GWAS studies as the most significant locus with the greatest effect size which account for approximately

Introduction

50% of psoriasis heritability. Additional MHC-I and MHC-II associations were identified doing step-wise conditional analysis for the HLA-C association and revealed contribution of HLA-A, HLA-B and HLA-DQA1 to disease risk (Okada et al. 2014)

One of the most informative GWAS studies was the one performed using the Immunochip genotyping platform, which includes greater genotyping density only at 186 immune relevant loci identified in previous GWAS studies across different inflammatory diseases (Tsoi et al. 2012). This study uncovered 15 new associations and it also included meta-analysis analysis with the largest available psoriasis cohorts (Tsoi et al. 2012). This meta-analysis has later been expanded, being the largest one the Tsoi and colleagues in 2017 (**Tsoi205; Tsoi2017**). This latest study has revealed additional 16 associations reinforcing NF κ B and cytotoxicity pathways in disease.

Meta-analysis of GWAS across Caucasian and Chinese populations have showed the value of this trans-ethnic approach to identify new association and understand the genetic differences between populations contributing to disease (Yin et al. 2015). In this study four new loci including *LOC144817*, *COG6*, *RUNX1* and *TP63* were associated with psoriasis and PsA in both populations at non-coding gene regions. Genetic heterogeneity was also observed for 10 of the GWAS reported loci such as *ELMO1* and *TYK2* among others, when looking at the association in Caucasian and Chinese cohorts independently.

Changes in the frequencies of HLA-C and HLA-B alleles had already shown evince of genetic differences between psoriasis and PsA supporting the importance of independent GWAS studies(**Winchester2012**; Okada et al. 2014). Therefore, although most of the GWAS include mixed psoriasis and PsA cohorts some of them have performed stratified analysis between the two subcohorts. As result, a new association for a some of the loci such as *TRAF3IP*, *IFNLR1*, *IFIH1* and *NFKBIA* were confirmed to be associated when using only PsA cases

Introduction

(**Stuart2015**; Ellinghaus et al. 2010). When performing association analysis of psoriasis versus PsA, a signal in chr18 overlapping the long non-coding RNA (lnc-RNA) *LOC100505817* was identified. Moreover, a PsA GWAS using the Immunochip platform revealed a PsA specific association in chromosome 5q31 (Bowes et al. 2015). Overall, GWAS studies have demonstrated shared genetic susceptibility between psoriasis and PsA, but have also highlighted some specificity that may support a difference in the genetic architecture of both diseases. It is important to take into account that these results are affected by imprecise phenotyping of cases, which entails one of the many challenges when trying to compare both diseases.

1.3.4 Relevance of non-coding versus coding variants in disease susceptibility

Approximately 88% of all GWAS associations map within non-coding regions and only the remaining 12% account for variants in coding regions that likely cause non-synonymous mutations impacting in the final protein products (**Welter2013**).

Exome psoriasis association studies in Chinese and Caucasian populations have increased the number of coding variants with putative effect in the protein structure (**Tang2014**; **Zuo2015**; Dand et al. 2017). These studies confirmed some of the previously identified missense associations in *CARD14*, *ERAP1* and *ZNF815A*, revealed new common variants at previously GWAS associated loci such as *IL23R*, and identified protective rare missense changes at *TYK2* and *IFIH1* (**Tang2014**; Dand et al. 2017). Nevertheless, results from extensive exome studies suggest that non-synonymous SNPs have a limited contribution to the overall genetic risk of psoriasis compared to non-coding variants (**Tang2014**).

Non-coding variants associated with disease are usually found at introns, intergenic regions or gene deserts and become causal through regulation of gene

Introduction

expression in different manners. The functional mechanism by which these variants affect gene expression depends on the type of regulatory element where the variants are located, including enhancer, silencers, promoters and the 5' and 3' untranslated region (UTR) of genes (**Ward2012**). In addition, the precise impact of the variant will also be influenced by the dynamic cell and context specific functional epigenome, which involves different DNA and DNA-protein bound modifications such as methylation and acetylation (**Feil2013**). Overall, the combination of different epigenetic marks determine the accessibility of the surrounding chromatin and the likelihood of a SNP having a role in the regulation of gene expression. Causal GWAS variants can alter the expression of target genes through different mechanisms including changes in chromatin accessibility, histone modifications, protein binding (for example TF), DNA methylation and binding of micro-RNAs (miRs) (Knight and C 2014) (See 1.4.X.Understanding the epigenetic landscape in complex diseases)

Expression quantitative trait loci (eQTL) analysis one of the most informative approaches to connect the effect of a SNP allele with the regulation in expression of a particular gene. This method performs an association between gene expression and SNPs in *cis* (<1Mb) or *trans* to the gene of interest (ref). Identification of *cis* and *trans* eQTLs could reveal gene networks, as in T2D where a *cis*-eQTL for the TF Kruppel-like factor 14 (*KLF4*) is also associated with other genes in *trans*, highlighting downstream targets regulated by that TF (**Small2011**). Similarly, *cis* effect of a variant in cytokine coding genes could be accompanied by a *textit{trans}* association with downstream members in the same signaling pathway. For example, previous studies in our laboratory described that, in monocytes, a *cis*-eQTL in the IFN- β gene is associated in *trans* with IFN-modulated genes, including IFN regulatory factor 9 (*IRF9*) amongst others (**Fairfax2014**). Nevertheless, eQTLs and chromatin interaction only provide indirect evidence of an association between a variant and regulation of gene

Introduction

expression. Functional assays are required to confirm and fully understand the mechanisms behind non-coding variant GWAS associations, as it will be detailed in section xxxx (Edwards et al. 2013).

1.3.5 The role of GWAS studies in highlighting immune-relevant cell types and pathways

GWAS represent a biologically unbiased approach to shed some light into pathophysiological relevant cell types and molecular pathways associated with disease. In the field of common immune-mediated diseases, GWAS have underscored some of the most important cell types for which genetic variation is functionally relevant. For example, in T2D the strongest GWAS associations showed to be enriched at pancreatic β cells and to affect genes involved in insulin secretion, consistently with the insulin resistance that characterises the disease pathology (Visscher2017).

Immune diseases have benefited from the use of the immune targeted genotyping array Immunochip to perform a systematic comparison of the genetic architecture across the different conditions. For example psoriasis and PsA share risk loci with AS, CD, MS, RA and T1D, among others (**ImmunoBase**). However, genetic overlap where the signal is the same across different diseases have not necessarily the same direction and a risk allele for one can be protective for another reflecting true differences in the pathophysiology of different immune mediated diseases. The better understanding of immune-related diseases has led to identification of shared susceptibility loci and the use of therapeutic interventions across diseases, such as anti IL-23 and anti IL-17 antibodies to treat psoriasis, PsA, AS and IBD (Visscher2017). Cross disease association studies have also been performed for the simultaneous analysis of AS, UC, primary sclerosing cholangitis (PSO), CD and psoriasis (Ellinghaus et al. 2016). This study revealed the greatest genetic pleiotropy of psoriasis with AS and CD, this meaning same

Introduction

alleles predisposing to disease risk (**ImmunoBase**). Among the 206 multi-trait associated loci, enrichment was found for regulatory elements in bone marrow, NK and T cells as well as for immune response pathways, supporting the contribution of GWAS to the biological understanding of disease. In the case of psoriasis and PsA, most of the GWAS risk loci highlight genes that belong to a small number of pathways and they are enriched for regulatory elements of several cell types (Capon 2017). Nevertheless, it is important to bear in mind that in most of the cases non-coding variants from GWAS studies lack of functional characterisation and they tend to be associated arbitrarily to the nearest gene or the closest gene which fits into current knowledge about pathophysiology. This bias to some extent the genes that contribute to enrichment of certain pathways and the efficacy of drugs developed to target some of those genes has helped to further confirm their truly involvement in disease.

Antigen presentation

HLA-Cw^{}0602*, the strongest GWAS association in psoriasis is also associated with other diseases such as Hepatitis C, PSO and Graves disease (Blais et al. 2011). *HLA-Cw^{*}0602* is involved in antigen presentation and the absence of differences at the transcript level between normal and psoriatic individuals suggests the association is not explained by alteration in regulation of gene expression (Hundhausen et al. 2012). The relevance of antigen presentation in psoriasis and PsA has been reinforced by the GWAS association of the endoplasmic reticulum aminopeptidase 1 *ERAP-1* gene, which codes for an aminopeptidase involved in the trimming of peptide antigens. GWAS studies identified that *ERAP-1* was associated with psoriasis and PsA only in individuals carrying one copy of the rs10484554 *HLA-C* risk allele (Strange et al. 2010). Moreover, the same study also found dependent association of an SNP nearby the zeta chain of T cell receptor associated protein kinase 70 *ZAP70* gene and

Introduction

HLA-Cw^{}0602*. ZAP70 is a tyrosin kinase that binds the CD3- ζ of the TCR and it is involved in the CD8 $^{+}$ cells auto-reactivity regulation (Picard et al. 2009), overall highlighting the role of HLA-dependent CD8 $^{+}$ dysregulation in psoriasis and PsA. At the clinical level, *HLA-Cw^{*}0602* and *ERAP-1* have also been associated with pediatric psoriasis onset together with other GWAS loci (Bergboer and of 2012). These epistatics phenomena, where association of one gene is dependent by the presence of another, has also been found between other HLA class I molecules such as *HLA-B^{*}27* and *HLA-B40* and *ERAP1* in AS (Cortes2015b; Evans et al. 2011). The signal at chromosome 5q15 for *ERAP1* allele is the same in psoriasis and AS and they also share the direction of the effect (ImmunoBase). Several studies have revealed that the disease associated polymorphisms in AS increased *ERAP-1* and *ERAP-2* gene expression and also altered splicing, resulting in ERAP-1 protein isoforms with increased activity (Constatino2015; Hanson et al. 2018). Regarding cell types, the role DC and macrophages involved in antigen presentation is reinforced by genetic evidence.

Skin barrier

GWAS have highlighted KC specific genes such the *LCE* gene cluster or genes with a very relevant role in the skin such as *CARD-14*, both previously mentioned. Further studies in the *PSORS4* have revealed that association with diseases is due to a deletion comprising two of the genes within this family, *LCE3B* and *LCE3C* (*LCE3C_LCE3B_del*)(Cid et al. 2009). In normal skin, expression of *LCE3B* and *LCE3C* is very low and it is induced upon barrier disruption. These are proteins that form the cornified envelope on the most external layer of the epidermis and are thought to be involved in epidermal terminal differentiation (Bergboer2011). Additionally, epistasis between this deletion and *HLA-Cw^{*}0602* has been identified in certain populations including Dutch and American (Riveira-Munoz2011; Cid et al. 2009). Overall, the lack

Introduction

of *LCE3B* and *LCE3C* expression in psoriasis patient could lead to an impaired repair following skin disruption and facilitate microorganisms infection and the triggering of the dysregulated immune response. Treatment of psoriasis patients with UVB radiation have proved upregulation of *LCE3E* expression after 48 hours contributing to amelioration of the skin lesions(Jackson et al. 2005). *CARD14* is primarily expressed in epithelial tissues where it is involved in recruitment and activation of the NF- κ B pathway (Blonska and research 2011). Common and rare pathogenic mutations of *CARD14* in KC cell lines led to increased activation of NF- κ B as well as overexpression of psoriasis-associated genes including *IL6*, *IL36*, *TNFA* and *TNFAIP2*, among others (**Jordan2012b**).

NF- κ B and TNF pathways

The NF- κ B pathway is involved in the regulation of the innate and adaptive immune response. Several psoriasis and PsA GWAS loci have been mapped to gene members of the NF- κ B and TNF signaling pathway such as *TNIP1*, *TNFAIP3*, *NFKBIA*, *REL*, *TRAF3IP2*, among others *NF- κ B* is a dimeric TF, formed by assembly of two of the five proteins from the NF- κ B family, that translocates into the nuclei upon cytokine stimuli, importantly TNF- α . Dysregulation of a feedback loop between TNF- α and NF- κ B contributes to the development of many chronic inflammatory diseases (**Liu2017**) and neutralisation of TNF- α is used for treatment of many immune-mediated diseases, as previously described. In psoriasis, elevated levels of NF- κ B are found in lesional skin compared to uninvolved and normal skin (Lizzul et al. 2005). Psoriasis and PsA GWAS association with *NFKBIA*and *REL*, two of the genes coding for *NF- κ B* subunits, are driven by SNPs at an intergenic region, not having yet directly evidence to their effect over regulation of these genes (**GWAS studies**). *REL* has been associated with other inflammatory diseases, including CD and RA (**ImmunoBase**) and, interestingly, the RA risk allele has a protective effect in

Introduction

PsA showing opposite direction effects (Bowes et al. 2012). The relevance of members downstream TNF- α signaling is highlighted by the GWAS association of *TNIP1* and *TNFAIP3*, protein products interact with each other and participate in the regulation of NF- κ B activation. SNP variants in these regions have also been identified for CD, UC and SLE, among other immune diseases, reinforcing the relationship between these two pathways in chronic inflammation (**ImmunoBase**). In mice, a chromosomal region including *Tnfaip3* induces psoriasis in a TNF- dependent manner and it also increases atherosclerosis risk, one of the most prevalent co-morbidities in psoriasis and PsA (**Wang2008; Idel2003**). The association with the interacting protein TRAF3IP2 is stronger in PsA than in psoriasis (**H”–u”ffmeier2010**) and a haplotype including two missense mutations and two intronic variants has been reported in two different studies (**H”–u”ffmeier2010; Ellinghaus et al. 2010**). The missense mutations rs33980500 located at a highly conserved region of the TRAF3IP2 protein has shown reduced affinity TRAF interacting proteins, which has a downstream effect in NF- κ B activation and the IL-17/IL-23 axis(**H”–u”ffmeier2010**). Exome-sequencing studies have also lately implicated variants with predicted evidence on protein structure and function at *TNFSF15*, a TNF ligand family protein induced by TNF- α , mostly expressed in endothelial cells and with a role in regulating NF- κ B and MAP kinases activation (**Wang2014; Dand et al. 2017**). The latest psoriasis and PsA meta-analysis study of Tsoi and colleagues has identified three additional associations with genes belonging to the NF- κ B pathway, reinforcing the implication of NF- κ B activation in psoriasis and PsA development (**Tsoi2017**). Nevertheless, approved drug for the treatment of psoriasis or PsA targeting directly any member of this pathway are lacking, since some studies have shown that naturally occurring constitutive deficiency in NF- κ B leads to immune related pathologies (**Orange2005; Puel2004**)

Introduction

Type I IFN and innate host defense

Psoriasis and PsA GWAS associations have highlighted genes involved in innate immunity including host response to virus and bacteria, which importantly involve genes from type I IFN signaling pathway. Mapping of several GWAS loci to genes from the type I IFN signaling pathway together with clinical and experimental data has reinforced the role of pathogen response psoriasis and PsA (**Nextle2005**). Some of the genes highlighted by GWAS studies include *IL28RA*, *IFIH1*, *TYK2*, *RNF114*, *ELMO1* and *DDX58*. Some of these genes are also susceptibility loci for other immune-mediated diseases. GWAS lead SNPs causing a missense mutations in *TYK2* have been identified in several immune-mediated diseases including CD, IBD, T1D, RA and MS, in addition to psoriasis and PsA (**ImmunoBase**). *TYK2* codes one of the Janus kinases (JAK) protein family which phosphorylates the IFN- α and IFN- β receptor α chain and initiates the IFN type I downstream response (**Calamontici1994**). Exome-sequencing and GWAS studies have identified two independent missense mutations predicted to impair its catalytic activity to phosphorylate the receptor and initiate the downstream inflammatory cascade, overall having a protective effect for the risk to develop disease (Strange et al. 2010; Tsoi et al. 2012; Dand et al. 2017). Currently, tofacitinib is the only inhibitor of all the JAKs that is used for RA treatment (**van Vollenhoven2012**) and despite its side-effects it is currently under clinical trials for approval to treat other immune-related diseases together with more specific JAK inhibitors (Baker and diseases 2017). Moreover, several drugs targeting type I IFN pathway members are also being developed. For example Monoclonal Ab against IFN- α subtypes have failed to suppress the IFN gene signature in psoriasis patients and new approaches towards blocking the IFN- α receptor have shown greater efficacy in SLE (Furie et al. 2017). Regarding upstream targets of the IFN I pathway, the psoriasis and PsA GWAS intronic variant at *ELMO1* is essential for activation of the pathogen-sensing receptors

Introduction

TLR7 and (**TLR9**) and the subsequent IFN- α production in pDC (Tsoi et al. 2012). Currently, clinical trials testing inhibitors of these TLR receptors are being conducted in SLE (Baker and diseases 2017).

IL-17/IL-23 axis

Together with the TNF pathway, the IL-17/IL23 axis is the most widely targeted by biological therapeutics. GWAS studies have suggested the relevance of this pathway in psoriasis and PsA by several associations including *IL23A*, *IL23R* and *IL12B*. The cytokine IL-23, involved on a wide range of pro-inflammatory processes as previously explained, is formed by two subunits: IL-23A/p19 and IL12-B/p40, also a component of IL-12. For both of them, GWAS association has been established by proximity of non-coding lead SNPs to these genes but direct functional evidence in regulation of their expression has not yet been established (Cargill et al. 2007; Strange et al. 2010; Tsoi et al. 2012). Nevertheless, transcriptional studies have shown increased levels of p40 and p19 in psoriasis lesional skin and a role of both subunits in the abnormal KC differentiation (Lee2004; Zhu2011). Similarly, GWAS associations with *IL23R* has been reported in several studies (Nair et al. 2008; Strange et al. 2010). Particularly, a study in German and American Caucasian confirmed a shared association between CD and psoriasis of a two SNPs haplotype, which includes a missense variant (Nair et al. 2008). This missense variant involves an arginine to glutamine exchange (Arg381Gln) which has a protective effect under an inflammatory environment, including CD (Duerr2006). Conversely, this haplotype is not associated with psoriasis risk in Chinese population where a different non-synonymous potentially damaging variant has been reported as the potential functionally meaningful (Tang2014). Interestingly, a secondary *IL23* signal to the reported by Tsoi et al., 2012 has been specifically associated with PsA and the independency from AS secondary signals for the same locus

Introduction

has also been demonstrated *Tsoi2012, Bowes2015*. The genetic relevance of the Th-17 pathway is partly explained by these associations with potential effects on the IL-23 response and its role in Th-17 cell differentiation and activation. However, GWAS associations at intronic variants nearby genes of the Th-17 pathway have also been identified, such as interferon regulatory factor 4 (*IRF4*) and the signal transducer and activator of transcription 3 (*STAT3*), also associated with CD and MS (**Immunobase**; Tsoi et al. 2012). Both transcription factors are involved in the overall control of the Th-17 differentiation process (**Huber2008; Harris2007**). Moreover, previously mentioned GWAS associated genes with the NF-B and TNF pathways such as *TRAF3IP2*, *NFKBIZ* and *TYK2* are shared with the IL-23/IL-17 axis, stressing not only the relevance of this pathway but also the importance of pathway cross-talk. The relevance of this axis in the aetiology of psoriasis and PsA is reinforced by the fact that the individual blockade of the IL-17A and IL-23 appeared to be more effective than the use of anti-TNF drugs (**Griffiths2015; Blauvelt2017**). Interestingly, the inhibition of IL-17A using secukinumab is effective in the treatment of psoriasis, PsA and AS, whereas it worsens CD, for which the treatment using antibodies against IL-12/23p40, as the previously mentioned ustekinumab, have a much prolonged benefit compared to the other diseases (**Patel2012; Hueber2012; Blauvelt2017b**). Overall, this stressed the importance of the Th17/IL-23 axis in inflammation and demonstrates that blocking the pathway at different levels translates into different effects within and across inflammatory diseases.

Intergenic regions and genome-wide pathway enrichment analysis

As previously mentioned, most of the GWAS associations are located at intergenic regions or gene deserts difficulting their functional characterisation and biological relevance. Some examples in psoriasis and PsA include chr1p36.23, chr2p15, chr6q25.3 and chr9q31.2. One of the most interesting

Introduction

regions is chr2p15, which lead SNP and direction of association is shared with AS (**Immunobase**). Within this locus, the closest gene to the association signal is *B3GNT2* but other genes with a role in the immune response like *CMMD1* are also proximal (**Maine2007**). Among the other intergenic associations chr1p36.23 is shared with UC and chr6q25.3 has also been reported in MS, CD and RA (**Immunobase**). The association at chr1p36.23 is proximal to a number of gene candidates including *RERE*, *SLC45A1*, *ERRFI1* and *TNFRSF9* (Tsoi et al. 2012). Unpublished capture-HiC data using the immortal KC cell line HaCaT has revealed interaction of SNPs in this locus with the promoter of the *ERRFI1* gene, which encodes an inhibitor of the epidermal growth factor receptor signaling required for normal KC proliferation (**Ray-Jones2017**). Nevertheless, the same locus could be an enhancer for other nearby genes when looking at a different cell type, which would reinforce the importance of the cell type specificity in functional studies. In the same lines of identifying relevant biological processes, new approaches using genetic association data have allowed the performance of genome-wide pathway analysis. This analysis represents a more powerful and biological meaningful way than GWAS to study the association of functionally related genes with disease risk. In psoriasis, genome-wide pathway analysis has revealed association of novel processes, such as retinol metabolism, transport of inorganic ions and aminoacids and post-translational protein modifications not previously related with the disease aetiology (**Aterido2015**). Interestingly, *B3GNT2* is one of the genes belonging to the post-translational protein modifications pathway validated in this study. Overall, these results have highlighted unexplored pathways in disease and open new biological mechanisms that may be contributing to the risk and progression of psoriasis and PsA.

1.3.6 Limitations and future of GWAS studies

Although GWAS have made a great contribution into the understanding of the genetic component of complex diseases, several limitations need to be considered when interpreting the results of these type of genetic approach.

One of the GWAS limitation is due to the LD block structure of the genome, as the disease associated loci are large and include hundreds or thousands of SNPs equally likely to be causal. Therefore, an association between a genetic locus and a disease does not reveal neither the causal variant, which could be any of the variants in high LD r^2 with the tag SNP, or the target gene and genetic mechanisms driving the association. Additional genotyping, statistical fine-mapping and epigenetic data are required in order to shed light towards the identification of the causal SNP.

GWAS are very much dependent on the sample size, which will have great impact on the effect size of the associated variants to disease that can be uncovered (Visscher2017). In addition to the effect size, GWAS have a higher statistical power to identify association with common SNPs than with rare variants for any sample size. Since for two variants to be in high LD r^2 their allele frequencies need to be similar, arrays tagging common SNPs lack of power to detect associations due to rare variants (Wray2005). This has partly tried to be overcome by improving the design of the genotyping arrays. For example, the Immunochip incorporated SNPs with MAF<1%. However, it has failed to identify association driven by rare SNPs in loci already reported in GWAS for different immune diseases (Visscher2017). Although the sample size and adequate coverage of rare variants may be contributing, the role of rare variants in common diseases have also been largely discussed and opposing views are reflected in the common disease common variant and common disease rare variant hypothesis.

Introduction

Another concern is the heritability missed to be explained by the risk alleles associates with different complex diseases by GWAS. For example, in T2D or height only 5% and 10% of the total heritability could be explained, respectively, with the early GWAS associations (Ku et al. 2010; Yang et al. 2010). Later studies in height proposed a model with two main sources of missing heritability: SNPs association with small effect not passing the genome-wide statistical significant threshold and the rare variants not tagged by common SNPs due to low LD (Yang et al. 2010). Exome SNP arrays and greater sample size uncovered 83 height associations of SNPs with MAF<5% that individually accounted for the same amount of variation as previously detected common variants (Marouli et al. 2017). Similarly in psoriasis new associations such us the intronic signal at (TNFSF15) and rare alleles at already identified signals could suggest that some of the unexplained heritability would come from new associations and rare alleles at previously reported loci (Dand2017.). Moreover, heritability could have been overestimated assuming additive effect instead of epistatic interaction between different associated loci contributing to trait heritability (Zuk2012).

In addition to rare SNPs, other sources of common variation such as CNV, small ($\geq 1\text{Kb}$) insertions/deletions (indels) and inversions could also contribute to the missing heritability. The 1000 Genome Project and HapMap have helped to better understand these other sources of variation and later genotyping platforms such as the Illumina Human 1M Beadchip, the Affymetrix 6.0 and the Immunochip have included probes for CNV and small indels (Ku et al. 2010). Incorporation of new genotyping platforms have allowed to identify genome-wide associations of CNV with autism and schizophrenia, among others (Glessner et al. 2009; Marshall et al. 2017). CNV in LCE has been proved to be the causal for the association to psoriasis and PsA, as previously mentioned (Cid et al. 2009). However, genome-wide studies have failed to yield reproducible results (Uebel2017).

Introduction

In the case of translocations and inversions, neither arrays or widely-used short reads NGS technologies are appropriate to detect this type of variation. Although this type of variation has a role in several disease genotypes (Feuk2010), detection of translocations and inversions at a genome-wide scale is still very and their real frequency underestimated (Ku et al. 2010). There are some statistical methods that use dense SNP genotyping to detect an unusual LD pattern among the SNPs as a read oy for chromosome rearrangements (Bansal et al. 2007). Nevertheless, implementation of whole-genome sequencing (WGS) using long reads are the best tool to accurately assess this genetic variability (Visscher2017). Overall, the state of the field has evidenced that WGS technologies will naturally replace genotyping arrays in GWAS as they become more affordable.

1.4 Functional interpretation of genome-wide association studies in complex diseases

1.4.1 Overcoming the limitations of GWAS: post-GWAS studies

GWAS studies shed limited light on the link between genetic variants and disease mechanisms. As previously mentioned, GWAS report associations with disease for a particular locus tagged by a SNP in LD with many other variants in the same region. Nevertheless, the most significant associated SNP may not be the functional causal variant and the results could also be biased by the inability of the genotyping platforms to capture all the genetic variability in each locus Edwards2013. Statistical fine-mapping approaches have been designed to partially overcome those limitations and refine each GWAS locus to the strongest associated SNPs. In addition to this, the overlap of the fine-mapped variants with functional data will further help to narrow down the set of candidate causal SNPs. The integration of cell type and context specific epigenetic data, including

Introduction

chromatin accessibility, histone modifications and DNA methylation can help to determine the chromatin state where the variant is located and the potential in regulating gene expression Petronis2010. Additionally, the incorporation of gene expression, eQTL analysis and chromatin interaction data will help to establish a relationship between non-coding variants and putative gene targets. Finally, establishment of the functional relationship between the genetic variant and the disease phenotype will involve the establishment of appropriate cellular assays and *in vivo* animal models.

1.4.2 The use of fine-mapping to prioritise causal variants

Fine-mapping strategies can partially overcome two of the main limitations of the GWAS studies: the association of hundreds of SNP per locus due to the extensive LD blocks inherited together (haplotype)and the incomplete coverage of the human genetic variation. The aim of fine-mapping analysis is reducing the size of the GWAS genomic intervals and yielding a minimal set of SNPs which will contain the causal variant and explain most of the association for that particular region (**Spain2015**).

Fine-mapping studies require extense genotyping that to meet the assumption that the causal variant will be likely interrogated in the analysis. This can be achieved by WGS, dense genotyping arrays and imputation using publicly available data. Recently, the Immunochip array has been extensively used for most of the immune-mediated inflammatory diseases GWAS, including psoriasis and PsA, since it has been customised to increase the density of genotyped variants at previously associated immune-relevant loci in a cost-effective manner (**Trynka2011**). Similarly, imputation methods using WGS reference panels, such as the aforementioned HapMap and 1000 Genomes Project, have offered genome-wide coverage for SNPs and CNVs with MAF >1% across different ancestry groups (Abecasis GR 2012). More recently, the UK10K project has improved the

Introduction

quality of imputation specifically for rare variants with MAF between 0.01% and 0.5% (**Chou2016**). Interestingly, exhaustive fine-mapping using a customised genotyping array has been conducted for eight psoriasis GWAS loci using a frequentist approach which measure the association of each SNP through p-values (Das et al. 2014).

Nevertheless, Bayesian statistical analysis has been chosen over the frequentist approach to increase the resolution of the GWAS associations and facilitate the identification of relevant genes and disease mechanism. Bayesian fine-mapping quantifies the evidence of association of each of the genotyped or imputed SNPs as Bayes Factor(BF) and used them to calculate posterior probabilities (PP), which in the context of the fine-mapping data represent the probability of each SNP to drive a particular association (**Wakefield2007**). The output of this type of studies is a credible set of SNPs accounting for 95 or 99% of the PP, since including only the most significant SNP would failed to report the causal variant in approximately 97.6% of the fine-mapped loci (Bunt et al. 2015). Bayesian fine-mapping has been systematically applied to the set of GWAS loci identified for several immune-mediated diseases, including T2D, IBD, AS and SLE (**Gaulton2015; Sun2016; Huang2017**; Maller et al. 2012; Bunt et al. 2015). In contrast, systematic fine-mapping studies for all the sixty-three psoriasis GWAS loci have not been performed. In PsA, Bayesian fine-mapping has been conducted for some of the Immunochip GWAS associations, including the 5q31 PsA-specific locus (Bowes et al. 2015). One of the main limitations of the traditional Bayesian approaches refers to another assumption made by the model where only one causal SNP is considered per locus. To address this issue, step-wise conditional analysis is performed at each locus followed by calculation of PP and identification of credible sets of SNPs Maller2012,Bunts2015. Improved stochastic Bayesian fine-mapping outperforms the step-wide Bayesian methods by avoiding the biases of the conditional analysis and considering all the possible

Introduction

models regarding number of putative causal SNPs driving the association of each locus (**Wallace2016**)

The resolution of fine mapping studies could be enhanced by the integration of trans-ethnic fine-mapping meta-analysis, particularly by the inclusion of Yoruba (YRI) and Chinese Han (CHB) descendants 1000 Genomes samples with reduced LD blocks, that can shed light on the true independence of secondary signals (**Kichaev2015**; Bunt et al. 2015). Additionally, inclusion of functional data from publicly available sources such as ENCODE or The Roadmap Epigenomics project, as priors of the approximate Bayesian model demonstrated reduction of the number of SNPs in the credible set and also increased the proportion of successfully fine-mapped loci (**Kichaev2015**; Bunt et al. 2015). These observations were further reinforced by a study integrating fine-mapping data generated by a Bayesian approach known as probabilistic identification of causal SNPs (PICS) with *cis*-regulatory elements for thirty-three immune cell types (**Farh2015**). Interestingly, the top fine-mapped causal variants presented the greatest enrichment (~60%) for enhancer elements, particularly for those in activated cell types and also for DHS and TF binding sites. In the particular case of psoriasis, PICS prioritised SNPs showed enrichment for Th0 naïve CD4⁺ T cells followed by Th1, Th2 and Th17 CD4⁺ subsets. Recently, publicly available tools such as fGWAS and PAINTOR have leveraged cell type-specific annotation to inform the Bayesian analysis and output a further refined credible set of SNPs with functional relevance (**Pickrell2014; Kichaev2015**).

1.4.3 Understanding the epigenetic landscape in complex diseases

Epigenetics modifications, previously defined, are responsible for heritable changes in gene function independent of mutations in the DNA (**Feil2012**). Environmental and intrinsic factors can trigger changes in the epigenome that

Introduction

will result in alteration of gene function through regulation of expression. For example, dietary components such as vitamin B12 intake can results in changes in methylation with locus specific effect (**Wolff1998**). In addition, the genetic background can increase the predisposition to epigenetic changes due to extrinsic factors. Consistently, studies have demonstrated differences in response to environmental factors by different mice breeds as well as greater differences in the epigenetic landscape in human dizygotic when compared to monozygotic twins (**Pogribny2009; Kaminsky2009**).

The disease-associated variants from GWAS studies have been consistently shown to be enriched for regulatory elements tagged by a wide range of epigenetic marks, such DHS, histone modifications and DNA methylation Trynka2013,Trynka2013b,Gusev2014. For example, 76.6% of all non-coding GWAS SNPs together with those in complete LD have been located within broadly defined DHS (Maurano et al. 2012). Moreover, significant enrichment (9.8-fold) in chromatin accessible enhancers, designated by the combination of DHS and a set of histone marks, was reported for combined GWAS SNPs from eleven common diseases (**Gusev2014**).

The plasticity of the epigenetic landscape is determinant for cell differentiation and identity and particularly important in the immune system to ensure adaptation and response to different pathogen infections (**Yosef2016**). The epigenetic landscape is responsible for the regulation of gene expression and its cell type specific effect has been demonstrated in eQTL studies, proving that between 50 to 90% of eQTLs are cell type and stimulus dependent(**Nica2011; Naranbhai2015; Kasela2017**; Dimas et al. 2009; Fairfax et al. 2012; Fairfax et al. 2014; Raj et al. 2014). For example, the variant rs17445836, also a lead SNP in GWAS MS, regulates expression of *IRF8* in monocytes only after two hours of LPS stimulation (Fairfax et al. 2014). Similarly, eQTLs from whole blood have shown only modest overlap with immune enhancers (14%) and immune-mediated

Introduction

GWAS fine-mapped SNPs, highlighting that disease associated variants are also more likely to exert functional effects in a tissue specific manner (**Brown2013; Farh et al. 2014**).

The importance of considering the diversity in the epigenome across cell types when addressing GWAS hits interpretation has increased the efforts to extensively characterise those differences. Varied epigenetic marks including DHS, histone modifications, TF binding and DNA methylation have been interrogated in a wide range of cell types by consortiums such as ENCODE, The Roadmap Epigenomics and Blueprint (**ENCODE2012; Adams2012; Bernstein et al. 2010**).

The integration of those datasets including ,multiple histone marks, transcription factor binding and DHS tracks has provided a more precise insight into the functionality of the genome allowing elaboration of cell type specific chromatin states maps. This has been achieved through development of algorithms such as ChromHMM that uses Hidden Markov Model to segment the genome and and label it with a chromatin state based on concurrence of several epigenetic marks (**Ernst2010; Hoffman2012**). Chromatin segmentation maps have been generated for several cell types, being the most comprehensive the release by The Roadmap Epigenome Project of 111 primary cell and tissue chromatin state maps, which include the definition of eight active and seven repressed states (**Kundaje2015; Ernst et al. 2011**).

The latest methodological advances in the field are enabling a more personalised study and understanding of the epigenome. The development of low cell input and high throughput techniques using NGS to interrogate chromatin accessibility, histone modifications, TF binding and chromatin interaction using as little as few hundred cells is opening the door to map the regulatory landscape in a large number of individuals and cell types (**Buenrostro2013; Schmidl2015; Oudelaar2017**). In the same lines, epigenetic

Introduction

plasticity is also inherent to cell populations and single-cell techniques can help to understand cell-to-cell variability (**Cusanovich2015; Rotem2015; Nagano2013; Smallwood2014**; Buenrostro et al. 2015). Characterising the epigenome at a single cell resolution can help to gain further understanding about disease mechanisms and to interpret the impact of genetic variability in gene regulation.

Since this epigenetic technical revolutions started, systematic studies have been conducted to identify inter- and intra- individual differences and pathological changes in chromatin accessibility and DNA methylation (**Qu2015; Corces2016; Liu2013. Add other ATAC**). Similarly, these new approaches have opened an avenue to ascertain the epigenome profile of clinical samples from different diseases and cell types which will help in the interpretation and understanding of non-coding GWAS variants. In addition to this, personalised epigenomes can also provide insight into disease activity and drug response. For example, differences in DNA methylation of genes responsible for CD4⁺ T cell activation correlated with clinical activity in juvenile idiopathic arthritis and different methylation patterns in RA also explained the failure to respond to DMARD therapy in some patients (**Spreafico2016; Glossop2017**). Overall, the technical feasibility of refining the specificity of the epigenetic maps in a cost effective manner will allow to expand the number of epigenomes available to inform the functional follow up and characterisation of GWAS variants.

1.4.4 The chromatin landscape

In cells nucleus, DNA is compacted into a highly organised structure known as chromatin. The basic repeating unit of chromatin is known as nucleosome. A nucleosome is formed by 147bp segment of DNA wrapped around an octameric core of histone proteins regularly spaced by 10bp of linker DNA (**Luger1997**). In general, highly compacted DNA will remain

Introduction

inaccessible for the assembly of the transcriptional machinery and will prevent gene expression. The chromatin accessibility can be altered by PTM of the histone proteins that will affect interactions between DNA and histones and also between nucleosomes in the vicinity (**Polach2000; Pepenella2014**). Chromatin structure is also influenced by adenosin triphosphate (ATP)-remodelling complexes that facilitate sliding of individual nucleosomes to neighboring DNA segments, increasing temporary chromatin accessibility at particular sites (**Cosma1999**). From the biochemical point of view, the signature of chromatin accessibility, histone modifications, transcription factor occupancy and DNA methylation has enabled the definition and state of DNA *cis*-regulatory elements including promoters, enhancers, silencers, insulators, and locus control regions, amongst others, as previously mentioned in the chromatin segmentation maps (**ref from chomatin segmentation**; Boyle et al. 2012).

Chromatin accessibility

Accessible chromatin constitutes about 1% of the human genome and represents a very robust marker for histone modifications, early replication regions, TSS and TFBS (**ENCODE2007**). The informativeness of accessible chromatin for understanding gene regulation has driven the development of several high-throughput techniques towards accurately tagging these parts of the genome. The "golden standard" technique is DNase-seq, which uses the non-specific doublestrand endonuclease DNase I to preferentially cut on nucleosome-free regions known as DHS sites. Isolation of the chromatin-free material is followed by further enzymatic digestion and DNA library preparation prior to NGS (**John2013**). DNase-seq provides high quality information regarding TFBS, generating footprints that allow to identify TF binding in relation to chromatin structure (**Hesselberth2009; Boyle2010**). Another method to interrogate the accessible genome is formaldehyde-assisted isolation of regulatory elements

Introduction

(FAIR-seq), which uses formaldehyde cross-linking, sonication and phenol-chloroform extraction to remove the DNA-protein complexes and retain only the nucleosome-depleted regions undergoing NGS (**Giresi2006**). Both methods have enabled ENCODE to map the regulatory elements in several cell lines, primary cells and in tissues when abundant collection was possible (**ENCODE2007; Buck2014; Gaulton2010**) Indirect measurement of the chromatin accessibility has been carried out using MNase-seq, which relies on the endo-exonuclease activity of the micrococcal nuclease (MNase) enzyme on cross-linked nuclei to degrade chromatin-free DNA and only retain the nucleosome-bound material for subsequent sequencing (**Axel1975; Ponts2010**). MNase-seq provides a qualitative and quantitative comprehensive map for nucleosome positioning and also TF occupancy. The main disadvantage of these three methods is the high number of cells (ideally 5 to 10 M) required for the assays to provide good quality data, becoming a limitation to apply these strategies to particular biological and clinical samples.

Recently, a new methodology known as assay for transposase-accessible chromatin using sequencing (ATAC-seq) has represented a groundbreaking step in characterising the regulatory landscape of the genome (**Buenrostro2013**). ATAC-seq is based on an engineered hyperactive transposase enzyme, known as Tn5, that preferentially access nucleosome-free and inter-nucleosomes DNA inserting sequencing adapters (Gradman and Reznikoff 2008; Adey et al. 2010). The main advantage of ATAC-seq over DNase-seq are the lower number of cells and the simplicity of the protocol. ATAC-seq requires as little as 500 to 50,000 cells and is a fast two-steps protocol that yields information regarding open chromatin and nucleosome positioning simultaneously. These two aspects make ATAC-seq a very versatile technique to interrogate the chromatin landscape in a clinical set-up, where sample availability and time-efficiency are key factors (**Scharer2016; Qu2015; Qu2017**). Regardless the strength of this new

Introduction

technique, ATAC-seq in many cell types sensitivity is not comparable to DNase-seq and further optimisations of the first released protocol by Buenrostro and colleagues have followed (**Corces2016; Sos2016; Corces2017**).

Histone modifications and TF occupancy

The combination of histone modifications and TF binding to the DNA are essential mechanisms to fully understand transcriptional regulation. Characterising the regulatory elements based on the co-localisation of histone marks in defined combinations is known as the "histone" code (**Jenuwein2001**). Histone modifications take place in the NH₂terminal tail that protrudes from the nucleosome and mediate variation in their interaction with neighboring nucleosomes and the affinity with DNA-binding proteins (**Bannister2011**). Amongst the most common modification are acetylation, phosphorylation and methylation, but others including SUMOylation, ubiquitination and adenosin di-phosphate (ADP) ribosilation have also been found (**Bayarsaihan2011**). These PTM are reversely catalysed histone acetylases and deacetylases (HATs and HDACs), histone kinases and histone methyl transferases and demethylases (HMTs and HDMs), which activity and recruitment gets affected by the surrounding histone modifications (**Bannister2011; Shi2006; Nelson2006**).

Acetylation of histones increases chromatin accessibility due to the negative charges reducing the affinity for the DNA and is associated with active transcription. Conversely, histones methylation is found in both active and repressed chromatin. Overall, the combination of histone modifications can be used to broadly divide chromatin into condensed non-transcribed heterochromatin and accessible transcriptionally active euchromatin. Further studies have identified facultative heterochromatin in genes that are spatially and temporally regulated and constitutive, for those regions which contain permanent silenced genes. Facultative heterochromatin is enriched

Introduction

for H3K27me3 and the polycomb repressor complexes, whilst constitutive heterochromatin is marked by H3K9me3 (**Hansen2008; Bannister2001**). Several types of chromatin corresponding to different regulatory elements have also been distinguished. Enhancer can be reliably tagged by high levels of H3K4me1 whereas promoters are enriched for H3K4me3 and both regulatory features co-localise with H3K4me2 modifications (**Heintzman2007; Hon2009**). Presence of H3K27ac indicates active gene regulation at promoters and enhancers, whereas H3K9ac is only found at active promoters (**Hon2009; Creyghton2010**). Conversely, H3K27me3 together with the heterochromatin mark H3K9me3 inform of gene repression at promoters (**Hansen2008; Bannister2001; Pan2007**). Nevertheless, the functional understanding and interpretation of the different combinations of histone marks on the same genomic region still remains under investigation. GWAS associated variants with different complex diseases have demonstrated to be relatively enriched for some of those modifications, importantly H3K4me3, H3K9ac, H3K79me2, H3K4me1 and H3K36me3 (Ernst et al. 2011; Trynka and Raychaudhuri 2013).

Regulation of transcription by TF is conducted through different mechanisms. TF also play a role, together with histone modifications, in nucleosome positioning as well as in acting as boundary elements to separate chromatin states (**Vierstra2014; Bell2000; Zhang et al. 2009**). TF occupancy is indirectly tagged by chromatin accessibility assays such as DHS mapping through reduced cutting sensitivity of DNase-I due to protein binding. The enrichment of GWAS variants within DHS regions highlights the potential role of many disease-associated SNPs to become pathological by altering TF binding and consequently gene expression.

Chromatin immunoprecipitation sequencing (ChIP-seq) has been widely used in the last few years, since NGS has become more available, in order to precisely locate histone modification and TF binding into the genome.

Introduction

This technique allows assaying protein-DNA binding *in vivo* using Ab that specifically recognise histone modifications or TF following DNA-protein cross-link and DNA sonication. After the immunoprecipitation of the desired DNA-protein complexes using the appropriate Ab, the cross-linking is reversed and the proteins digested prior to DNA library preparation and sequencing (**Solomon1988; Barski2007; Johnson2007**). ChIP-seq has been used to analyse a wide range of histone modifications and TF binding in different cell lines, primary cells and tissues (**ENCODE2012; Adams2012; Bernstein et al. 2010**) Similarly to the first generation of chromatin accessibility techniques, ChIP-seq requires the use of large number of cells ranging between 5 to 10×10^6 cells per experiment, limiting its application to the available biological material. In order to overcome this limitation, a wide range of protocols have been developed of which ChIPmentation (ChIPm) represents the simplest and more cost-effective one, requiring 10,000 and 100,000 cells for histone modifications and TF assay, respectively (**Schmidl2015**). ChIPm has incorporated the use of the Tn5 transposase to simultaneously fragment and add adapters to the immunoprecipitated DNA, accelerating library preparation and increasing the sensitivity of the results. ChIPmentation has been successfully used to identify subtype-specific epigenome signatures in chronic lymphocytic leukaemia (**Rendeiro2015**).

DNA methylation

Together with histone modifications, DNA methylation has a pivotal role in the immune system as the differentiation of haematopoietic stem cells into the different lineages depend on the transition between different transcriptional programs that require genome-wide changes in DNA methylation (**Sellars2015; Lai2013**). DNA methylation involves the transference of a methyl group from the S-adenosyl-L-methionine (SAM) to the 5' carbon in the cytosine

Introduction

of the di-nucleotide CpG by a group of enzymes known as DNA methyl-transferase (DNMTs). CpG islands are found along the entire genome and when methylated they are associated with repression of gene expression (**Herman2003**). Interestingly, DNA methylation is tightly coordinated with histone methylation and the repressive mark H3K9me3 is involved in driving DNA methylation at the same site (**Rottach2009**). From the functional aspect, maturation and activation of immune cells require a complex network of DNMTs that alter the methylation signature in spatial and temporal specific manner, such as in the TNF- α locus upon inflammatory stimuli(**Sullivan2007**). The pathogenecity of changes in the methylome has been studies in a wide range of complex diseases including RA, SLE, psoriasis and PsA, amongst others (**Lei2009; Liu2013; Zhang2010**). For example, an genome-wide study of the methylome in PsA PBMCs revealed a hypomethylated pattern in naïve patients compared to those under MTX treatment (**Kim1996**). The currently most widely used methods are whole-genome bisulfite sequencing (WGBS) and the bead array hybridisation. Both are based the bisulfite treatment of DNA that converts cytosine into uracil (**Frommer1992; Miura2014; Dedeurwaerder2013**). The use of methylome arrays such as the Illumina HumanMethylation450 Bead Chip is the most cost-effective strategy to detect functionally relevant differences in methylation focusing in the main location of regulatory CpG islands such as promoters, 5 UTR, 3 UTR and coding regions as well as miRNA genes.

Chromatin interactions and gene expression

If I look at it as the result of epigenetics. Make a brief summary of the mechanisms and talk about the techniques with special mention to promoterHiC, capture C and capture HiC.

1.4.5 Transcriptional profiles in disease

The role of environmental and genetic factors on altering gene expression in complex diseases has yielded to extensive comparison of case-control transcriptional profiles. The informativeness of approach is also conditioned by the study of the relevant disease tissue, which sometimes remains challenging due to the lack of pathophysiological understanding or the difficulties in accessing it. In immune-mediate diseases, PBMCs differential gene expression (DGE) analysis between patients and controls has enabled to identify causal pathways and biochemical functions in RA, UC, SLE, AS, psoriasis and PsA, amongst others (**Miao2013; Junta2009; Baechler2003; Assassi2010; Batliwala2005**). In addition to this some more cell and tissue specific studies have been performed, targeting synovial synovial macrophages in RA, B cells and monocytes in SLE or uninvolved and lesional skin in psoriasis (**Katschke2001; Dozmorov2015; Jabbari et al. 2012**). The extensive overlap of GWAS variants with non-coding regions potentially dysregulating gene expression has highlighted the role of eQTL studies, previously explained, as a very informative tool to link GWAS variants with particular genes. In this lines, consortium such as the Genotype Tissue Expression (GTEx) have made a great contribution by generating publicly accessible comprehensive tissue-specific eQTL studies that have contributed to the functional understanding of GWAS risk alleles in many complex diseases Londsdale2013,Fagny2017. Lately, eQTL studies have been performed in the context of disease. For example, an eQTL study in five immune relevant cell types isolated from IBD and anti-neutrophil cytoplasmic antibody-associated vasculitis have revealed disease specific eQTLs, some of which disappear following treatment (**Peters2016**).

Introduction

micro-RNAs and long non-coding RNAs

What they are and the importance in disease. Not very specific as I can include more info in the specific intro

Methods to assay gene expression The use of micro-arrays to perform genome-wide expression studies have been replaced by RNA sequencing (RNA-seq) in the last few years as NGS technologies have become more cost-effective. RNA-seq has overcome the use of pre-designed complementary sequence detection probes and is an unbiased method that performs sequencing of all the RNA species present in the sample at the time. Extracted RNA undergoes retro-transcription to cDNA, quantitative PCR (qPCR) amplification preserving relative abundance of each transcript and library preparation prior to sequencing. Systematic comparison of gene expression analysis between micro-arrays and RNA-seq have shown the superior sensitivity of RNA-seq to detect low abundance transcripts (**Zhao2014**). Moreover, the use of RNA-seq allows capturing additional information, importantly identification of new exons, alternative splicing and allele-specific expression (ASE). Alternative splicing refers to the process by which different exons are removed together with introns leading to a variety of mature RNA molecules from the same gene that will consequently result into different protein isoforms (**Pan2008**). Regulation of gene expression through isoform abundance is very common at different tissues and during particular biological processes. For example, RNA-seq isoform quantification has highlighted that differentiation of CD4⁺ T cells into the pro-inflammatory Th-17 is driven by the isoform ROR γ t not ROR γ , both coded by the RORC gene (**Zhao2014**). Several methods have been developed to perform differential exon and isoform quantification with different strengths and limitations in their performance (**Steijger2013; Ding2017**). In individuals heterozygous for a particular exonic SNPs haplotype in a gene, RNA-seq has also allowed to quantify the allele-specific abundance of transcripts, which previously

Introduction

required of additional molecular assays (**Yan2002**). Importantly, ASE has provided direct evidence for local/*cis*-eQTLs to be explained by allele-specific mechanism, showing significant differences in haplotype transcript abundance for up to 88% of the genes with an associated *cis*-eQTLs (**Pickrell2010**).

CAGE

single cell

enabling new applications, such as transcriptome profiling of non-model organisms [10, 11], novel transcripts discovery [12], investigation of RNA editing [13, 14] and quantification of allele-specific gene expression [15].

microarrays have been replaced by RNA-seq that allow also to perform differential splicing analysis CAGE and enhancer RNAs single-cell RNA techniques not in lots of detail because will talk a bit more in the specific section.

Pathway enrichment analysis and gene signature

What is it, how can be performed and what information facilitates This procedure provides a comprehensive understanding of the molecular mechanisms that cause complex diseases

1.4.6 Transcriptional regulation in complex diseases

Non-coding variants at enhancers, silencers and promoters can dysregulate gene expression by altering affinity for transcription factor (TF) binding and driving subsequent alterations in histone modifications and chromatin accessibility (ref). For example, in thyroid autoimmunity, the risk allele of an intronic SNP in the thyroid stimulating hormone receptor (TSHR) gene reduces TSHR protein expression by increasing the affinity of the repressor promyelocytic leukemia zinc finger protein (PLZF), which impairs tolerance to thyroid auto-antigens (Stefan et al. 2014). Alterations in TF binding can also affect looping and long-range chromatin interactions between enhancers and promoter , which

Introduction

for example instance in prostate cancer causes upregulated expression of the oncogene *SOX9* due to increased interaction and enhancer activity (Zhang2012). Alternatively, Non-coding SNP can regulate gene expression by creating a new promoter-like element, as it is in the case of in α - thalassemia where this phenomenon leading to dysregulated downstream activation of all α -like globin genes (Gobbi et al. 2006). Non-coding variants placed at the 3' UTR SNP can affect binding of micro-RNA and lnc-RNA and therefore alter transcript stability and translation. For example there is a Crohns-disease-associated variant at the 3'UTR of the gene immunity related GTPase M *IRGM* that reduces binding of the miR-196 which increases stability of the transcript and enhancing mRNA transcript ultimately resulting in altering alterations in autophagy (Brest et al. 2011). In psoriasis and PsA, some specific SNPs located at 3' UTR of genes such as *IL-23*, *TRAF3IP2* or *SOCS1* have been hypothesised to create or alter miRNA binding sites but there is lack of no experimental evidence has been provided yet (Pivarcsi et al. 2014).

The functional relevance of epigenetic changes in the regulation of gene expression has stressed the relevance of performing epigenome-wide association studies(EWAS), which also allow a more cell type specific approach, instrumental to understand complex diseases. These studies are particularly relevant given the plasticity of the epigenome that would allow using those risk associated changes as potential drug targets, alike genetic variants that are more challenging for alteration. As an example, a DNA methylation EWAS in psoriasis skin samples revealed nine disease-associated differentially methylated sites as result of disease status and environmental factors rather than genetic effects (Zhou2016).

as well as epigenetic QTL studies.

Lastly, transcriptional regulation can also be altered by iintronic SNPs can that lead to also influence mRNA splicing, alterations with examples in the a protective effect of the variant in TNF Receptor Superfamily Member 1A

Introduction

(TNFRSF1A) associated with multiple sclerosis (MS) (Gregory et al. 2012) or an increase of risk to migraine by exon skipping due to an intronic variant in acyl-coenzyme A synthetase 5 (ASCL5) (Matesanz et al. 2016).

SLE variant altering NFkB binding site Trynka 2013

1.4.7 Integration and interpretation of data

Talk about GWAS and transcriptomic profiles in psoriasis GWAS Swindell paper as a way to integrate both Ward and Kellis 2012 paper includes methodology to incorporate this data for the interpretation of genetic variation at a functional level Pathway enrichment analysis Example of chromatin states integration data in psoriasis Tsoi 2017: In psoriasis, integration of chromatin states with GWAS variants have Our results (Fig. 3b; Supplementary Table 7) show that the psoriasis signals are most enriched among enhancers in CD4 T-helper (Th0, Th1 and Th17) and CD8 cytotoxic T cells, in concordance with the previous study

-Tools such as Regulome DB or XGR to integrate functional data with GWAS genetic variants in a more cell type specific manner

1.4.8 Approaches to establish disease mechanisms and causality of genetic variant

massive parallel reporter assays (Inoue and Ahituv 2015)

Wiley and sons 2010 Determining the function of these polymorphisms requires methods which do not necessarily lend themselves to high-throughput techniques, such as those which identified the variants in the first place; and thus, at present, there is a bottle-neck between the polymorphisms with putative functional effects and with those proven effects in vitro.

FUMA method McCarthy <https://www.nature.com/articles/s41467-017-01261-5>

Chapter 2

Material and Methods

2.1 Ethical approval and recruitment of study participants

Sample recruitment for the two different phenotypes and the healthy volunteers were conducted under different ethics.

2.1.1 Psoriasis patient recruitment

Patient blood samples and normal or psoriatic skin biopsies were collected in collaboration with the Dermatology Department research nurses at the Churchill Hospital, Oxford University Hospitals NHS Trust and Professor Graham Ogg at the Weatherall Institute of Molecular Medicine, University of Oxford under approval from the Oxfordshire Research Ethics Committee (REC 09/H0606/71 and 08/H0604/129). After written informed consent, up to 60 mL of blood from eligible psoriasis patients were collected into 10 mL anticoagulant EDTA-containing blood tubes (Vacutainer System, Becton Dickson).

Psoriasis patients were eligible for recruitment when meeting the following criteria:

- over 18 years old
- previously or newly diagnose, in a flare and going into biologic therapy for the first time

Material and Methods

- fulfillment of the clinically accepted Psoriasis Area and Severity Index (PASI) classification for psoriasis diagnosis (Fredriksson and Dermatology 1978)
- moderate to severe disease (PASI \geq 5)
- less than 2 weeks without antibiotics unless used for prophylaxis
- available clinical information and written consent

Detailed clinical information of the psoriasis cohort is included in

2.1.2 PsA patient recruitment

Sample recruitment was performed as part of the Immune Function in Inflammatory Arthritis (IFIA) study established in 2013 (REC/06/Q1606/139) in collaboration with research nurses at the Nuffield Orthopaedic Centre, Oxford University Hospitals NHS Trust and Dr Hussein Al-Mossawi at the Botnar Research Centre. Following informed written consent, blood (30 mL) and synovial fluid aspirate (variable upon disease severity) were recruited into 10 mL anticoagulant sodium heparin coated tubes (Vacutainer System, Becton Dickson).

Eligibility of the PsA patients was upon fulfillment of the following criteria:

- over 18 years old
- previously or newly diagnose, with concomitant psoriasis, in a flare and going into biologic therapy for the first time
- fulfillment of the clinically accepted PsA Response Criteria (PSARC) including a physician global assessment questionnaire (Philipp 2011; Clegg et al. 1996)
- oligoarticular phenotype and naïve for any treatment
- less than 2 weeks without antibiotics unless used for prophylaxis

Material and Methods

- available clinical information and written consent

Further details about the cohort and clinical information can be found in (Chapter ??)(Table ??).

2.1.3 Healthy volunteer recruitment

Recruitment of healthy volunteers was conducted as part of the study Genetic Diversity and Gene Expression in White Blood Cells with approval from the Oxford Research Ethics Committee (REC 06/Q1605/55). Up to 80 mL of blood were collected into 10 mL anticoagulant EDTA-containing blood tubes, similarly to the psoriasis sample recruitment.

The criteria for healthy individuals to participate in the study was:

- over 18 years old and preferably British or European
- no family history of psoriasis, PsA, RA or SpA
- matched sex and age with the psoriasis cohort
- less than 2 weeks since last infectious process
- available clinical information and written consent

2.2 Sample processing

Blood, synovial fluid and skin biopsies were processed straight after recruitment following the appropriate protocols.

2.2.1 PBMC and synovial fluid cells isolation

PBMC were isolated from blood samples through density gradient separation using Ficoll-Paque. Total synovial fluid (SF) cells (SFC) were isolated by centrifugation at 500g for 5 min. Both were washed twice in Hanks balanced

Material and Methods

salt solution without calcium or magnesium (Thermo Fisher Scientific) and resuspended in phosphate saline buffer (PBS, Gibco) supplemented with 0.5% fetal calf serum (FCS, Invitrogen) and 2mM ethylenediaminetetraacetic acid (EDTA, Sigma) prior to cell types separation. Cell numbers and viability were determined by manual count using a haemocytometer and trypan blue (Sigma).

2.2.2 Primary cell isolation using magnetic-activated cell sorting

For the work related to psoriasis and healthy volunteers, primary cell subpopulations were separated using magnetic-activated cell sorting (MACS, Miltenyi). Positive selection was performed for consecutive isolation of CD19⁺ B cells, CD8⁺ T cells, CD14⁺ monocytes and CD4⁺ T cells with AutoMACS Pro (Miltenyi) and cells were manually counted as previously described. MACS separation was chosen over Fluorescence-associated cell sorting (FACS) due to time and logistic constraints in the sample processing and therefore cell numbers in downstream application may not be as exact.

2.2.3 Primary cell isolation using fluorescence-activated cell sorting

Primary cell subpopulations from controls to study the effect of cryopreservation in chromatin states (Chapter 3) or PsA blood and SF samples were isolated by FACS. PBMC and SFC were resuspended in PBS 1mM EDTA (FACS buffer) at 10x10⁶ cells/mL, stained with the appropriate antibody cocktail (Table 2.1) for 30 min at 4°C, washed with FACS buffer and centrifuged at 500g for 5 min at 4°C. For the samples used in Chapter 3, a modified FACS buffer supplemented with 3 mM EDTA, 2% FCS and 25mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, Invitrogen) was used

Material and Methods

to avoid cell clumping after cryopreservation and short recovery. After removing the supernatant, cells were resuspended in FACS buffer prior to separation.

In the controls samples of Chapter 3 only CD14⁺ monocytes and CD3⁺ CD14⁻ CD4⁺ T cells were isolated in the SONY SH800 cell sorter. For the PsA samples, separation of CD19⁺ B cells, memory T cells (CD3⁺ CD14⁻ CD4⁺ CD45RA⁻ and CD3⁺ CD14⁻ CD8⁺ CD45RA⁻) ,CD14⁻ monocytes and CD56⁻ NK was performed using FACS Aria (BD) cell sorter from both PBMC and SFC. Bulk sorted cells were collected in 1.5mL tubes in PBS 1% FCS, whilst single cell and small bulk sorting was performed in 96-well plates in the appropriate buffer (See RNA-seq section). Different nozzle sizes were chosen for bulk and single-cell sorting and OneComp eBeads (eBioscience) were used for compensation of fluorescence spill over.

Surface marker	Fluorochrome PsA/CTL	Manufacturer PsA/CTL	Clone PsA/CTL	Dilution PsA/CTL
Viability	eFluor780	-	eBioscience	1:250
CD3	FITC/AF700	SK7/UCHT1	BioLegend	xxx/1:50
CD4	APC	RPA-T4/RPA-T4	BioLegend	1:50/1:50
CD8a	PE	RPA-T8	BioLegend	xxx
CD45RA	BV421	HI100	BioLegend	xxx
CD19	PerCP-Cy5.5	SJ25C1	BioLegend	xxx
CD14	Pe-Cy7/FITC	M5E2/TUK4	BioLegend/Miltenyi	xxx/1:100
CD56	BV510	NCAM16.2	BD	xxx

Table 2.1: Details regarding target molecule, fluorochrome, clone, supplier and dilution used for PBMC and SFC staining are provided for each of the antibodies in the panel. In controls only CD3, CD4 and CD4 markers were used.

2.2.4 Skin biopsies processing and adherent assay

KC enrichment from skin biopsies was performed as described in Gutowska-Owsiaik and colleagues (GutowskaOwsiaik and Schaupp 2012). Skin biopsies (approximately 4mm) were washed with PBS, cut in 1mm width strips and incubated in 2U/mL of dispase II (Sigma) overnight at 4°C. The epidermis

Material and Methods

was separated from the dermis and either snap-frozen in liquid nitrogen (for RNA extraction) or further digested in trypsin (Invitrogen) at 37°C for 5 min, when used for chromatin accessibility assay. After digestion the resulting cell suspension was filtered through a 70 μ m nylon strainer (BD) and washed with PBS. In some instances cells were manually counted and aliquoted for ATAC-seq processing. In others, cell from each of the biopsies were resuspended in KGM-2 BulletKit (Lonza) supplemented with 0.06mM Ca²⁺ and cultured in a collagen IV coated 96-well plate at 4°C for 10 min or 3 hours, upon experimental requirements (see Chapter X). After culturing, cells were washed twice with 200 μ L of PBS and kept at 37°C for downstream processing.

2.3 Experimental protocols

2.3.1 Cryopreservation and cell culture

For the controls samples in Chapter 3, 40-50x10⁶ of PBMC were freeze-thawing using a modified version of the (Kent 2009) protocol, where cells were pre-conditioned in RPMI 1640 (brand) complete medium supplemented with 2 mM L-glutamine, 100U penicillin and strep 100 μ g/mL and 50% FCS for 30 minutes and afterwards diluted 1 in 2 in complete RPMI 1640 (supplemented as previously described) with 20% dimethyl sulfoxide (DMSO, Sigma). PBMC underwent slow cryopreservation at -80°C in isopropanol at -1°C per minute and stored for a minimum of two weeks in liquid nitrogen. PBMC were thaw, resuspended in supplemented complete RPMI 1640 with 10% FCS at a density of 10⁶ cells/mL and rested for 30 min at 37C, 5% CO₂ in 25mL non-adherent polypropylene cell culture flasks followed by filtering through a 40 μ m to obtain an homogenous cell suspension undergoing FACS separation. Frozen Normal human epidermal keratinocytes (NHEK) in passage 3 were recovered and cultured at a cell density of 5x10⁶ cells/mL in a 75 mL adherent cell

Material and Methods

culture flask (brand) in EpiLife basal medium (Gibco) following manufacturer's instructions. After recovery NHEK were trypsinised at room temperature for 8 minutes followed by trypsin inactivation with EpiLife 10% FCS, centrifugation at 180g for 10 min at room temperature and manual counting with trypan blue. NHEK were seeded in a 96-well plate in 100uL of medium at a cell density of 160 cells/ μ L. NHEK were cultured for 2 days to a 90-100% confluence before being used downstream.

2.3.2 ATAC-seq, Fast-ATAC and Omni-ATAC

Improved versions of the ATAC-seq protocol were progressively used in the thesis for assessment of chromatin accessibility in different primary cells, including CD14 $^{+}$ monocytes, CD4 $^{+}$ and CD8 $^{+}$ T cells, CD19 $^{+}$ B cells and CD56 $^{+}$ NK cells. The subsequent version aimed to reduce the amount of mitochondrial DNA and improve the ratio of signal to noise for this technique.

After MACS separation, primary cells were manually counted as above specified and they were resuspended in PBS with 1% FCS. As previously stated, due to reduced accuracy of manual cell counting compared to FACS sorting, in my experiments ATAC-seq was performed using an estimated number of cells between 50,000 to 100,000. ATAC-seq was performed as described in Buenrostro *et al.*, 2013 with minor modifications. Cells were centrifuged at 500g for 5 min at 4°C. After removing the supernatant cells were lysed for 10 min, the nuclei were transposed using the Nextera Tn5 transposase (Illumina) for 40 min at 37°C and DNA was purified using the PCR MinElute kit (Qigen). Additional modifications and performance in 96-well plates were implemented for KC and they will be described in

After appropriate determination of the amount of DNA amplification using qPCR, samples were amplified and singled indexed for 11 PCR cycles using modified Nextera primers from Buenrostro *et al.*, 2013 (Table 2.2). The resulting

Material and Methods

DNA libraries were purified using the MinElute PCR purification kit (Qiagen) and additional Agencourt AMPure XP Magentic Beads (Beckman Coulter), according to the manual specifications, to remove the remaining adaptors and primer dimers.

Material and Methods

Primer name	Full sequence
Ad1.noMx	AATGATACTGGGACCCAGAGATCTACACTCGTCGGCAGCGTCAGATGT
Ad2.1	CAAGCAGAAAGACGGCATACGGAGATTCGCCCCTAGTCTCTCGTGGCTCGGAGATGT
Ad2.2	CAAGCAGAAAGACGGCATACGGAGATCTAGTACGGTCTCGTGGCTCGGAGATGT
Ad2.3	CAAGCAGAAAGACGGCATACGGAGATTCTGCCTGTCGTTGGCTCGGAGATGT
Ad2.4	CAAGCAGAAAGACGGCATACGGAGATGCTCAGGAGTCTCGTGGCTCGGAGATGT
Ad2.5	CAAGCAGAAAGACGGCATACGGAGATAGGAGTCCGTCCTCGTGGCTCGGAGATGT
Ad2.6	CAAGCAGAAAGACGGCATACGGAGATCATGCTCTAGTCTCGTGGCTCGGAGATGT
Ad2.7	CAAGCAGAAAGACGGCATACGGAGATGTAGAGAGGTCTCGTGGCTCGGAGATGT
Ad2.8	CAAGCAGAAAGACGGCATACGGAGATCCTCTCTGGTCTCGTGGCTCGGAGATGT
Ad2.9	CAAGCAGAAAGACGGCATACGGAGATAGCGTAGGTCTCGTGGCTCGGAGATGT
Ad2.10	CAAGCAGAAAGACGGCATACGGAGATCAGCCCTCGGTCTCGTGGCTCGGAGATGT
Ad2.11	CAAGCAGAAAGACGGCATACGGAGATGCCCTCTACGTTCTCGTGGCTCGGAGATGT
Ad2.12	CAAGCAGAAAGACGGCATACGGAGATTCCTCTACGTTCTCGTGGCTCGGAGATGT
Ad2.13	CAAGCAGAAAGACGGCATACGGAGATCACGACGTTCTCGTGGCTCGGAGATGT
Ad2.14	CAAGCAGAAAGACGGCATACGGAGATACAGTGGTGTCTCGTGGCTCGGAGATGT
Ad2.15	CAAGCAGAAAGACGGCATACGGAGATCAGATCCAGTCTCGTGGCTCGGAGATGT
Ad2.16	CAAGCAGAAAGACGGCATACGGAGATACAAACGGGTCTCGTGGCTCGGAGATGT
Ad2.22	CAAGCAGAAAGACGGCATACGGAGATTGTGACCACTCGTGGCTCGGAGATGT
Ad2.23	CAAGCAGAAAGACGGCATACGGAGATAGGGTCAAGTCTCGTGGCTCGGAGATGT

Table 2.2: Name and full sequence of the PCR primers used for amplification, indexing and pooling of the ATAC-seq and ChIP-seq samples in this thesis. These primers were designed by Buenrostro *et al.*, 2013 and they are an modified version of the Nextera Illumina primers optimised for larger molecular weight DNA fragments from low input samples. All samples were indexed with the universal primer Ad1.noMx and one of the additional 18 primers. The indexing sequence of each of the primers is in blue text.

Material and Methods

Following the Nature Methods publication of Corces et al., 2016, the initial ATAC-seq protocol was replaced by a modified version named Fast-ATAC. It was specifically optimised for hematopoietic cells and combined cell lysis and transposition in a single step. Fast-ATAC was performed as described by Corces et al., 2016 with minor modifications. Since 5,000 cells was considered the lower limit to generate good quality data in Fast-ATAC, in my experiments I used 20,000 MACS or FACS sorted cells, to account for inaccurate manual cell counting as well as possible cell loss over centrifugation steps. The Fast-ATAC reaction was performed for 30 min at 37°C and agitation at 400rpm. DNA was purified as in ATAC-seq and libraries were generated following 13 cycles of PCR amplification, following appropriate cell cycle determination. Purification following PCR were performed using Agencourt AMPure XP Magentic Beads only.

Omni-ATAC, a third generation of ATAC-seq was published by Corces et al., 2017. It consisted in an universal protocol with individual cell lysis and transposition reactions intercalated with a washing step, to remove mitochondrial DNA and other cell debri Omni-ATAC was performed as described by Corces and colleagues (**Corces2017**) using 50,000 cells.

Following either of the three protocols, DNA tagmentation profiles were assessed with the D1000 high sensitivity DNA tape (Agilent) as part of the quality control and quantified using the Kapa kit (Roche), following the manufacturer's instructions. Pools of 12 to 16 libraries were sequenced in one to 3 lanes of the HiSeq4000 Illumina platform by the Oxford Genomics Centre at the Wellcome Trust Centre for Human Genetics to achieve approximately 50 million paired-end reads.

2.3.3 Chromatin Immunoprecipitation with sequencing library preparation by Tn5 transposase

Assessment of histone marks modification in the chromatin of psoriasis patients from Cohort 1B and four age matched healthy volunteers was performed using a low cell input Chromatin Immunoprecipitation (ChIP) method known as ChIPmentation (ChIPm). For each individual and cell type three histone marks, including H3K27ac, H3K4me1 and X were tested in chromatin isolated from 100,000 cells and compared to an input chromatin sample processed in parallel. Samples were processed following the protocol published by Schmidl and colleagues (**Schmidl2015**) with some modifications. Aliquots of 600,000 cells of MACS sorted cell types, as described in 2.2, were fixed with 1% formaldehyde (Sigma) and snap frozen in dry ice and ethanol prior to storage at -80°C. Chromatin sonifications of the different individuals and cell types were performed in one batch using Covaris M220(Covaris). Each of the aliquots was resuspended in 130µL of SDS lysis buffer (Table2.3), sonicated for 8 min using a duty factor of 5% and aliquoted for single ChIPm reactions prior to long term storage at -80°C.

Sonicated chromatin aliquots were thawed and 1.5 volumes of ChIP equilibration buffer (Table2.3) was added in order to reduce the SDS concentration to 0.1% and to achieve the appropriate concentration of NaCl and Triton-X100. For the immunoprecipitation step, samples were incubated with the appropriate amount of antibody (Table 2.4) overnight in rotation at 4°C . Protein-A Dynabeads (Invitrogen) were also washed in beads wash buffer (Table2.3), blocked with yeast tRNA (Ambion) and BSA (NEB) overnight in rotation at 4°C and added to the sample-antibody mix for incubation. After appropriate washes, fragmentation was performed on the bead-antibody bound crosslinked complexes, which prevents overtagmentation of the DNA. This was followed by protein K digest, reverse crosslinking, elution of the DNA and from the beads

Material and Methods

and purification using PCR MinElute kit. The chromatin input control samples were quantified with Qbit after reverse crosslinking and 1ng of chromatin was used for tagmentation.

Amplification by qPCR was done in each of the samples and control inputs to identify the number of full cycles required to reach one-third of the final fluorescence. Libraries were amplified for the number of cycles minus one determined with this strategy to minimise the total number of PCR replicates. The primers used for amplification and indexing were the ones optimised by Buenrostro and colleagues (Table 2.2). The number of amplification cycles for each of the samples is recorded in

2.3.4 RNA extraction and RNA-seq

Bulk RNA-seq

Following MACS isolation of the different cell types from the psoriasis and matched healthy controls, between $2-3 \times 10^6$ cells were resuspended in $350\mu L$ of RNAProtect (Qiagen) in *mercaptoethanol* (BM, Sigma) and snap frozen in dry ice before storage at -80°C . Cells isolated from psoriasis and control Cohort 1A (Chapter ??) were preserved in RNAProtect, which stops any biochemical reaction and transcriptional activity maintaining cell integrity. At early stages of the project, when I was uncertain of time frames to process the different material from the acquires samples, I decided to use RNAProtect to preserve cells for future RNA extraction to guarantee high quality in case storage exceeded 6 months. In the psoriasis and control samples from Cohort 1B, cells were resuspended in 0.1% BM supplemented RLT buffer, which lysates cells and prevents RNA degradation. Cell lysates were homogenised using the QIAshredder (Qiagen) prior to RNA extraction. When starting from RNAProtect preserved material, cells were centrifuged at 300g for 10 min at room temperature, the supernatant were removed and the pellets were

Material and Methods

Reagent	Final concentration
SDS lysis buffer	
SDS	0.25%
Sigma	
EDTA	1mM
X	
Tris-HCl pH 8	10mM
Sigma	
PI	1X
Roche	
Water	-
ChIP equilibration buffer	
Triton-X100	1.66%
EDTA	1mM
NaCl	233mM
Tris-HCl pH 8	10mM
PI	1X
Water	-
Beads washing buffer	
SDS	0.1%
EDTA	50mM
NaCl	150mM
NP-40	1%
Tris-HCl pH 8	10mM
PI	1X
Water	-
ChIP buffer	
SDS	0.1%
Triton-X100	1%
EDTA	1mM
NaCl	140mM
Tris-HCl pH 8	10mM
PI	1X
Water	-

Table 2.3: Composition of the three modified buffers in house for the ChIPm protocol: SDS lysis buffer, ChIP equilibration buffer, beads washing buffer and ChIP buffer. For each of the buffers the reagents, composition and supplier are indicated. The final volume prepared for each buffer was adjusted depending on the number of samples processed at the time. Sodium dodecyl sulfate (SDS), PI (proteinase inhibitor), Supplier for each of the reagents as follows: SDS (Sigma), EDTA(xxx), Tris-HCl pH8 (xx), Triton-X100 (xxx), NP-40 (Sigma) NaCl(xx), PI (Roche) and water (Ambion)

Material and Methods

Histone mark	Feature	µg per sample	Manufacturer
H3K27ac	Active enhancer, promoter	2	Diagenode (C1)
H3K4me1	Enhancer	1	Diagenode (C1)
H3K4me3	Active promoter, enhancer	1	Diagenode (C1)

Table 2.4: Details regarding the histone marks, the the most likely chromatin state delineated, the amount of antibody required per reaction and the supplier and catalog num of the antibodies.

resuspended in 350µL of RLT 0.1% BM buffer, prior to homogenisation with QIAshredder.

Total RNA were extracted using the AllPrep DNA/mRNA/microRNA Universal kit (Qiagen) following the manufacturer's instructions. RNA extractions were performed in batches of 12 samples, including all cell types from each individual and a balanced numbers of psoriasis and control samples, to minimise batch effect correlation with phenotype groups. Basic quantification was performed with NanoDrop (Thermo Scientific) before storage at -80°C.

RNA-seq quality control (QC), quantification, library preparation and sequencing were carried out by Oxford Genomics Centre at the Wellcome Trust Centre for Human Genetics in two independent batches of samples, each including Cohort 1A or Cohort 1B, respectively. Processing of samples in two batches was due to logistics of patients recruitment in the project. Quality control and quantification were assessed with the Bioanalyzer (Agilent). Preparation of RNA-seq libraries was performed using Ribo-Zero rRNA Removal kit (Illumina), based on ribosomal RNA depletion. Unlike strategies using polyadenylated transcripts selection, this method allowed to preserve non-polyadenylated transcripts including nascent pre-mRNA (unspliced) and functionally relevant lncRNAs. For each of the cohorts, all libraries were pooled together and sequenced over several lanes of HiSeq4000 to a depth of 50 million total reads

Material and Methods

per sample in order to maintain an appropriate level of sensitivity for subsequent expression analysis given the greater complexity of these libraries.

PsA memory CD4⁺ and CD8⁺ single-cell RNA-seq

Single CD4⁺ or CD8⁺ memory T cells were sorted in 2µL of cell lysis buffer into 96-well plates. Four and three plates were prepared for CD4⁺ and CD8⁺, respectively, including wells containing 50 cells in each of the plates for control purposes. Libraries were prepared, indexed and sequenced at the Sanger Institute (Cambridge) using the Smart-seq2 methodology as described by Picelli and colleagues (Picelli *et al.* 2014). Samples were sequenced using 50-bp single-end Illumina HiSeq2500 at a depth of approximately 4 million reads per cell.

scRNA-seq was also generated in FACS sorted 1:1 ratio CD4⁺ or CD8⁺ memory T cells and in bulk PBMC using 10X Genomics technology Chromium single cell 3' and 5' expression library prep kits (PN-120267 and PN-1000014, respectively) by the Oxford Genomics Centre at the Wellcome Trust Centre for Human Genetics. Libraries were sequenced with Illumina HiSeq4000 xxxx bp single-end at a depth of 50,000 reads per cell.

Small-bulk RNA-seq

Between 100 to 500 cells of the five populations isolated from PsA patients were FACS sorted into 2µL of cell lysis buffer and processed for library prep as in Picelli *et al.*, 2014 by the Oxford Genomics Centre at the Wellcome Trust Centre for Human Genetics. Libraries were sequenced with Illumina HiSeq4000 xxxx bp single-end at a depth of XXX reads per cell.

2.3.5 Single-cell analysis of the V(D)J T cell receptor repertoire

Single-cell sequencing of the V(D)J segments from TCR transcripts was performed simultaneously with the 10X Genomics technology Chromium 5'

Material and Methods

expression library kit (PN-1000014) by the Oxford Genomics Centre at the Wellcome Trust Centre for Human Genetics. In short, full-length cDNA was amplified by PCR with primers against the 5 and 3 ends of the barcode sequences inherent to the 10X Genomics bead technology. The amplified material was divided for use in 5' total scRNA-seq (as specified previously in 2.3.4) and also for enrichment of the TCR by PCR amplification with specific primers. TCR enriched cDNA was followed by enzymatic fragmentation and size selection in order to obtain variable length fragments spanning the V(D)J segments. Library prep and indexing was followed by Illumina HiSeq4000 xxxx bp single-end sequencing at a depth of XXX reads per cell.

2.3.6 DNA extraction and SNP genotyping

DNA isolation was performed using the AllPrep DNA/mRNA/microRNA Universal kit (Qiagen) following the manufacturer's instructions. Basic quantification was performed using NanoDrop (Thermo Scientific) and kept at -80°C for long term storage. For each of the patients and controls, the extracted DNA was amplify by PCR using forward (5'-CACTGTGGAGGGAGGAACAA-3') and reverse (5'-CGTGTGGCCAGGATAGTCT-3') primers annealing up and down stream the SNP rs4672505, respectively. The 390 bp PCR product was purified using MinElute PCR purification kit, quantified by dsDNA Qbit kit (Invitrogen) according to the manufacturer's instructions and prepared for Sanger sequencing using the Mix2Seq service (Eurofins). The forward and reverse sequences were analysed with BioEdit software.

2.3.7 Mass cytometry analysis

Mass cytometry analysis was performed by Dr Nicole Yager in collaboration with UCB following their in-house protocol. Briefly, PBMCs and SFMCs were split in three ways for 5 min fixation with 1.4% paraformaldehyde, fixation

Material and Methods

and 6h treatment with monensin (MN) and brefeldin A (BFA) or fixation and 4h treatment with ionomycin and phorbol-12-myristate-13-acetate (PMA). All samples were treated with cisplatin before fixation to facilitate discrimination of dead cells in the staining. MN and BFA are both inhibitors of protein transport that prevent the release of cytokines from the cells and allow measuring the intrinsic cytokine production in basal conditions whereas treatment with ionomycin and PMA induces unspecific activation of immune cells. After appropriate treatment, cell suspensions were washed with PBS and stained with a phenotyping panel of forty four cell surface markers (including viability) or permeabilised and stained with the intra-cellular staining (ICS) panel formed by fourty three markers including surface and intracellular targets Table xxxx. Data analysis was conducted using xxxxxx.

Phenotyping panel

CD248, CD19, GP38, FAP, CD4, CD8a, IL8, CD16, CD25, IL4, CD123, IL-21, FcεR, IL-17F, IL-2, TNF α , IL-17A, IL-10, CD11c, CD14, CD161, IL6, IFN- γ , GM-CSF, FCeR, CD44, IL-17FF, IL-17AF, CD3, CD45RO, CD-11b, CD56, HLA-DR, IL-13, CD117, CD45	CD25, C
---	---------

Table 2.5: Molecules targeted by the phenotyping and cytokine production antibody staining panel in PBMCs and SFMCs.

2.4 Computational and statistical analysis

2.4.1 ATAC-seq data analysis

ATAC-seq, Fast-ATAC and Omni-ATAC data were analysed using an in house pipeline towards which development I made an important contribution. The pipeline performs single sample data processing and it also builds a

Material and Methods

combined master list for each of the comparisons of interest to later perform differential analysis.

Next generation sequencing data analysis

NGS data for each of the samples was trimmed for low quality base pairs and Nextera adapter sequences using cutadapt () before general QC assessment using fastqc (). Trimmed data was alignment to the reference genome built hg19 using bowtie2 (**Langmead2006**) and the following parameters -k 4 -X 2000 -I 38 –mm -1, consistently with other publications (**Buenrostro2013; Corces2016**). Samtools () was used to remove PCR duplicates reads previously marked with Picard Tools () as well as low MAPQ (<30) non-uniquely and non-properly paired reads. The resulting bam file was additional filtered to remove mitochondrial DNA and reads were adjusted by +4 bp in the plus strand and by -5 bp in the minus strand to represent the center of the transposition binding event. Pileup tracks (bigWig files) for visualisation were generated using bedtools genomeCoverageBed where each value represents number of reads per position () and the UCSC genome browser bedGraphToBigWig tool ().

Peak calling and filtering

Peak calling was performed using MACS2 callpeak () and the parametres –nomodel –shift -100 –extsize 200 –p 0.1 –keep-dup all –call-summits and filtered for those overlapping with the blacklisted features from ENCODE project (). The –shift and –extsize parameters were set according to the recommendations of MACS2 for DHS and following other ATAC-seq publications (**Corces2016; Buenrostro et al. 2015**). The pval cut off for peak filtering was determine for each of the cell types separately. For a particular cell type, the bam file of each sample (patients and controls) was partitioned in two (pseudoreplicates) and peak calling was performed in each of them followed by Irreproducibility

Material and Methods

Discovery Rate (IDR) analysis to assess the percentage of peaks sharing IDR at each corresponding pval. Median of the pval showing the greater percentage across all the samples was used to filter each peak list. Peak summits were replaced by the median of the summits for those with multiple called summits and extended +/- 250bp to create a non-overlapping homogenous 500bp peak list (**Corces2016**; Buenrostro et al. 2015).

Sample quality was determined by the fold-change enrichment of ATAC-seq signal across all the TSS identified by Ensembl, since chromatin is expected to be more accessible at the sites of transcriptional initiation compared to the flanking regions.

Combined peak master list and differential analysis

To perform differential open chromatin analysis a non-overlapping 500bp peak master list was built for each cell type separately by union of all the peaks that were present in at least 30% of the samples, regardless patient and control group. Reads overlapping each of those peaks were retrieved for each sample using HTSeq-count algorithm () to build a combined count matrix. An empirical 95% confidence cut-off for the number of counts in absent peaks was calculated on the raw count matrix and used to filter out some peaks in the master list before proceeding to differential analysis ((**Xinmin2005**; Jonker et al. 2014)). Differential analysis was performed using DESeq2 R packaged taking into account paired samples for the PsA analysis or correcting for covariates for the case-control psoriasis analysis (which covariates???) (**Love2014**). A combined master list for all cell types and tissues (when applicable) was also built following the same strategy, normalised and used for principal component analysis (PCA).

Material and Methods

2.4.2 ChIPm data analysis

Next generation sequencing data analysis

ChIPm NGS data from samples and inputs were processed similarly to ATAC-seq ([??ATAC_analysis](#)) for trimming, mapping and filtering with minormodifications. *Parmsubtract* and *bedGraphToBigWig* tools.

Peak calling and filtering

Peak calling for each sample was performed accounting for the background in the input samples using MACS2 callpeak –bw 200 –p 0.1 –keep-dup all –call-summits. In this case the average library fragment size(–bw) was used by MACS2 to first empirically find the model that better represent the precise protein-DNA interaction and then calling peaks with statistical confidence (pval) using the input sample to calculate the local background. Filtering and down stream peak homogenisation was performed similarly to ATAC-seq analysis ([??ATAC_analysis](#)).

Sample quality for H3K27ac and H3K4me3 ChIPm samples was determined by the fold-change enrichment of the histone mark across all the TSS identified by Ensembl, since both marks are enriched at promoters.

Combined peak master list and differential analysis

To perform differential histone modification analysis a peak master list was built and counts at each of the locations were performed as described for ATAC-seq ([??ATAC_analysis](#)). PCA analysis was performed for samples from all cell types and individuals.

Differential analysis

2.4.3 RNA-seq data analysis

Bulk RNA-seq analysis

The ribo-depleted RNA-seq data generated was mapped using the aligner STAR (**Dobin2013**) against the Gencode hg19 annotation file containing reference chromosomes, scaffolds and assembly patches. The annotation file comprised 2,840,283 gene entities, including lncRNAs. Mapping allowed multiple alignments and only retained those with the best score and a mismatch percentage lower than 0.04%. Duplicates were marked and removed using Picard Tools. The filtered de-duplicated and sorted bam files were used to retrieve counts at each of the genes location of the annotation file using HTSeq-count. Differential gene expression analysis was performed with DESeq2 R package filtering parameters

Single-cell RNA-seq analysis

2.4.4 Genomic region annotation, pathway enrichment analysis and visualisation

Annotation of genomic regions and signalling-pathways visualisation were performed with two functionalities of the R package and web-app Atlas and Analysis of systems-biology-led pathways, developed by Dr. Hai Fang and towards which I have made some contribution (manuscript in preparation). ATAC-seq and ChIPm peaks were annotated with gene entities based on publicly available promoter-HiC data in 17 immune cell types (“**parencite**”) Javierre2016. The interactions were weighted based on the number of cell types for which the same bait-target interaction was reported as well as on the confidence of each of those interaction measured by the CHiCAGO score. This approach integrates the knowledge in the field about regulatory regions affecting the expression of distal genes and was not biased by the physical vicinity to a gene when

Material and Methods

annotating genomic intervals located at intergenic regions or gene deserts. For visualisation, manually curated KEGG pathways including all the genes for each gene family were colored based on the fold-change from the corresponding differential analysis and highlighted in bold when passing the FDR threshold for significance.

2.4.5 Enrichment analysis for genomic annotation features

-Includes the ATAC all peaks enrichment for eQTLs -ATAC all peaks enrichment for GWAS -ATAC for TFBS, chromatin annotation segments

2.4.6 Statistical fine-mapping

Chapter 3

Establishment of laboratory methods and analytical tools to assess genome- wide chromatin accessibility in clinical samples

3.1 Introduction

**Previous and current methods to identify the accessible genome
in cells and tissues**

Implementation of ATAC-seq to define the chromatin landscape

Technical limitations and recent advances in optimisation

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4473780/>

Talk about ATAC being more variable, a native chromatin accessibility assessment without cross-linking. Role of transposase ability in accessing the chromatin, debris and DNA from dead cells adding noise

Paper to justify peak calling: A comparison of peak callers used for DNase-Seq data.

New ATAC but also explanations of the limitations: Characterization of chromatin accessibility with a transposome hypersensitive sites sequencing (THS-seq) assay

Challenges of working with clinical samples

3.2 Results

3.2.1 Establishment of an ATAC-seq data analysis pipeline based on current knowledge

When the first ATAC-seq publication (**Buenrostro2013**) appeared, there were not well established protocols for the complete processing of the data. Since then, several publications have used ATAC-seq and modifications of this protocol together with a wide range of data analysis strategies to answer different biological questions (Table 3.1). There are several limiting aspects in the process of analysing ATAC-seq data, including QC assessment, peak calling/filtering and differential analysis of chromatin accessibility regions between groups. Using the current knowledge in the field as well as on my own analysis, I agreed on the most appropriate criteria and parameters to implement in our in-house pipeline. For this purpose, I used ATAC data generated with the first protocol (**Buenrostro2013**) in paired CD14⁺ monocytes and CD4⁺ total T cells from the same three healthy individuals, all of them downsampled to 30 million of reads, in order to facilitate the comparison across all of them.

Table 3.1: .

Publication	Peak calling and filtering	Master list	Differential analysis
Corces <i>et al.</i> , 2016	MACS2 (-nomodel), summit extension +/-250bp, rank summits by pval	Maximally significant overlapping peaks	Quantile non- unsupervised normalisation and hierarchical clustering.
ENCODE	MACS2 -nomodel, pairwise IDR analysis, filtering IDR<10%	Choosing longest pairwise	NA
Turner <i>et al.</i> , 2018	MACS2 (-nomodel -q 0.01)	Merging all filtered called peaks from the different cell types.	De novo:DiffReps with fragment size 50bp.

Alasoo <i>et al.</i> , 2018	MACS2 (-nomodel -shift -25 -extsize 50 -q 0.01	Merge of peaks from all conditions present in at least three samples of the same condition.	Peak based: TMM normalisation and lima voom (FDR<0.01).
Qu <i>et al.</i> , 2017	ZINBA PP>0.99.	Merging of filtered peaks from each individual sample.	Quantile normalisation and peak based in house Pearson correlation method.
Rendeiro <i>et al.</i> , 2016	MACS2 (-nomodel -extsize 147)	Merge of peaks from all samples in an iterative process including permutations	Peak based: quantile normalisation and Fisher exact test (FDR<0.05).
Schareret <i>et al.</i> , 2016	HOMER (-style dnase)	Merge of all overlapping peaks between all samples using HOMER mergePeaks	Peak based: TMM normalisation and edgeR package (FDR<0.05).

Sample quality control

Regarding QC measurements, the variability in performance of the methodology, particularly ATAC-seq and Fast-ATAC, has required to agree on appropriate parameters to determine the quality of the samples before proceeding with downstream differential analysis. After reviewing the different read-outs implemented across different publications as well as the recently ENCODE update, I have identified the most informative ones showing supporting correlation between them.

Firstly, I analysed the fragment size distribution for each of the six samples in order to determine if they recapitulated the expected periodicity of nucleosomes protecting the DNA during the transposition event (Figure 3.1a). All the samples showed periodicity every ~200bp up to 600bp, clearly distinguishing chromatin organisation into mono-, di- and tri-nucleosomes. The relative intensity of nucleosome-free DNA fragments (<~147pb) compared to nucleosome-bound DNA was greater for some of the samples (e.g CTL1 CD4⁺ and CD14⁺) and similar or lower for others (e.g CTL3 CD4⁺ and CD14⁺). Nucleosome-free fragments(peak<~147bp) are also clearly distinguished in all of the samples, meeting the ENCODE QC recommendations (**ENCODE**).

Another QC measurement was the enrichment of ATAC-seq signal over a random background of reads across all the TSS identified for Ensemble genes (Figure 3.1b). It is well established that nucleosome repositioning and an increase in chromatin accessibility take place at TSS to allow formation of the transcriptional machinery and initiation of transcription. Fold-enrichment signals ranged between 5-7 for the CD4⁺ samples and they were much higher(between 17-20) for the CD14⁺ samples. The lower sample quality of the CD4⁺ compared to CD14⁺ shown by the TSS signal were recapitulated by the ATAC-seq signal at the promoter of the constitutively expressed gene glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) (Figure 3.1c).

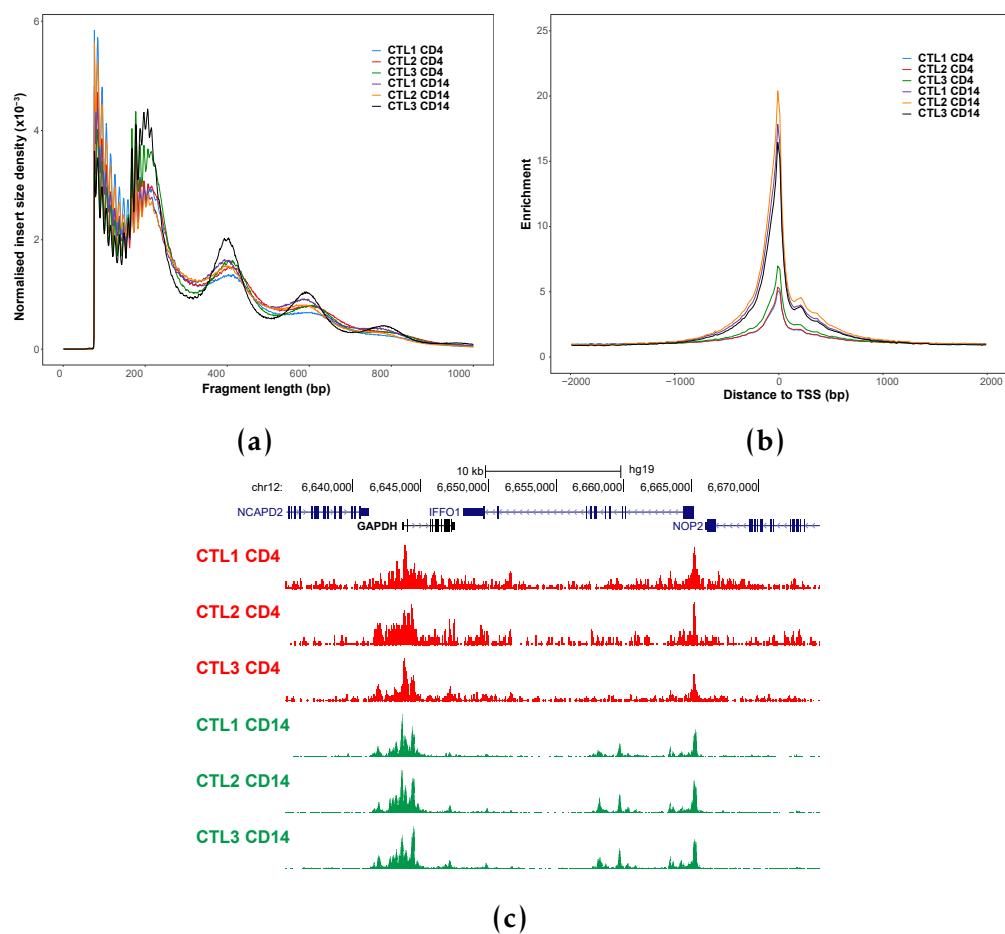


Figure 3.1: Measurements for quality control assessment in ATAC-seq samples

As part of the QC assessment I looked at the percentage of mitochondrial reads and the fraction of reads in peaks (FRiP)(Table 3.2).

Sample	% MT reads	Fraction of reads in peaks
CTL1 CD4	14.9	9.8
CTL2 CD4	30.5	11.2
CTL3 CD4	28.8	11.6
CTL1 CD14	43.3	32.2
CTL2 CD14	36.8	57.0
CTL3 CD14	37.6	49.9

Table 3.2:

FRiP score is a way of assessing the background signal in different types of assays that are based on peak calling, including ChIP-seq. Positive correlation between the TSS fold-change enrichment and FRiP was observed (data not shown), being both appropriate inter-dependent QC measures to evaluate sample noise. Regarding TSS and FRiP cut-off values, Alsooo *et al.*, 2018 and, recently, ENCODE have recommended minimum FRiP between 10-20% and TSS between 6-10. ENCODE has prioritised the use of TSS over FRiP as the measurement to determine the noise in the sample (**ENCODE**). According to this recommendations all these samples passed QC; however clear differences were seen between CD4⁺ and CD14⁺ samples. The mitochondrial content ranged between 14.9-43.3% and, alike FRiP and TSS, it was higher in CD14⁺ than in CD4⁺ and not directly related with any of the other QC measurements.

Peak calling and filtering

As part of the ATAC-seq pipeline implementation, peak calling and the criteria for filtering where another two aspects to determine. Although different peak callers have been used, most of the publications as well as ENCODE has been using MACS2 as the preferred methodology (Table 3.1). MACS2 has been initially developed for ChIP but it has also been used for DHS and ATAC-seq

with disabling the model and agreeing in an extension size (`-extsize`) and a shift (`-shift`), which indicate the direction and number of bp for reads to be shifted and the number of bp for them to be extended, respectively. The `-extsize` should correspond to the average fragment size, which in my libraries is \sim 200bp and the `-shift` is set to -100, as it is recommended to be set to $-1/2$ of the fragment size for chromatin accessibility assays. This parameter could be further optimised but it escapes from the aim of this thesis.

I was interested in understanding the effect of sequencing depth and the sample quality on the peak calling to have a better control of both variables in the downstream analysis. I performed random read sub-sampling every 5M total reads (from 5M to 30M) followed by peak calling with arbitrary filtering for FDR<0.01 in each of the six aforementioned samples.

The number of called peaks passing filtering showed an steady increased over the read depth which seemed to reach a *plateau* around 25M reads (Figure 3.2a). This was consistent with the decay in the increments of called peaks over read depth, almost invariable, from 20M reads onwards (Figure 3.2b). Moreover, lower number of peaks were detected in CD4 $^{+}$ samples compared to CD14 $^{+}$ highlighting the influence of sample quality on the total number of called peaks. Interestingly, sample quality measured by FRiP reflected very low changes over read depth and was stable from 15M reads for all six samples (Figure 3.2c). Overall, this confirmed that measurement of sample quality by FRiP or TSS is not biased by sequencing depth.

Regarding peak calling filtering, most of the ATAC-seq publications using MACS2 have arbitrarily used an FDR<0.01 (Table 3.1). In collaboration with Dr. Gabriele Migliorini and following ENCODE pipeline, we explored the use of IDR to experimentally identify the most appropriate p-val for filtering each individual sample. Each sample was partitioned in two, peaks were called in each half and the percentage of peaks (over the total number shared peaks)

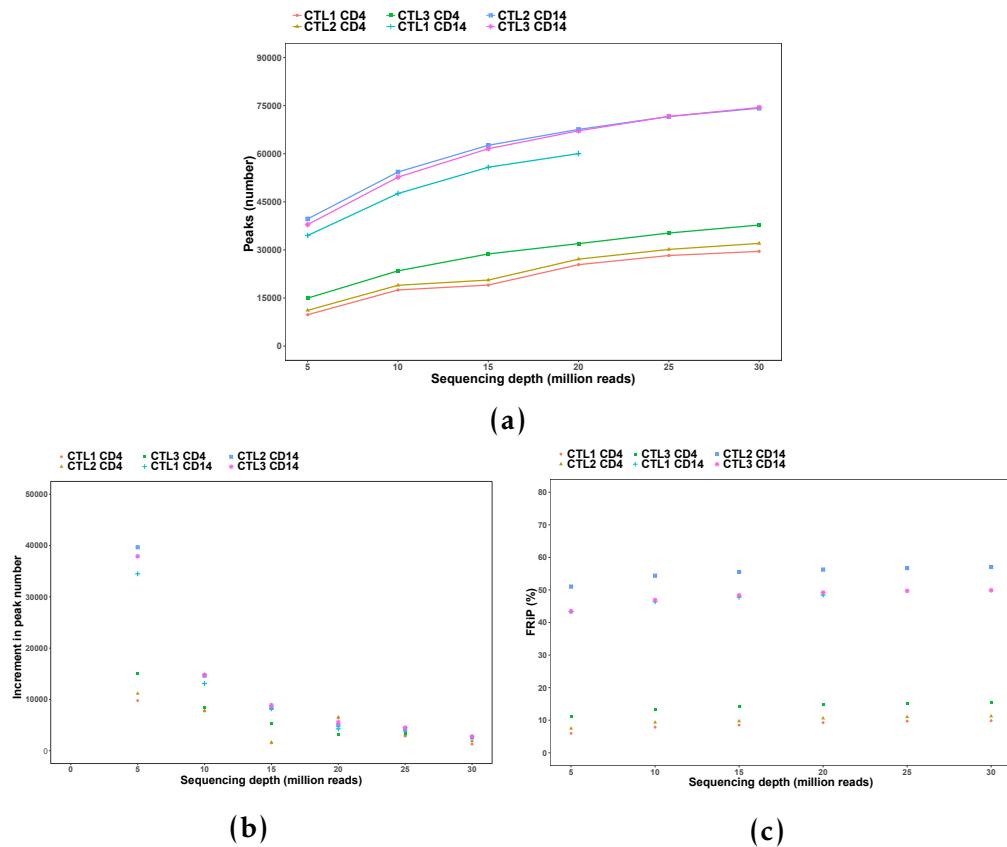


Figure 3.2: Peak calling at different sequencing depth in ATAC-seq samples

sharing IDR at a particular p-val was calculated (Figure 3.3 a and b). Both of the representative samples showed variation in the percentage of shared peaks upon sequencing depth under 10M reads, being the effect more pronounced and extended in the lower quality (CTL2 CD4⁺ Figure 3.3 a) compared to the counterpart CD14⁺ (Figure 3.3 b). The shape of the curves was also influenced by the sample quality, presenting a smoother profile reaching a single maximum percentage of shared IDR peaks for samples with TSS enrichment >~10 compared to samples with lower quality. All the CD14⁺ samples reached the maximum percentage of IDR shared peaks at approximately -log₁₀ pval 8 (data not shown). Filtering the CD4⁺ peaks at the -log₁₀ pval of the first maximum of IDR shared peaks reduced the percentage of peaks overlapping noise (e.g heterochromatin, repetitive sequences and repressed regions) when compared to peaks filtered based on FDR<0.01 (Figure 3.3 b). In summary, this

Establishment of methods to assess genome-wide chromatin accessibility

IDR analysis appeared as systematic method to identify an optimum p-val to perform individual filtering in a sample-specific manner and in a less arbitrary way than the extended 1% FDR.

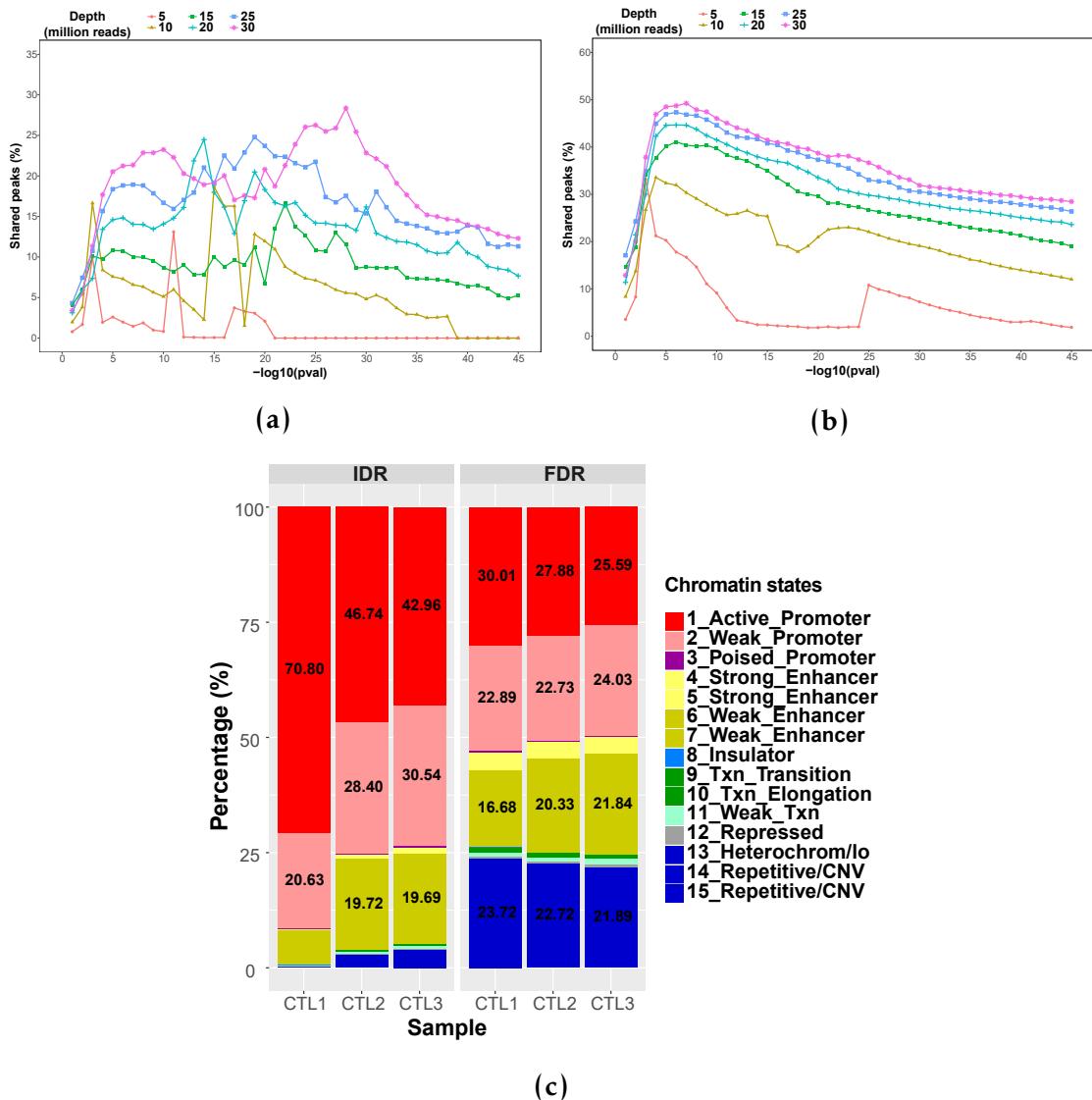


Figure 3.3: Peak calling filtering using IDR analysis in ATAC-seq samples

Differential chromatin accessibility analysis

From the methods that can be used to perform differential chromatin accessibility analysis (Table 3.1), I chose a peak-based approach where a consensus master list between all samples was built and the number of reads overlapping the master list peaks were retrieved for each sample. As previously

mentioned in the Chapter the master list was composed of non-overlapping 500bp with peaks present in at least 30% of the samples, regardless the group they belonged to (e.g patients or controls). One of the main limitations of the ATAC-seq and FAST-ATAC protocols (discussed in the next section) is the background signal. Therefore, it was calculation of an empirical cut-off, similarly to the strategy use in micro-array technology, was performed to minimise the impact of background read counts on the differential analysis (Xinmin2005; Jonker et al. 2014). Moreover, due to the lack of consistency found across the ATAC-seq publications, two methods for normalisation/differential analysis were assayed.

From the count matrix of the same six samples as before, the combined distribution of read density from all the absent peaks in each sample was used to define a sequence of twenty cut-offs. These cut-offs corresponded to the number of counts showed by a particular percentage of absent peaks (supplementary info). Each cut-off was used to filter out from the raw count matrix those peaks from the master list for which the number of counts was \leq than that particular cut-off in more than three samples (being three the number of the smallest group of replicates in this particular experimental design). Quantile normalisation followed by differential analysis with limma voom showed greater number of differential open chromatin regions (DOCs) at an FDR <0.01 compared to DESeq2 across all the cut-offs (Figure 3.4 a). The two approaches presented progressive decrease in the number of DOC sites from the 75% cut-off. Conversely, the proportion of DOC calculated over the total number of regions considered in the differential analysis for each cut off significantly increased from the 50% cut-off onwards, indicating a progressive reduction in the false positive hits reported 3.4 b). From this analysis, 80% was chosen as a conservative filtering cut-off for which almost all the 19,855 DOCs identified by the most conservative method

(DESeq2) at an FDR<0.01 were recapitulated by limma voom at the same FDR (Figure 3.4 c).

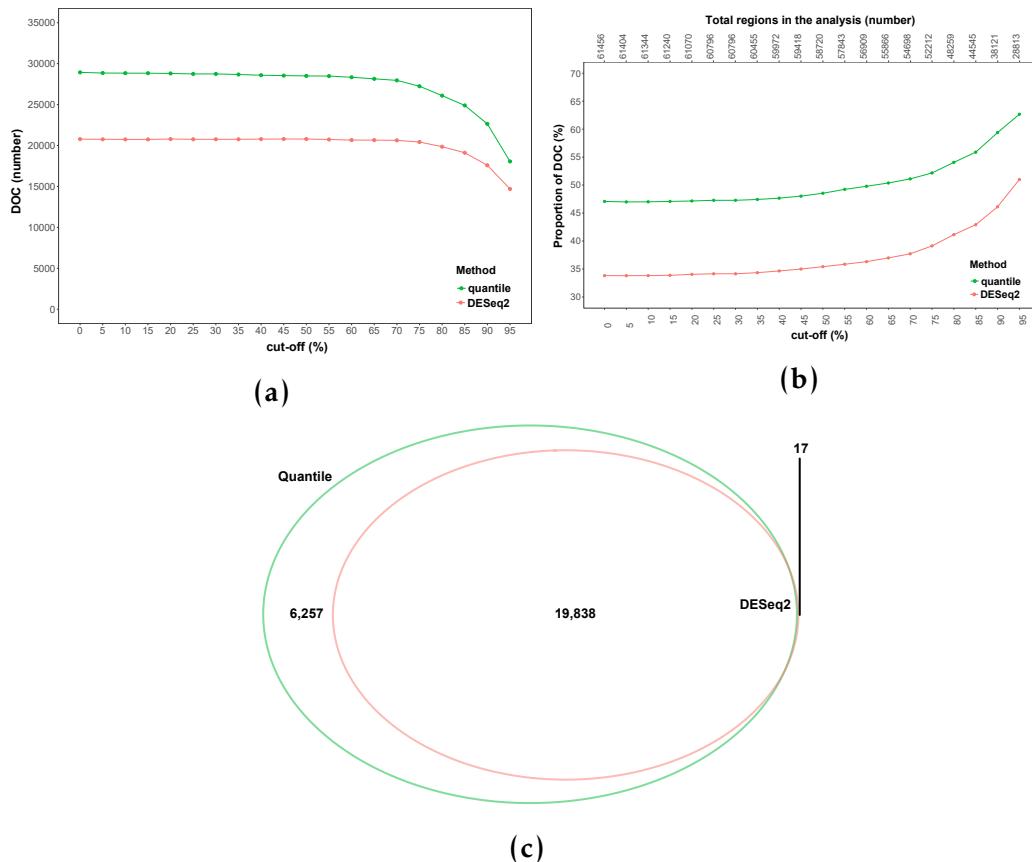


Figure 3.4: Differential chromatin accessibility analysis using limma voom and DESeq2.

Both methods performed appropriate normalisation of the counts at each of the master list peaks across the six samples, being the median of the quantile normalisation slightly more consistent across the two cell types compared to DESeq2 (Figure 3.5 a). When looking at the first FDR ranked 19,855 limma voom DOCs, 18,768 of them were the same as the retrieved by DESeq2. Moreover, very significant positive correlation was found between the fold changed of those 18,768 significant DOCs in both differential analysis methods ($r^2=0.999$, $p\text{-val}=2.2^{-16}$) (Figure 3.5 b). These observations suggested that the differences in the number of FDR significant DOCs reported by each of the methods could be partly due to differences in the way of calculating the false-positive rate.

Clustering and heat map and pathway analysis-briefly

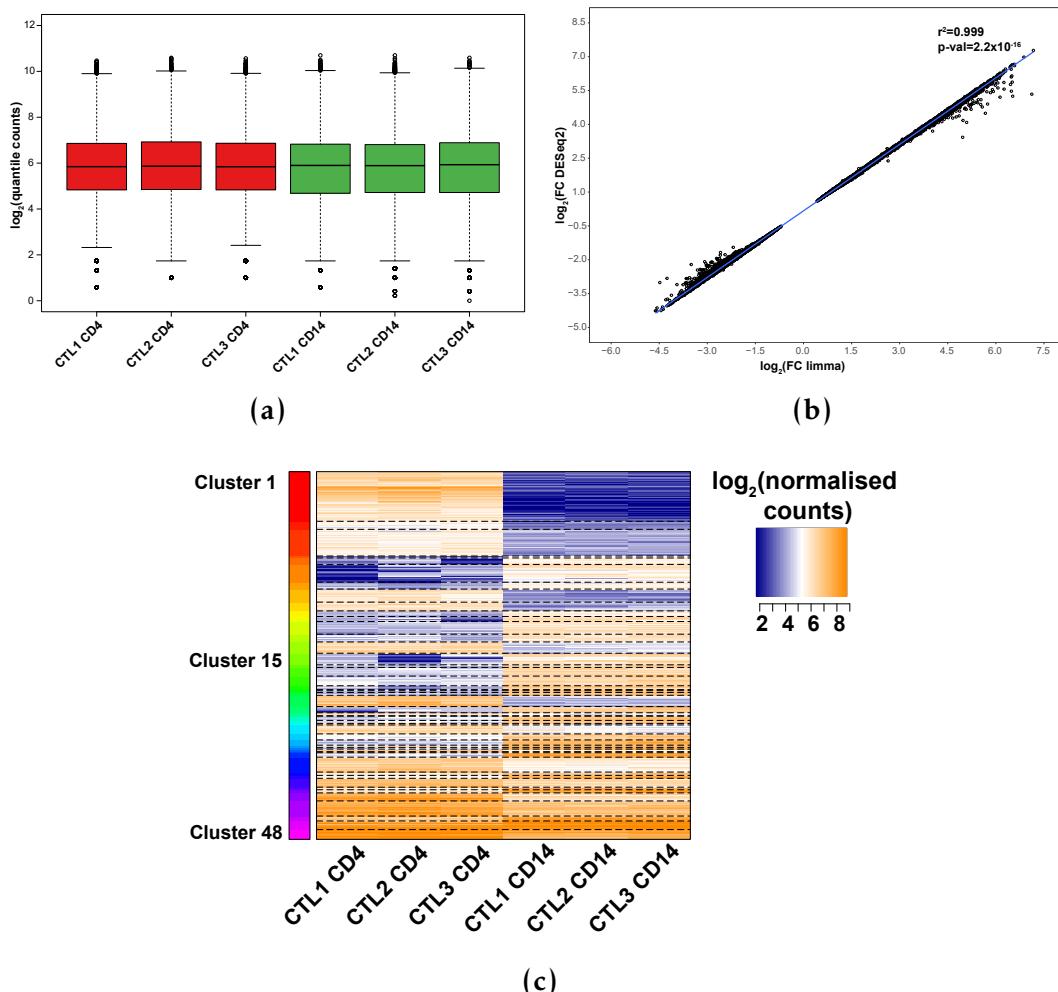


Figure 3.5: Exploration of the differential chromatin accessibility analysis using 80% as the empirical cut-off.

3.2.2 Assessment of ATAC-seq transposition times and comparison with FAST-ATAC protocol in relevant cell types

3.2.3 Limitations of ATAC-seq and FAST-ATAC to assess chromatin accessibility in KC

Due to the fact that KC is one of the most relevant cell types in psoriasis pathophysiology, ATAC-seq as described in Buenrostro *et al.*, 2013 (named as ATAC-seq 1 here) was performed in 50,000 cells of a suspensions isolated from a

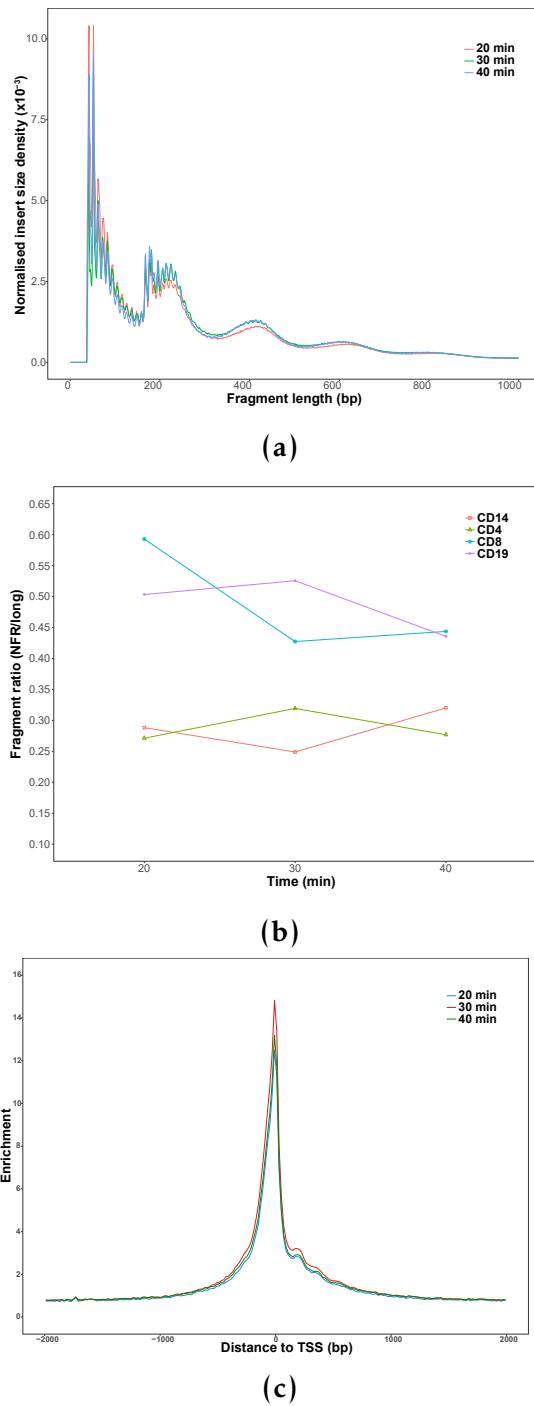


Figure 3.6: Assessment of the effect of transposition times on the ATAC-seq QC parameters

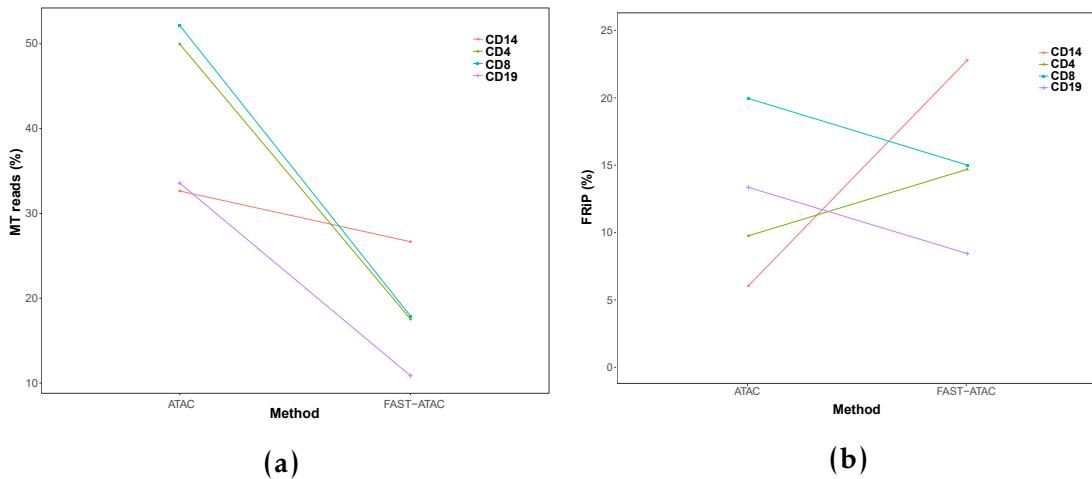


Figure 3.7: Differences in MT DNA abundance and signal specificity between ATAC-seq and FAST-ATAC protocols

psoriasis lesional skin biopsy. Two different transposition times (30 and 40 min) were tested. Since biopsy handling and lesional epidermal KC are particularly challenging this was considered the best system to test the performance of the standard protocol in the clinical setting of interest for the study. Two transposition times (30 and 40 min) were tested.

Although cell suspensions obtained from biopsies using trypsinisation of the epidermal sheet are 90% enriched in KC, they also contain significant amounts of dead cells and free-DNA released by apoptotic cells. In order to overcome this problem and the impact that it may have over ATAC-seq background signal, viable KC were selected by adherence assay. Biopsy cell suspensions were cultured for 3h in a 96-well plate and washed afterwards to ensure that only the viable and less differentiated KC would remain for downstream analysis. In parallel cultured NHEK were also used to assess the performance of the different ATAC-seq protocols.

Table for the conditions: done Tapestation profiles of the chosen condition. done Send the others to supplementary. QC measurements: for ATAC1, ATAC2 and NHEK, mention frag size distribution done DHS enrichment for p and q done but not convincing. The complex network of keratin filaments

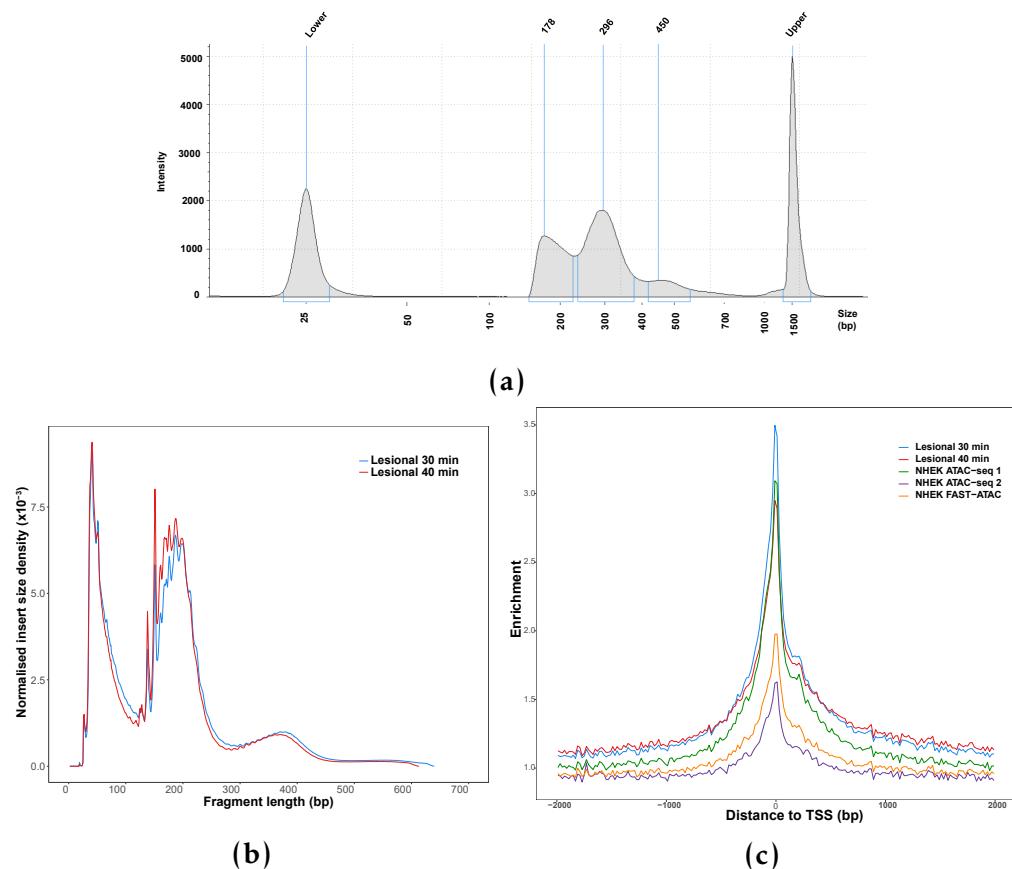


Figure 3.8: QC assessment of ATAC-seq in KC enriched cell suspension derived from a psoriatic lesional skin biopsy. Two transposition times (30 and 40 min) were tested using the standard ATAC-seq protocol (Buenrostro *et al.*, 2013 in 50,000 cells from the same suspension.

in stratified epithelia is tightly regulated during squamous cell differentiation. Keratin 14 (K14) is expressed in mitotically active basal layer cells, along with its partner keratin 5 (K5), and their expression is down-regulated as cells differentiate.

Protocol	Lysis and transposition	Key parameters
Buenrostro et al., 2013	Two steps	0.1% NP-40 and 2.5µL Tn5
Bao et al., 2015	Two steps	0.05% NP-40 and 5µL Tn5
Corces et al., 2016	One step	C1: 0.01% digitonin, 0.5µL Tn5 C2: 0.01% digitonin, 2.5µL Tn5 C3: 0.025% digitonin, 0.5µL Tn5 C4: 0.025% digitonin, 2.5 µL Tn5

Table 3.3: Description of the most relevant parameter from the ATAC-seq and FAST-ATAC protocols assayed in NHEK and skin biopsies. Transposition for all the different protocols was 30 min.

Omni-ATAC Tapestation profiles of the chosen condition include it with the supplementary that includes all other tapestation profiles. done QC measurements: frag size distribution and TSS done Track including all skin samples

Think of what to include about the biopsies in supplementary done

3.2.4 Discussion

Maybe justify in the dicussion the use of DESeq2 and limma shared based on Alasoo observation of noise effect in limma

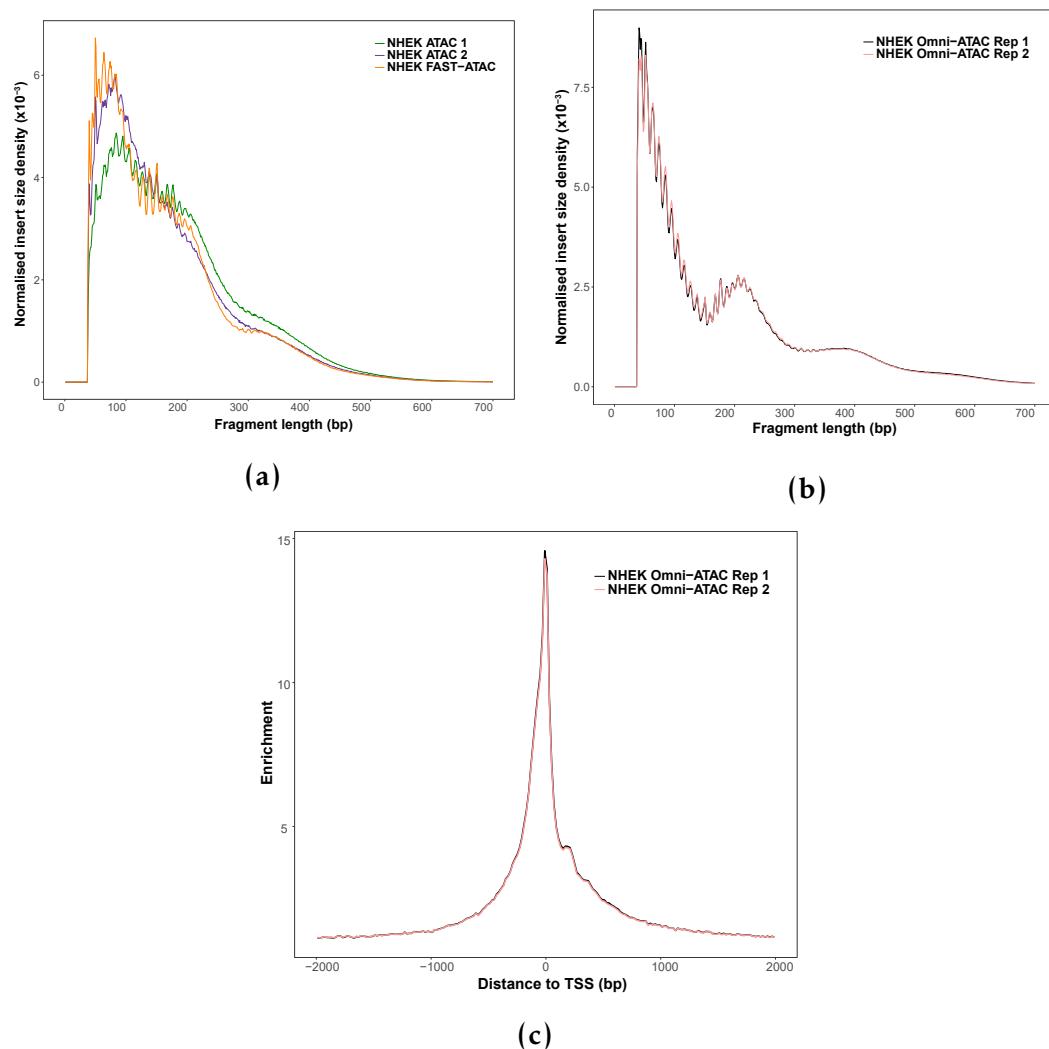


Figure 3.9: QC assessment of FAST-ATAC and Omni-ATAC in cultured NHEK.

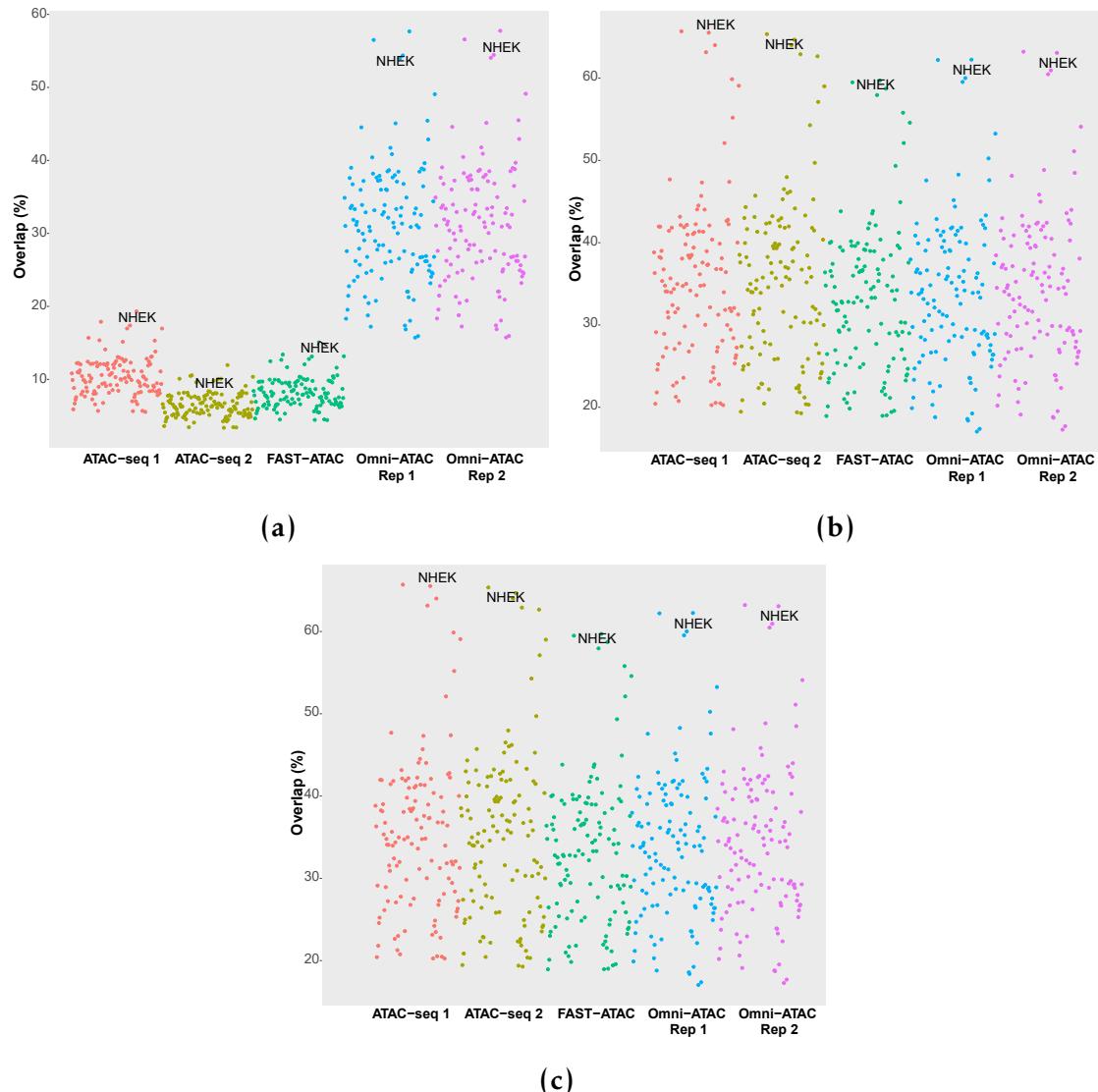


Figure 3.10: QC assessment of Omni-ATAC in NHEK and chromatin accessibility signal for the samples generated with the different ATAC-seq protocols.

Chapter 4

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

4.1 Introduction

4.2 Results

4.2.1 PsA patients cohort description and datasets

For this study blood and SF samples were collected from six PsA patients, with equal number of male and female (Table 4.1). All the patients presented oligoarticular joint affection and had been first diagnosed with psoriasis. Maybe add sth about oligoarticular?

The cohort presented a mean of 1.5 tender or swollen affected joints (TJC66 and SJC66), which is characteristic of the oligoarticular form of disease, and joint pain of xxxx. Regarding global assessment, the mean scores for the patient and physician evaluation were X and 3, respectively, in a scale of 1 to 5. These four measurements including joints and global assessment compose the PsARC

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

disease activity scores, used by clinicians as the main indicator of response to treatment by recommendation of the xxxx, as previously explained in Chapter 1.

The mean age of the cohort at the time of diagnosis was 44.3 years old and the mean disease duration 8.8 years. Interestingly, PsA1728 was diagnosed at a later age compared to the other patients in the cohort (late PsA onset clinical significance??). Moreover, CRP levels, other marker of inflammation, was also measured in all the patients presenting an average of 17.45 mg/L and being particularly higher in PSA1719 and PSA1728, compared to the other patients. At the time of sample recruitment all the PsA patients were naive for treatment and only PSA1505 had been on methotrexate therapy in the past for xxx months/years (how many years ago?). Post-visit, most of the patients qualified for TNFi biologic therapy xxxx.

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

Table 4.1: Description of PsA patients cohort recruitment and metadata. PsARC disease activity score is composed of tender joint count 66 (TJC66) and swollen joint count 66 (SJC66), joint pain (4 point score) and self-patient and physician global assessment (5 point score). Joint pain and global assessment use a likert scale based on questionnaire answers that measure the level of agreement with each of statements included. C-reactive protein (CRP).

Sample ID	Sex	Age at diagnosis	Disease duration (months)	Type	TJC66/SJC66	Physician assessment	CRP (mg/l)
PSA1718	Female	17	180	Oligo	2/2	3	6
PSA1719	Male	33	24	Oligo	1/1	3	36.6
PSA1607	Male	42	108	Oligo	1/1	4	8
PSA1728	Female	72	48	Oligo	2/2	3	43.2
PSA1801	Female	53	168	Oligo	2/2	3	9.9
PSA1505	Male	35	108	Oligo	1/1	2	1
Total	-	44.3	106	-	1.5/1.5	3	17.45

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

For each of the patients, paired data in blood and SF was generated from bulk mononuclear cells or the isolated cell types of interest (detailed in Table 4.2 and Chapter 3.2.1). However, not all types of data including ATAC-seq, PCR gene expression array, scRNA-seq and mass cytometry were generated for all six individuals of the cohort due to project constraints.

Sample ID	% FAST-ATAC	RNA PCR array	scRNA-seq	mass cytometry
PSA1718	Yes	Yes	No	Yes
PSA1719	Yes	Yes	No	Yes
PSA1607	Yes	No	Yes	Yes
PSA1728	No	Yes	No	Yes
PSA1801	No	No	Yes	Yes
PSA1605	No	No	Yes	Yes

Table 4.2: Datasets generated for the PsA cohort samples. Four types of data were generated in a paired way between blood and SF from the same individual. The types of data available varies between individuals due to project constraints. FAST-ATAC data was generated for CD14⁺, mCD4⁺, mCD8⁺ and NK cells. RNA PCR array was performed in CD14⁺, mCD4⁺ and mCD8⁺. scRNA-seq data was generated using 10X technology in bulk PBMCs, bulk SFMCs and sorted mCD4⁺ and mCD8⁺ from both tissues.

4.2.2 Differences in the chromatin accessibility landscape between circulating and SF immune cells

Quality control of open chromatin regions

The twenty four PsA samples form four cell types and two different tissues (PB and SF) were sequenced to a median of 158M reads (79M paired-end) per sample. After filtering for low quality mapping, duplicates and MT reads, the median total number of reads were 70.2M, 50.6M, 46.6 and 66.7M for CD14⁺, mCD4⁺, mCD8⁺ and NK cells, respectively (Figure 4.1 a). The differences between cell types and samples in the median of total reads remaining after filtering was inversely related to the percentage MT and duplicated reads identified (Figure 4.1 b). For example, mCD1⁺ and mCD8⁺, presented the lower

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

median of total number of reads after filtering concomitantly with the greater percentage of MT and duplicated reads. In combination, MT and duplicated reads accounted for a median of 42, 57.6, 62.2 and 40% in CD14⁺, mCD4⁺, mCD8⁺ and NK cells, respectively, importantly contributing to the loss of reads in this experiment. As previously mentioned, the MT DNA in ATAC-seq is one of the main sources of read loss, which is more accessible to the Tn5transposase due to the absence of nucleosomes. Although the FAST-ATAC protocol represented an improvement, the percentage of MT reads across amongst all the samples ranged between 2.1 and 25.4%. Similarly, despite initial optimisation of the number of PCR cycles used in the library amplification, the duplicated reads still represented between 22.9 to 55% of the total number of the pre-filtered reads.

Regarding sample quality determination, TSS enrichment analysis showed variation in the levels of background noise across cell types and highlighted the variability in performance of FAST-ATAC (Figure 4.1 c). A trend towards greater TSS enrichment in PB samples compared to SF can be observed in all four cell types. In terms of cell types,mCD4⁺ and mCD8⁺ presented the best signal-to-noise ratios, with median of 19.1 and 23.1 fold enrichment, respectively. In contrast, NK was the cell type with the lower TSS enrichment values. Particularly, the fold enrichment for PSA1719 and PSA1607 were 7.3 and 6.2, respectively, both just above the 6 to 10 acceptable range from ENCODE. If a bigger sample size was available it would be appropriate to drop this samples from the differential analysis since greater background levels will reduce the power of this approach.

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

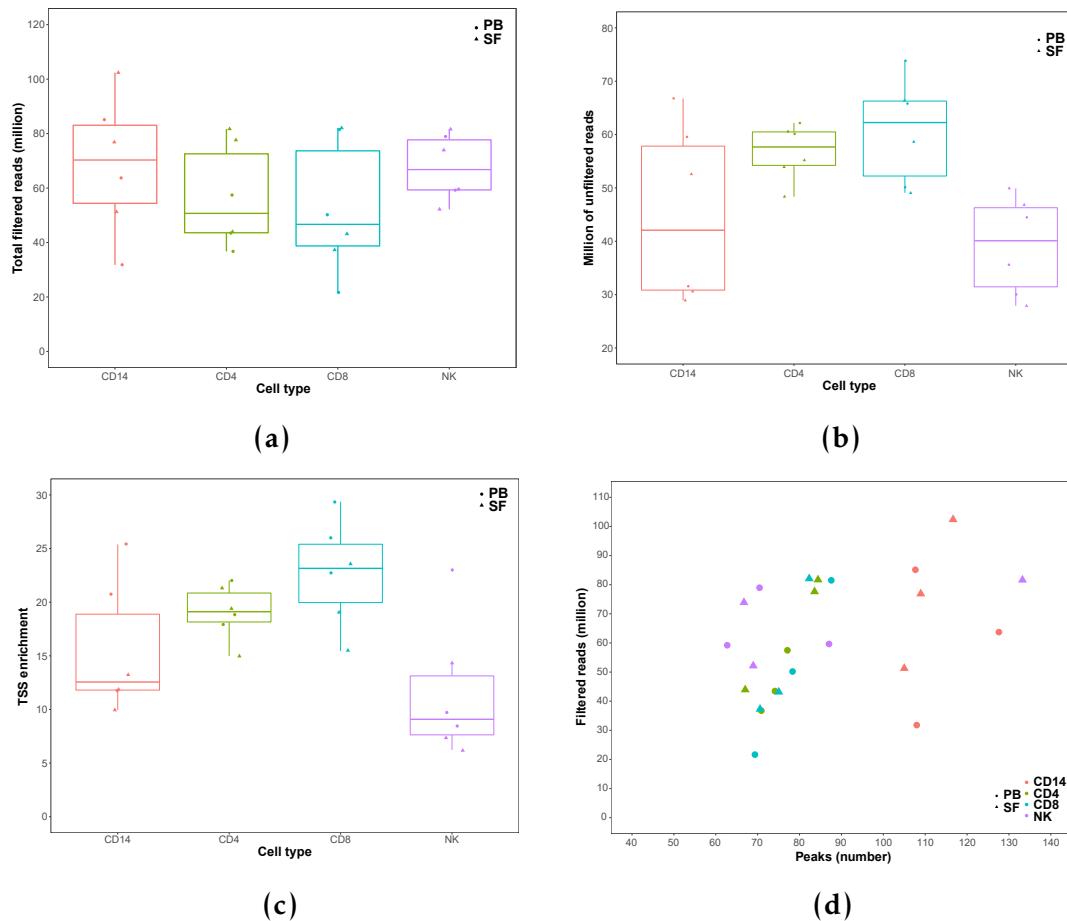


Figure 4.1: QC of FAST-ATAC PsA samples in four cell types.

When identifying open chromatin regions through peak calling and standard filtering for FDR<0.01 (not the IDR sample-specific filtering), the number of peaks ranged from $\sim 62 \times 10^3$ to $\sim 133 \times 10^3$ peaks per sample (Figure 4.1 d). A clear positive correlation between the number of called peaks and number of reads after filtering could be observed in the data. For example, CD14⁺ was the cell type with greatest number of called peaks (108.4×10^3) as well as the greater median of reads remaining after filtering when compared to the other three cell types (Figure 4.1 a). For the NK, the two samples with the greatest TSS enrichment (PSA1718 SF and PB) showed greater number of called peaks when compared to the other NK samples with similar number of reads. This observation was consistent with the correlation between sample quality and the number of identified accessible chromatin regions previously demonstrated in

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

Chapter ??ch:Results1). Overall, appropriate number of peaks were called in all the samples and no concerning outliers were identified.

Open chromatin reflects cell type specificity and functional relevance

In order to determine the ability of the open chromatin identified by the in house pipeline in the PsA sample cohort, a combined master list including all four cell types and the two tissues was built. Following Chapter 3.2.1 and Chapter 3, the combined master list contained open chromatin regions identified in at least 30% of the samples (in this case 7 samples) regardless cell type and tissue to avoid any bias.

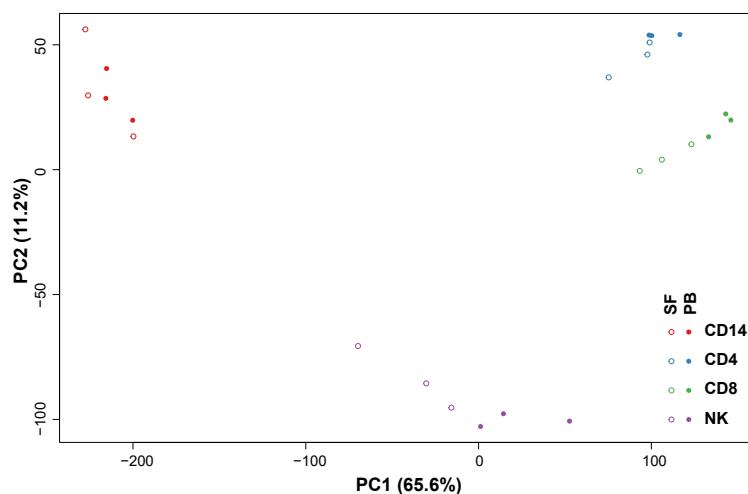


Figure 4.2: Combined PCA analysis of all four cell types isolated from blood and SF.

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

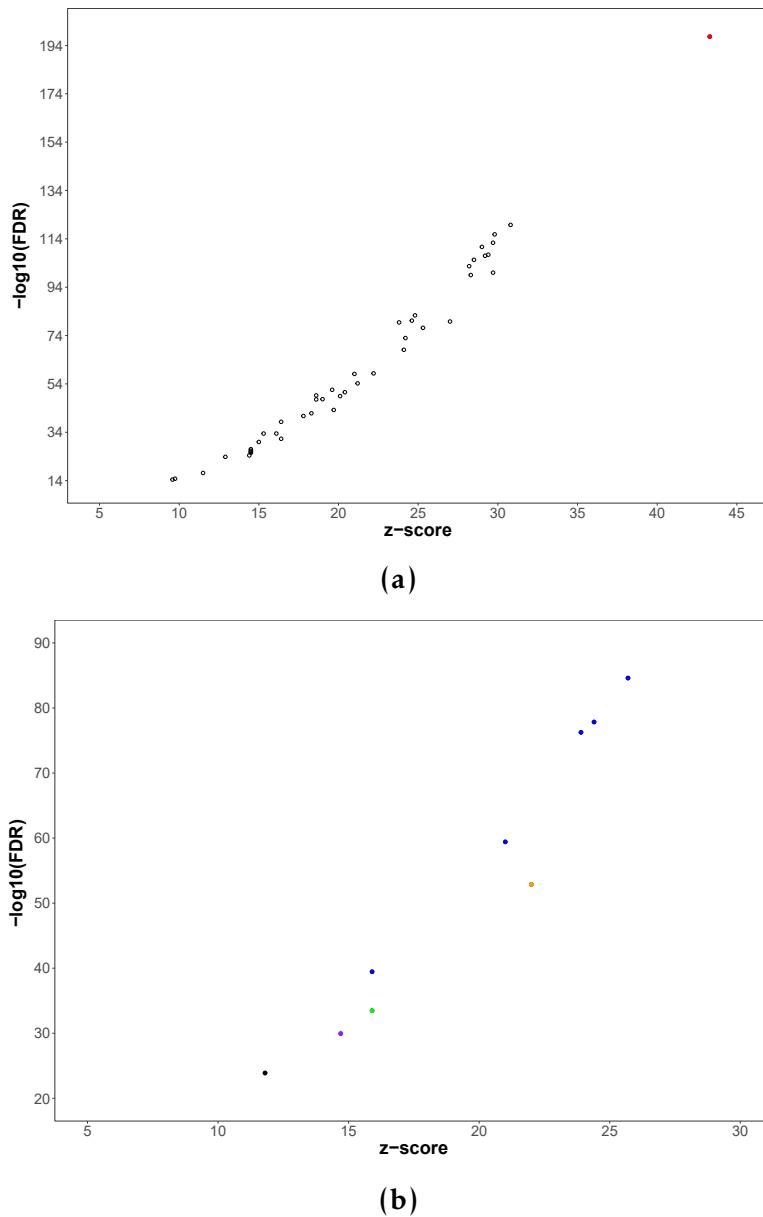


Figure 4.3: Enrichment of eQTLs in the combined cell types PsA accessible chromatin master list xxxx

Differential open chromatin analysis between blood and SF

Differential chromatin accessibility analysis was performed using a paired design between SF and PB for each of the four cell types 4.4. In all for analysis an 80% cut-off for background noise was applied in the count matrix as previously explained in Chapter 3. Only DOCs identified with DESeq2 and also shared by quantile normalisation and limma voom where taken forward.

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

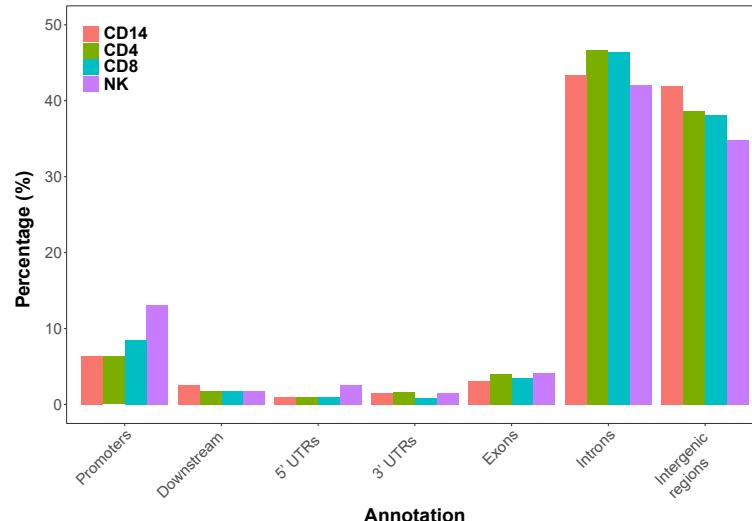
The CD14⁺ monocytes and NK showed a greater proportion of differentially accessible regions (23.3 and 8.9%, respectively) compared to CD4⁺ and CD8⁺ T cells. In CD14⁺ monocytes 3,779 out of 5,285 DOCs were more accessible in SF versus the 1,506 that were more open in the cells isolated from PB. Conversely, the number of DOCs more accessible in each of the tissues for CD4⁺, CD8⁺ and NK were evenly distributed between SF and PB.

Cell type	Total DOCs	Proportion DOCS (%)	DOCs open in SF	DOCs open in PB
CD14 ⁺	5,285	23.3	3,779	1,506
CD4 ⁺	1,329	4.3	621	708
CD8 ⁺	1,570	4.5	807	763
NK	2,314	8.9	1,223	1,091

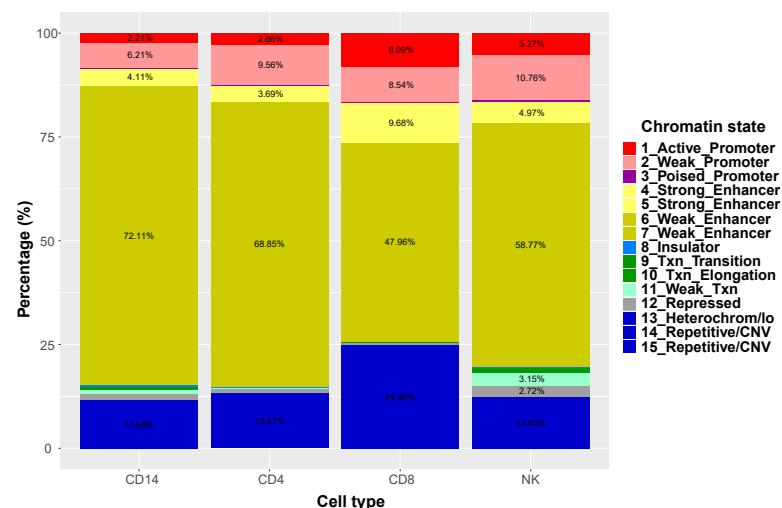
Table 4.3: xxxxxxxxx

When performing genomic annotation of the DOCs, intronic and intergenic regions included together 80% or more of all the DOCs in the four cell types (fig). DOCs annotated in universal promoter regions represented approximately between 5 to 15%. DOCs were also annotated with the fifteen states chromatin segmentation maps from the Epigenome Roadmap appropriate for each cell type (figure).

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis



(a)



(b)

Figure 4.4: Annotation with genomic regions and chromatin states of the PsA DOCs from the four cell types differential analysis. xxxx

The location of DOCs

Cell type	DOCs in gene body	Gene with > one DOC	Enhancers	Introns
CD14 ⁺	2,357	744	1,775	1,920
CD4 ⁺	700	99	504	577
CD8 ⁺	831	118	503	666
NK	1,246	235	782	937

Table 4.4: xxxxxxxxxxxx

Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in psoriatic arthritis

Here I am planning an overview of the differential analysis including -total number of doc split between open in SF and open in PB, -annotation according to chromatin accessibility in terms of all the hits -mention how many of them are annotated in gene bodies and if more than one region within the same gene and choose an example of an interesting gene, include UCSC track showing the differences

-enrichment for TFBS and others, either global or per open in SF and PB, maybe combined plots for all cell types kind of fig6 of XGR paper

DOCs highlight relevant functional pathways in a cell type and tissue specific manner

-pathway enrichment analysis if possible per open chromatin in each cell type -maybe include an A2 pathway which is different and unique between open in SF and PB in one cell type

4.2.3 Differential gene expression analysis in paired circulating and synovial immune cells

Array data

4.3 Discussion

fGWAS analysis as Matthias did would be of interest but needs appropriate GWAS data I am going to try using XGR to do some of this

Appendices

A Establishment of methods to assess genome-wide chromatin accessibility 108

Bibliography 125

List of Figures

A.1	FAST-ATAC and Omni-ATAC NHEK tapestation profiles.	109
A.2	Assessment of TSS enrichment from ATAC-seq and FAST-ATAC in healthy and psoriasis skin biopsies samples.	110

List of Tables

Appendix A

Establishment of methods to assess genome-wide chromatin accessibility

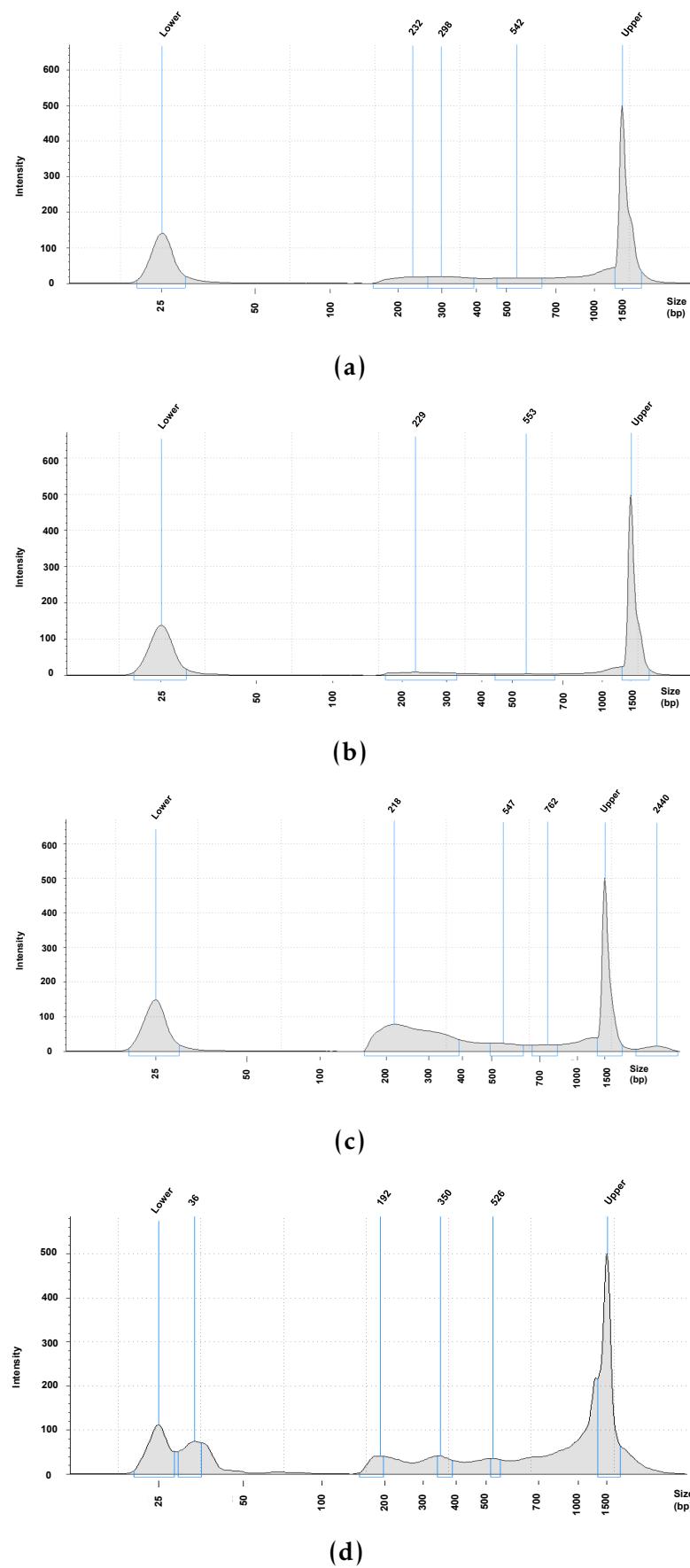


Figure A.1: FAST-ATAC and Omni-ATAC NHEK tapestation profiles.

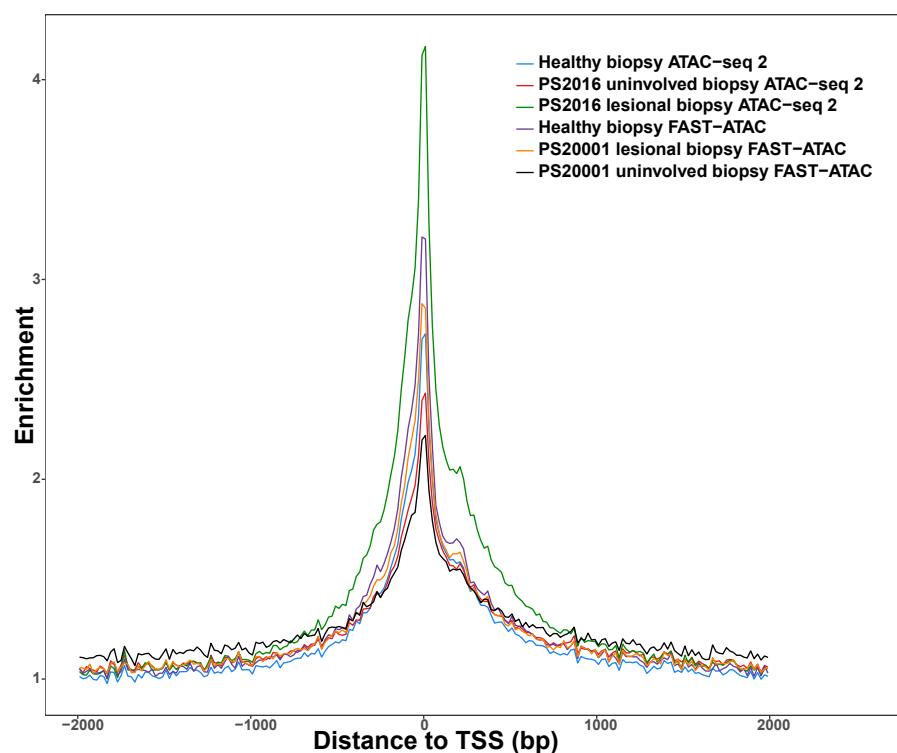


Figure A.2: Assessment of TSS enrichment from ATAC-seq and FAST-ATAC in healthy and psoriasis skin biopsies samples.

Bibliography

- Abecasis GR Auton A, Brooks LD DePristo MA Durbin RM Handsaker RE Kang HM Marth GT McVean GA. (2012). "An integrated encyclopedia of DNA elements in the human genome". *Nature* 489: pp. 57–74. issn: 0028-0836. doi: 10.1038/nature11247.
- Abramson, S. R. and and, Khattri-S. of (2016). "A Review of Current Evidence for TNF Inhibitor Switching in Psoriatic Arthritis". *Journal of Psoriasis and*. URL: <http://journals.sagepub.com/doi/abs/10.1177/247553031600100304>.
- Adey, Andrew, Morrison, Hilary G., Asan, Xun, Xu, Kitzman, Jacob O., Turner, Emily H., Stackhouse, Bethany, MacKenzie, Alexandra P., Caruccio, Nicholas C., Zhang, Xiuqing, et al. (2010). "Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition". *Genome Biology* 11. issn: 1465-6906. doi: info : pmid / 21143862. URL: <http://genomebiology.com/2010/11/12/r119>.
- Appel, H., Maier, R., and Arthritis, Wu-P. (2011). "Analysis of IL-17+ cells in facet joints of patients with spondyloarthritis suggests that the innate immune pathway might be of greater relevance than the Th17". *Arthritis*. URL: <https://arthritis-research.biomedcentral.com/articles/10.1186/ar3370>.
- Arend, W. P., Palmer, G., and reviews, Gabay-C. (2008). "IL1, IL18, and IL33 families of cytokines". *Immunological reviews*. doi: 10.1111/j.1600-065X.2008.00624.x. URL: <http://dx.doi.org/10.1111/j.1600-065X.2008.00624.x>.
- Austin, Lisa M., Ozawa, Maki, Kikuchi, Toyoko, Walters, Ian B., and Krueger, James G. (1999). "The majority of epidermal T cells in psoriasis vulgaris lesions can produce type 1 cytokines, interferon-, interleukin-2, and tumor necrosis factor-, defining TC1 (Cytotoxic T Lymphocyte) and TH1 effector populations: 1 a type 1 differentiation bias is also measured in circulating blood T cells in psoriatic patients". *Journal of Investigative Dermatology* 113: pp. 752–759. URL: [http://scholar.google.com/scholar?q=ThemajorityofepidermalTcellsinpsoriasisvulgarislesionscanproducetype1cytokines,interferon-,interleukin-2, andtumornecrosisfactor-,definingTC1\(CytotoxicT Lymphocyte\)andTH1effectorpopulations:1atype1differentiationbiasisalsomeasuredincirculatingbloodTcellsinpsoriaticpatients&hl=en&num=20&as_sdt=0%2C22](http://scholar.google.com/scholar?q=ThemajorityofepidermalTcellsinpsoriasisvulgarislesionscanproducetype1cytokines,interferon-,interleukin-2, andtumornecrosisfactor-,definingTC1(CytotoxicT Lymphocyte)andTH1effectorpopulations:1atype1differentiationbiasisalsomeasuredincirculatingbloodTcellsinpsoriaticpatients&hl=en&num=20&as_sdt=0%2C22).
- Baeten, D., Breban, M., Lories, R., and, Schett-G. (2013). "Are spondylarthritides related but distinct conditions or a single disease with a heterogeneous phenotype?" *Arthritis &*. doi: 10.1002/art.37829. URL: <http://dx.doi.org/10.1002/art.37829>.

BIBLIOGRAPHY

- Baker, K. F. and diseases, Isaacs-J. D. of the rheumatic (2017). "Novel therapies for immune-mediated inflammatory diseases: What can we learn from their use in rheumatoid arthritis, spondyloarthritis, systemic lupus erythematosus". *Annals of the rheumatic diseases*. doi: 10.1136/annrheumdis-2017-211555. URL: <http://dx.doi.org/10.1136/annrheumdis-2017-211555>.
- Bansal, V., Bashir, A., and research, Bafna-V. (2007). "Evidence for large inversion polymorphisms in the human genome from HapMap data". *Genome research*. doi: 10.1101/gr.5774507. URL: <http://dx.doi.org/10.1101/gr.5774507>.
- Bergboer, J. G. M. and of, Oostveen-A. M. (2012). "Paediatric onset psoriasis is associated with ERAP1 and IL23R loci, LCE3C_LCE3B deletion and HLAC * 06". *British Journal of*. doi: 10.1111/j.1365-2133.2012.10992.x. URL: <http://dx.doi.org/10.1111/j.1365-2133.2012.10992.x>.
- Bernstein, B. E., Stamatoyannopoulos, J. A., Costello, J. F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M. A., Beaudet, A. L., Ecker, J. R., et al. (2010). "The NIH Roadmap Epigenomics Mapping Consortium". *Nat Biotechnol* 28: pp. 1045–8. ISSN: 1087-0156. doi: 10.1038/nbt1010-1045.
- Black, A. P. B. and of, ArdernJones-M. R. journal (2007). "Human keratinocyte induction of rapid effector function in antigenspecific memory CD4+ and CD8+ T cells". *European journal of*. doi: 10.1002/eji.200636915. URL: <http://dx.doi.org/10.1002/eji.200636915>.
- Blais, M. E., Dong, T., and Immunology, RowlandJones-S. (2011). "HLAC as a mediator of natural killer and Tcell activation: spectator or key player?" *Immunology*. doi: 10.1111/j.1365-2567.2011.03422.x. URL: <http://dx.doi.org/10.1111/j.1365-2567.2011.03422.x>.
- Blonska, M. and research, Lin-X. (2011). "NF-B signaling pathways regulated by CARMA family of scaffold proteins". *Cell research*. URL: <https://www.nature.com/articles/cr2010182>.
- Bowes, J., Ho, P., Flynn, E., and the, Ali-F. of (2012). "Comprehensive assessment of rheumatoid arthritis susceptibility loci in a large psoriatic arthritis cohort". *Annals of the*. doi: 10.1136/annrheumdis-2011-200802. URL: <http://dx.doi.org/10.1136/annrheumdis-2011-200802>.
- Bowes, John, Budu-Aggrey, Ashley, Huffmeier, Ulrike, Uebe, Steffen, Steel, Kathryn, Hebert, Harry L., Wallace, Chris, Massey, Jonathan, Bruce, Ian N., Bluett, James, et al. (2015). "Dense genotyping of immune-related susceptibility loci reveals new insights into the genetics of psoriatic arthritis". *Nature communications* 6: p. 6046. ISSN: 2041-1723. doi: 10.1038/ncomms7046. URL: <http://dx.doi.org/10.1038/ncomms7046>.
- Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., Karczewski, K. J., Park, J., Hitz, B. C., Weng, S., et al. (2012). "Annotation of functional variation in personal genomes using RegulomeDB". *Genome Res* 22: pp. 1790–7. ISSN: 1088-9051. doi: 10.1101/gr.137323.112.

BIBLIOGRAPHY

- Boyman, O., Hefti, H. P., Conrad, C., Nickoloff, B. J., Suter, M., and Nestle, F. O. (2004). "Spontaneous development of psoriasis in a new animal model shows an essential role for resident T cells and tumor necrosis factor-alpha". *J Exp Med* 199: pp. 731–6. ISSN: 0022-1007 (Print)0022-1007. doi: 10.1084/jem.20031482. URL: <http://dx.doi.org/10.1084/jem.20031482>.
- Brest, P., Lapaquette, P., Souidi, M., and Nature, Lebrigand-K. (2011). "A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease". *Nature*. doi: 10.1038/ng.762. URL: <http://dx.doi.org/10.1038/ng.762>.
- Buenrostro, Jason D., Wu, Beijing, Chang, Howard Y., and Greenleaf, William J. (2015). "ATAC seq: A Method for Assaying Chromatin Accessibility Genome Wide". *Current Protocols in Molecular Biology* 109: pp. 1–9. doi: 10.1002/0471142727.mb2129s109. URL: <http://dx.doi.org/10.1002/0471142727.mb2129s109>.
- Bunt, M. van de, Cortes, A., Brown, M. A., and Morris, A. P. (2015). "Evaluating the Performance of Fine-Mapping Strategies at Common Variant GWAS Loci". *PLoS Genet* 11: pp. 1–14. doi: 10.1371/journal.pgen.1005535. URL: <http://dx.doi.org/10.1371/journal.pgen.1005535>.
- Burgos-Pol, R. and Dermo, Martnez-Sesmero-J. M. (2016). "The cost of psoriasis and psoriatic arthritis in 5 European countries: a systematic review". *Actas Dermo*. URL: <http://www.sciencedirect.com/science/article/pii/S1578219016301524>.
- Cameron, A. L. and of, Kirby-B. (2003). "Circulating natural killer cells in psoriasis". *British Journal of*. doi: 10.1046/j.1365-2133.2003.05319.x. URL: <http://dx.doi.org/10.1046/j.1365-2133.2003.05319.x>.
- Capon, Francesca (2017). "The Genetic Basis of Psoriasis". *International Journal of Molecular Sciences* 18: p. 2526. ISSN: 1422-0067. doi: 10.3390/ijms18122526. URL: <http://dx.doi.org/10.3390/ijms18122526>.
- Cargill, Michele, Schrodi, Steven J., Chang, Monica, Garcia, Veronica E., Brandon, Rhonda, Callis, Kristina P., Matsunami, Nori, Ardlie, Kristin G., Civello, Daniel, and Catanese, Joseph J. (2007). "A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes". *The American Journal of Human Genetics* 80: pp. 273–290. URL: <http://www.sciencedirect.com/science/article/pii/S0002929707626858>.
- Carter, K. W., Pluzhnikov, A., Timms, A. E., Miceli-Richard, C., Bourgoin, C., Wordsworth, B. P., Jean-Pierre, H., Cox, N. J., Palmer, L. J., and Breban, M. (2007). "Combined analysis of three whole genome linkage scans for ankylosing spondylitis". *Rheumatology* 46: pp. 763–771. doi: 10.1093/rheumatology/ke1443. URL: <http://dx.doi.org/10.1093/rheumatology/ke1443>.

BIBLIOGRAPHY

- Chandran, Vinod, Schentag, Catherine T., Brockbank, John E., Pellett, Fawnda J., Shanmugarajah, S., Toloza, Sergio M. A., Rahman, Proton, and Gladman, Dafna D. (2009). "Familial aggregation of psoriatic arthritis". *Annals of the rheumatic diseases* 68: pp. 664–667. doi: 10.1136/ard.2008.089367. url: <http://dx.doi.org/10.1136/ard.2008.089367>.
- Cid, R. de, Riveira-Munoz, E., Zeeuwen, P. L., Robarge, J., Liao, W., Dannhauser, E. N., Giardina, E., Stuart, P. E., Nair, R., Helms, C., et al. (2009). "Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis". *Nat Genet* 41: pp. 211–5. issn: 1061-4036. doi: 10.1038/ng.313.
- Clark, R. A., Chong, B., Mirchandani, N., Brinster, N. K., Yamanaka, K., Dowgiert, R. K., and Kupper, T. S. (2006). "The vast majority of CLA+ T cells are resident in normal skin". *J Immunol* 176: pp. 4431–9. issn: 0022-1767 (Print)0022-1767. url: <http://dx.doi.org/>.
- Clegg, Daniel O., Reda, Domenic J., Mejias, Edwin, Cannon, Grant W., Weisman, Michael H., Taylor, Thomas, Mak, Elly, Blackburn, Warren D., Vasey, Frank B., and Mahowald, Maren L. (1996). "Comparison of sulfasalazine and placebo in the treatment of psoriatic arthritis. A Department of Veterans Affairs Cooperative Study". *Arthritis & Rheumatology* 39: pp. 2013–2020. doi: 10.1002/art.1780391210. url: <http://dx.doi.org/10.1002/art.1780391210>.
- Coates, L. C., FitzGerald, O., Helliwell, P. S., and and, Paul-C. in arthritis (2016). "Psoriasis, psoriatic arthritis, and rheumatoid arthritis: Is all inflammation the same?" *Seminars in arthritis and*. url: <http://www.sciencedirect.com/science/article/pii/S0049017216300646>.
- Costello, P., Bresnihan, B., and of, O'Farrelly-C. (1999). "Predominance of CD8+ T lymphocytes in psoriatic arthritis". *The Journal of*. url: <http://europepmc.org/abstract/med/10332977>.
- Cui, Y., Li, G., Li, S., and Biology, Wu-R. in (2010). "Designs for linkage analysis and association studies of complex diseases". *Statistical Methods in Molecular Biology*. doi: 10.1007/978-1-60761-580-4_6. url: http://dx.doi.org/10.1007/978-1-60761-580-4_6.
- Dand, N., Mucha, S., Tsoi, L. C., and molecular, Mahil-S. K. (2017). "Exome-wide association study reveals novel psoriasis susceptibility locus at TNFSF15 and rare protective alleles in genes contributing to type I IFN signalling". *Human molecular*. doi: 10.1093/hmg/ddx328. url: <http://dx.doi.org/10.1093/hmg/ddx328>.
- Das, Sayantan, Stuart, Philip E., Ding, Jun, Tejasvi, Trilokraj, Li, Yanming, Tsoi, Lam C., Chandran, Vinod, Fischer, Judith, Helms, Cynthia, and Duffin, Kristina (2014). "Fine mapping of eight psoriasis susceptibility loci". *European Journal of Human Genetics*. doi: 10.1038/ejhg.2014.172. url: <http://dx.doi.org/10.1038/ejhg.2014.172>.

BIBLIOGRAPHY

- Diani, M., Altomare, G., and reviews, Reali-E. (2015). "T cell responses in psoriasis and psoriatic arthritis". *Autoimmunity reviews*. URL: <https://www.sciencedirect.com/science/article/pii/S1568997214002845>.
- Diluvio, L., Vollmer, S., Besgen, P., and of, Ellwart-J. W. (2006). "Identical TCR -chain rearrangements in streptococcal angina and skin lesions of patients with psoriasis vulgaris". *The Journal of*. doi: 10.4049/jimmunol.176.11.7104. URL: <http://dx.doi.org/10.4049/jimmunol.176.11.7104>.
- Dimas, Antigone S., Deutsch, Samuel, Stranger, Barbara E., Montgomery, Stephen B., Borel, Christelle, Attar-Cohen, Homa, Ingle, Catherine, Beazley, Claude, Arcelus, Maria, and Sekowska, Magdalena (2009). "Common regulatory variation impacts gene expression in a cell typedependent manner". *Science* 325: pp. 1246–1250. URL: <http://www.sciencemag.org/content/325/5945/1246.short>.
- Duffy, D. L., Spelman, L. S., and of, Martin-N. G. of the (1993). "Psoriasis in Australian twins". *Journal of the American Academy of*. URL: <https://www.sciencedirect.com/science/article/pii/0190962293702069>.
- Edwards, S. L., Beesley, J., French, J. D., and Dunning, A. M. (2013). "Beyond GWASs: illuminating the dark road from association to function". *Am J Hum Genet* 93: pp. 779–97. ISSN: 0002-9297. doi: 10.1016/j.ajhg.2013.10.012.
- Eedy, D. J., Burrows, D., Bridges, J. M., and Jones, F. G. (1990). "Clearance of severe psoriasis after allogenic bone marrow transplantation". *BMJ* 300: p. 908. ISSN: 0959-8138 (Print)1468-5833 (Electronic). URL: <http://dx.doi.org/>.
- Ellinghaus, David, Jostins, Luke, Spain, Sarah L., Cortes, Adrian, Bethune, Jrn, Han, Buhm, Park, Yu, Raychaudhuri, Soumya, Pouget, Jennie G., Hbenthal, Matthias, et al. (2016). "Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci". *Nature Genetics* 48: pp. 510–518. ISSN: 1061-4036. doi: 10.1038/ng.3528. URL: <http://dx.doi.org/10.1038/ng.3528>.
- Ellinghaus, E., Ellinghaus, D., Stuart, P. E., and Nair, R. P. (2010). "Genome-wide association study identifies a psoriasis susceptibility locus at TRAF3IP2". *Nature genetics* 42: pp. 991–996. doi: 10.1038/ng.689. URL: <http://dx.doi.org/10.1038/ng.689>.
- Eppinga, H. and rheumatology, Konstantinov-S. R. (2014). "The microbiome and psoriatic arthritis". *Current rheumatology*. doi: 10.1007/s11926-013-0407-2. URL: <http://dx.doi.org/10.1007/s11926-013-0407-2>.
- Ernst, J., Kheradpour, P., Mikkelsen, T. S., and Shores, N. (2011). "Mapping and analysis of chromatin state dynamics in nine human cell types". *Nature* 473: pp. 43–49. URL: <http://www.nature.com/nature/journal/v473/n7345/abs/nature09906.html>.

BIBLIOGRAPHY

- Evans, D. M., Spencer, C. C. A., Pointon, J. J., Su, Z., and Nature, Harvey-D. (2011). "Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility". *Nature*. URL: <https://www.nature.com/articles/ng.873>.
- Fairfax, Benjamin P., Makino, Seiko, Radhakrishnan, Jayachandran, Plant, Katharine, Leslie, Stephen, Dilthey, Alexander, Ellis, Peter, Langford, Cordelia, Vannberg, Fredrik O., and Knight, Julian C. (2012). "Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles". *Nature genetics* 44: pp. 502–510. ISSN: 1061-4036. doi: 10.1038/ng.2205. URL: <http://dx.doi.org/10.1038/ng.2205>.
- Fairfax, Benjamin P., Humburg, Peter, Makino, Seiko, Naranbhai, Vivek, Wong, Daniel, Lau, Evelyn, Jostins, Luke, Plant, Katharine, Andrews, Robert, McGee, Chris, et al. (2014). "Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression". *Science (New York, N.Y.)* 343: pp. 1118–1132. ISSN: 0036-8075. doi: 10.1126/science.1246949. URL: <http://dx.doi.org/10.1126/science.1246949>.
- Farh, Kyle, Carrasco-Alfonso, J., Dita, Mayer, Luckey, C., Nikolaos, A., Philip, L., Vijay, K., Charles, B., Mark, J., David, A., et al. (2014). "Genetic and epigenetic fine mapping of causal autoimmune disease variants". *Nature* 518: 337343. doi: 10.1038/nature13835. URL: <http://dx.doi.org/10.1038/nature13835>.
- Finlay, Andrew (2005). "Current severe psoriasis and the rule of tens". *British Journal of Dermatology* 152: pp. 861–867. doi: 10.1111/j.1365-2133.2005.06502.x. URL: <http://dx.doi.org/10.1111/j.1365-2133.2005.06502.x>.
- Fredriksson, T. and Dermatology, Pettersson-U. (1978). "Severe psoriasisoral therapy with a new retinoid". *Dermatology*. doi: 10.1159/000250839. URL: <http://dx.doi.org/10.1159/000250839>.
- Furie, R., Khamashta, M., and, Merrill-J. T. (2017). "Anifrolumab, an AntiInterferon Receptor Monoclonal Antibody, in ModerateSevere Systemic Lupus Erythematosus". *Arthritis &*. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/art.39962>.
- Gelfand, J. M., Gladman, D. D., Mease, P. J., and Academy, Smith-N. (2005). "Epidemiology of psoriatic arthritis in the population of the United States". *the American Academy*. URL: <http://www.sciencedirect.com/science/article/pii/S0190962205010108>.
- Gelfand, J. M., Neumann, A. L., Shin, D. B., and Jama, Wang-X. (2006). "Risk of myocardial infarction in patients with psoriasis". *Jama*. URL: <http://jamanetwork.com/journals/jama/fullarticle/203598>.

BIBLIOGRAPHY

- Gladman, D. D., Anhorn, K. A., Schachter, R. K., and Mervart, H. (1986). "HLA antigens in psoriatic arthritis". *The Journal of Rheumatology* 13: pp. 586–592. URL: http://scholar.google.com/scholar?q=HLAantigeninspsoriaticarthritis.&btnG=&hl=en&num=20&as_sdt=0%2C22.
- Gladman, D. D., Antoni, C., Mease, P., and rheumatic, Clegg-D. O. of the (2005). "Psoriatic arthritis: epidemiology, clinical features, course, and outcome". *Annals of the rheumatic*. URL: http://ard.bmjjournals.org/content/64/suppl_2/i14.short.
- Glessner, J. T., Wang, K., Cai, G., Korvatska, O., and Nature, Kim-C. E. (2009). "Autism genome-wide copy number variation reveals ubiquitin and neuronal genes". *Nature*. doi: 10.1038/nature07953. URL: <http://dx.doi.org/10.1038/nature07953>.
- Gobbi, De M., Viprakasit, V., Hughes, J. R., and, Fisher-C. (2006). "A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter". doi: 10.1126/science.1126431. URL: <http://dx.doi.org/10.1126/science.1126431>.
- Gossec, L., Smolen, J. S., Ramiro, S., Wit, M. de, Cutolo, M., Dougados, M., Emery, P., Landew, R., Oliver, S., Aletaha, D., et al. (2016). "European League Against Rheumatism (EULAR) recommendations for the management of psoriatic arthritis with pharmacological therapies: 2015 update". *Annals of the rheumatic diseases* 75: pp. 499–510. ISSN: 0003-4967. doi: 10.1136/annrheumdis-2015-208337. URL: <http://dx.doi.org/10.1136/annrheumdis-2015-208337>.
- Gradman, R. J. and Reznikoff, W. S. (2008). "Tn5 synaptic complex formation: role of transposase residue W450". *J Bacteriol* 190: pp. 1484–7. ISSN: 0021-9193. doi: 10.1128/jb.01488-07.
- Gregory, Adam P., Dendrou, Calliope A., Attfield, Kathrine E., Haghikia, Aiden, Xifara, Dionysia K., Butter, Falk, Poschmann, Gereon, Kaur, Gurman, Lambert, Lydia, and Leach, Oliver A. (2012). "TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis". *Nature* 488: p. 508. doi: 10.1038/nature11307. URL: <http://dx.doi.org/10.1038/nature11307>.
- Griffiths, C. E. M. and Lancet, Jnwn (2007). "Pathogenesis and clinical features of psoriasis". *The Lancet*. URL: <http://www.sciencedirect.com/science/article/pii/S0140673607611283>.
- Gudjonsson, J. E. and of, Karason-A. journal (2003). "Psoriasis patients who are homozygous for the HLACw* 0602 allele have a 2.5fold increased risk of developing psoriasis compared with Cw6 heterozygotes". *British journal of*. doi: 10.1046/j.1365-2133.2003.05115.x. URL: <http://dx.doi.org/10.1046/j.1365-2133.2003.05115.x>.

BIBLIOGRAPHY

- GutowskaOwsiak, D. and Schaupp, A. L. (2012). "IL17 downregulates filaggrin and affects keratinocyte expression of genes associated with cellular adhesion". *Experimental dermatology* 21: pp. 104–110. doi: 10.1111/j.1600-0625.2011.01412.x. URL: <http://dx.doi.org/10.1111/j.1600-0625.2011.01412.x>.
- Haddad, Amir and Chandran, Vinod (2013). "Arthritis Mutilans". *Current Rheumatology Reports* 15. issn: 1523-3774. doi: 10.1007/s11926-013-0321-7. URL: <http://dx.doi.org/10.1007/s11926-013-0321-7>.
- Hammer, R. E., Maika, S. D., Richardson, J. A., and Cell, Tang-J. P. (1990). "Spontaneous inflammatory disease in transgenic rats expressing HLA-B27 and human 2m: an animal model of HLA-B27-associated human disorders". *Cell*. URL: <http://www.sciencedirect.com/science/article/pii/009286749090512D>.
- Hanson, A. L., Cuddihy, T., Haynes, K., and, Loo-D. (2018). "Genetic Variants in ERAP1 and ERAP2 Associated With ImmuneMediated Diseases Influence Protein Expression and the Isoform Profile". *Arthritis &.* URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/art.40369>.
- Henseler, T. and Christophers, E. (1985). "Psoriasis of early and late onset: characterization of two types of psoriasis vulgaris". *J Am Acad Dermatol* 13: pp. 450–6. issn: 0190-9622 (Print) 0190-9622.
- Hirschhorn, Joel N. (2005). "Genetic Approaches to Studying Common Diseases and Complex Traits". *Pediatr Res* 57: pp. 74–77. issn: 0031-3998. URL: <http://dx.doi.org/10.1203/01.PDR.0000159574.98964.87>.
- Hitchon, Carol A., Danning, Carol L., Illei, Gabor G., El-Gabalawy, Hani S., and Boumpas, Dimitrios T. (2002). "Gelatinase expression and activity in the synovium and skin of patients with erosive psoriatic arthritis". *The Journal of rheumatology* 29: pp. 107–117. URL: [http://scholar.google.com/scholar?q=Gelatinaseexpressionandactivityinthesynoviumandskinofpatientswi&bbtnG=&hl=en&num=20&as_sdt=0%2C22](http://scholar.google.com/scholar?q=Gelatinaseexpressionandactivityinthesynoviumandskinofpatientswi&btnG=&hl=en&num=20&as_sdt=0%2C22).
- Hu, Stephen, Yu, Hsin-Su, Yen, Feng-Lin, Lin, Chi-Ling, Chen, Gwo-Shing, and Lan, Cheng-Che E. (2016). "Neutrophil extracellular trap formation is increased in psoriasis and induces human -defensin-2 production in epidermal keratinocytes". *Scientific Reports* 6: p. 31119. issn: 2045-2322. doi: 10.1038/srep31119. URL: <http://dx.doi.org/10.1038/srep31119>.
- Hundhausen, C., Bertoni, A., Mak, R. K., and Investigative, Botti-E. of (2012). "Allele-Specific Cytokine Responses at the HLA-C Locus: Implications for Psoriasis". *Journal of Investigative*. URL: <http://www.sciencedirect.com/science/article/pii/S0022202X15356189>.
- Husted, J. A., care, and, Thavaneswaran-A. (2011). "Cardiovascular and other comorbidities in patients with psoriatic arthritis: a comparison with patients with psoriasis". *Arthritis care &*. doi: 10.1002/acr.20627. URL: <http://dx.doi.org/10.1002/acr.20627>.

BIBLIOGRAPHY

- Ibrahim, G., Waxman, R., and Helliwell, P. S. (2009). "The prevalence of psoriatic arthritis in people with psoriasis". *Arthritis Rheum* 61: pp. 1373–8. issn: 0004-3591 (Print) 0004-3591. doi: 10.1002/art.24608.
- Jabbari, Ali, Surez-Farias, Mayte, Dewell, Scott, and Krueger, James G. (2012). "Transcriptional profiling of psoriasis using RNA-seq reveals previously unidentified differentially expressed genes". *The Journal of investigative dermatology* 132: pp. 246–249. issn: 0022-202X. doi: 10.1038/jid.2011.267. url: <http://dx.doi.org/10.1038/jid.2011.267>.
- Jackson, B., Tilli, Cmlj, Hardman, M. J., and Investigative, Avilion-A. A. of (2005). "Late cornified envelope family in differentiating epithelial response to calcium and ultraviolet irradiation". *Journal of Investigative*. url: <https://www.sciencedirect.com/science/article/pii/S0022202X15322934>.
- Jacobson, Christine C., Kumar, Sandeep, and Kimball, Alexa B. (2011). "Latitude and psoriasis prevalence". *Journal of the American Academy of Dermatology* 65: pp. 870–873. issn: 0190-9622. doi: 10.1016/j.jaad.2009.05.047. url: <http://dx.doi.org/10.1016/j.jaad.2009.05.047>.
- Jonker, M. J., Leeuw, W. C. de, and acids, Marinkovi-M. (2014). "Absence/presence calling in microarray-based CGH experiments with non-model organisms". *Nucleic acids*. doi: 10.1093/nar/gku343. url: <http://dx.doi.org/10.1093/nar/gku343>.
- Jordan, C. T., Cao, L., Roberson, E. D. O., and of, Duan-S. (2012). "Rare and common variants in CARD14, encoding an epidermal regulator of NF-kappaB, in psoriasis". *The American Journal of*. url: <https://www.sciencedirect.com/science/article/pii/S0002929712001577>.
- Karason, A., Gudjonsson, J. E., and of, Upmanyu-R. (2003). "A susceptibility gene for psoriatic arthritis maps to chromosome 16q: evidence for imprinting". *The American Journal of*. url: <https://www.sciencedirect.com/science/article/pii/S0002929707605102>.
- Kent, L. (2009). "Freezing and thawing human embryonic stem cells". *Journal of visualized experiments: JoVE*. url: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3152242/>.
- Kim, G. K., of, clinical, and aesthetic, Del J. Q. (2010). "Drug-provoked psoriasis: is it drug induced or drug aggravated?: understanding pathophysiology and clinical relevance". *The Journal of clinical and aesthetic*. url: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2921739/>.
- Knight, Benjamin P Fairfax and C, Julian (2014). "Genetics of gene expression in immunity to infection". 30: 6371. url: <https://www.sciencedirect.com/science/article/pii/S0952791514000909>.
- Ku, C. S., Loy, E. Y., Salim, A., and genetics, Pawitan-Y. of human (2010). "The discovery of human genetic variations and their use as disease markers: past, present and future". *Journal of human genetics*. doi: 10.1038/jhg.2010.55. url: <http://dx.doi.org/10.1038/jhg.2010.55>.

BIBLIOGRAPHY

- Lande, R., Gregorio, J., Facchinetti, V., Chatterjee, B., Wang, Y. H., Homey, B., Cao, W., Su, B., Nestle, F. O., Zal, T., et al. (2007). "Plasmacytoid dendritic cells sense self-DNA coupled with antimicrobial peptide". *Nature* 449: pp. 564–9. ISSN: 0028-0836. doi: 10.1038/nature06116. URL: <http://dx.doi.org/10.1038/nature06116>.
- Langewouters, A. M. G., Erp, P. E. J., Jong, Emgj, and Kerkhof, P. C. M. (2008). "Lymphocyte subsets in peripheral blood of patients with moderate-to-severe versus mild plaque psoriasis". *Archives of dermatological research* 300: pp. 107–113. doi: 10.1007/s00403-007-0819-9. URL: <http://dx.doi.org/10.1007/s00403-007-0819-9>.
- Leijten, Emmerik F. A., Kempen, Tessa S., Boes, Marianne, Amelsfort, Jocea M. R., Hijnen, Dirkjan, Hartgring, Sarita A. Y., Roon, Joel A. G., Wenink, Mark H., and Radstake, Timothy (2015). "Brief report: enrichment of activated group 3 innate lymphoid cells in psoriatic arthritis synovial fluid". *Arthritis & Rheumatology* 67: pp. 2673–2678. doi: 10.1002/art.39261. URL: <http://dx.doi.org/10.1002/art.39261>.
- Lin, Andrew M., Rubin, Cory J., Khandpur, Ritika, Wang, Jennifer Y., Riblett, MaryBeth, Yalavarthi, Srilakshmi, Villanueva, Eneida C., Shah, Parth, Kaplan, Mariana J., and Bruce, Allen T. (2011). "Mast Cells and Neutrophils Release IL-17 through Extracellular Trap Formation in Psoriasis". *The Journal of Immunology* 187: pp. 490–500. ISSN: 0022-1767. doi: 10.4049/jimmunol.1100123. URL: <http://dx.doi.org/10.4049/jimmunol.1100123>.
- Lizzul, P. F., Aphale, A., Malaviya, R., and Investigative, Sun-Y. of (2005). "Differential expression of phosphorylated NF-B/RelA in normal and psoriatic epidermis and downregulation of NF-B in response to treatment with". *Journal of Investigative*. URL: <https://www.sciencedirect.com/science/article/pii/S0022202X15322983>.
- Mahil, Satveer K., Capon, Francesca, and Barker, Jonathan N. (2016). "Update on psoriasis immunopathogenesis and targeted immunotherapy". *Seminars in Immunopathology* 38: pp. 11–27. ISSN: 1863-2297. doi: 10.1007/s00281-015-0539-8. URL: <http://dx.doi.org/10.1007/s00281-015-0539-8>.
- Maller, J. B., McVean, G., Byrnes, J., Vukcevic, D., Palin, K., Su, Z., Howson, J. M., Auton, A., Myers, S., Morris, A., et al. (2012). "Bayesian refinement of association signals for 14 loci in 3 common diseases". *Nat Genet* 44: pp. 1294–301. ISSN: 1061-4036. doi: 10.1038/ng.2435. URL: <http://www.nature.com/ng/journal/v44/n12/full/ng.2435.html>.
- Marouli, E., Graff, M., Medina-Gomez, C., Lo, K. S., and Nature, Wood-A. R. (2017). "Rare and low-frequency coding variants alter human adult height". *Nature*. doi: 10.1038/nature21039. URL: <http://dx.doi.org/10.1038/nature21039>.
- Marrakchi, S., Guigue, P., and of, Renshaw-B. R. (2011). "Interleukin-36receptor antagonist deficiency and generalized pustular psoriasis". *England Journal of*. doi: 10.1056/NEJMoa1013068. URL: <http://dx.doi.org/10.1056/NEJMoa1013068>.

BIBLIOGRAPHY

- Marshall, C. R., Howrigan, D. P., and Nature, Merico-D. (2017). "Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects". *Nature*. URL: <https://www.nature.com/ng/journal/v49/n1/abs/ng.3725.html>.
- Matesanz, F., Fedetz, M., and of, Barrionuevo-C. (2016). "A splice variant in the ACSL5 gene relates migraine with fatty acid activation in mitochondria". *European Journal of*. doi: 10.1038/ejhg.2016.54. URL: <http://dx.doi.org/10.1038/ejhg.2016.54>.
- Maurano, M. T., Humbert, R., Rynes, E., and, Thurman-R. E. (2012). "Systematic localization of common disease-associated variation in regulatory DNA". doi: 10.1126/science.1222794. URL: <http://dx.doi.org/10.1126/science.1222794>.
- McGonagle, Dennis, Ash, Zoe, Dickie, Laura, McDermott, Michael, and Aydin, Sibel (2011). "The early phase of psoriatic arthritis". *Annals of the Rheumatic Diseases* 70. issn: 0003-4967. doi: 10.1136/ard.2010.144097. URL: <http://dx.doi.org/10.1136/ard.2010.144097>.
- Mease, P. J., Goffe, B. S., Metz, J., VanderStoep, A., and Lancet, Finck-B. (2000). "Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial". *The Lancet*. URL: <https://www.sciencedirect.com/science/article/pii/S0140673600025307>.
- Meglio, Paola, Villanova, Federica, and Nestle, Frank O. (2014). "Psoriasis". *Cold Spring Harbor Perspectives in Medicine* 4. doi: 10.1101/cshperspect.a015354. URL: <http://dx.doi.org/10.1101/cshperspect.a015354>.
- Menon, B., Gullick, N. J., and, Walter-G. J. (2014). "Interleukin17+ CD8+ T cells are enriched in the joints of patients with psoriatic arthritis and correlate with disease activity and joint damage progression". *Arthritis &*. doi: 10.1002/art.38376. URL: <http://dx.doi.org/10.1002/art.38376>.
- Mensah, K. A., Schwarz, E. M., and reports, Ritchlin-C. T. rheumatology (2008). "Altered bone remodeling in psoriatic arthritis". *Current rheumatology reports*. URL: <http://link.springer.com/article/10.1007/s11926-008-0050-5>.
- Menter, Alan, Korman, Neil J., Elmets, Craig A., Feldman, Steven R., Gelfand, Joel M., Gordon, Kenneth B., Gottlieb, Alice, Koo, John, Lebwohl, Mark, Lim, Henry W., et al. (2009). "Guidelines of care for the management of psoriasis and psoriatic arthritis Section 3. Guidelines of care for the management and treatment of psoriasis with topical therapies". *Journal of the American Academy of Dermatology* 60: pp. 643–659. issn: 0190-9622. doi: 10.1016/j.jaad.2008.12.032. URL: <http://dx.doi.org/10.1016/j.jaad.2008.12.032>.
- Moll, J. M. H., in, arthritis, and rheumatism, Wright-V. (1973). "Psoriatic arthritis". *Seminars in arthritis and rheumatism*. URL: <http://www.sciencedirect.com/science/article/pii/0049017273900358>.

BIBLIOGRAPHY

- Nair, R. P., Ruether, A., Stuart, P. E., Jenisch, S., Tejasvi, T., Hiremagalore, R., Schreiber, S., Kabelitz, D., Lim, H. W., Voorhees, J. J., et al. (2008). "Polymorphisms of the IL12B and IL23R genes are associated with psoriasis". *J Invest Dermatol* 128: pp. 1653–61. ISSN: 0022-202X. doi: 10.1038/sj.jid.570125510.1038/sj.jid.5701255. Epub 2008 Jan 24.. URL: <http://dx.doi.org/10.1038/sj.jid.5701255>.
- Nestle, Frank O., Conrad, Curdin, Tun-Kyi, Adrian, Homey, Bernhard, Gombert, Michael, Boyman, Onur, Burg, Gnter, Liu, Yong-Jun, and Gilliet, Michel (2005). "Plasmacytoid dendritic cells initiate psoriasis through interferon- production". *The Journal of Experimental Medicine* 202: pp. 135–143. ISSN: 0022-1007. doi: 10.1084/jem.20050500. URL: <http://dx.doi.org/10.1084/jem.20050500>.
- Nestle, Frank O., Kaplan, Daniel H., and Barker, Jonathan (2009). "Psoriasis". *New England Journal of Medicine* 361: pp. 496–509. doi: doi: 10.1056/NEJMra0804595. URL: <http://www.nejm.org/doi/full/10.1056/NEJMra0804595>.
- Nickoloff, B. J. and Nestle, F. O. (2004). "Recent insights into the immunopathogenesis of psoriasis provide new therapeutic opportunities". *J Clin Invest.* Vol. 113: pp. 1664–75. ISBN: 0021-9738 (Print). doi: 10.1172/jci200422147. URL: <http://dx.doi.org/10.1172/jci200422147>.
- Nickoloff, B. J. and Wrone-Smith, T. (1999). "Injection of pre-psoriatic skin with CD4+ T cells induces psoriasis". *Am J Pathol* 155: pp. 145–58. ISSN: 0002-9440 (Print)0002-9440. doi: 10.1016/s0002-9440(10)65109-7. URL: [http://dx.doi.org/10.1016/s0002-9440\(10\)65109-7](http://dx.doi.org/10.1016/s0002-9440(10)65109-7).
- Noordenbos, T., Yeremenko, N., and, Gofita-I. (2012). "Interleukin17positive mast cells contribute to synovial inflammation in spondylarthritis". *Arthritis & Rheumatism*. doi: 10.1002/art.33396. URL: <http://dx.doi.org/10.1002/art.33396>.
- Okada, Y., Han, B., Tsoi, L. C., and of, Stuart-P. E. (2014). "Fine mapping major histocompatibility complex associations in psoriasis and its clinical subtypes". *The American Journal of*. URL: <https://www.sciencedirect.com/science/article/pii/S0002929714003115>.
- Oliveira, Mfsp, Rocha, B. O., and de, Duarte-G. V. brasileiros (2015). "Psoriasis: classical and emerging comorbidities". *Anais brasileiros de*. doi: 10.1590/abd1806-4841.20153038. URL: <http://dx.doi.org/10.1590/abd1806-4841.20153038>.
- Organization, World Health (2016). *Global report on psoriasis*. Report.
- Perera, G. K., Di Meglio, P., and Nestle, F. O. (2012). "Psoriasis". *Annu Rev Pathol* 7: pp. 385–422. ISSN: 1553-4006. doi: 10.1146/annurev-pathol-011811-132448.
- Picard, C., Dogniaux, S., and of, Chemin-K. journal (2009). "Hypomorphic mutation of ZAP70 in human results in a late onset immunodeficiency and no autoimmunity". *European journal of*. doi: 10.1002/eji.200939385. URL: <http://dx.doi.org/10.1002/eji.200939385>.

BIBLIOGRAPHY

- Picelli, S., Faridani, O. R., Bjorklund, K., and protocols, Winberg-G. (2014). "Full-length RNA-seq from single cells using Smart-seq2". *Nature protocols*. doi: 10.1038/nprot.2014.006. url: <http://dx.doi.org/10.1038/nprot.2014.006>.
- Pivarcsi, A., Sthle, M., and dermatology, Sonkoly-E. (2014). "Genetic polymorphisms altering microRNA activity in psoriasis key to solve the puzzle of missing heritability?" *Experimental dermatology*. doi: 10.1111/exd.12469. url: <http://dx.doi.org/10.1111/exd.12469>.
- Polachek, A., Li, S., care, and, Chandran-V. (2017). "Clinical Enthesitis in a Prospective Longitudinal Psoriatic Arthritis Cohort: Incidence, Prevalence, Characteristics, and Outcome". *Arthritis care &*. doi: 10.1002/acr.23174. url: <http://dx.doi.org/10.1002/acr.23174>.
- Poole, C. D., Lebmeier, M., Ara, R., and Rheumatology, Rafia-R. (2010). "Estimation of health care costs as a function of disease severity in people with psoriatic arthritis in the UK". *Rheumatology*. doi: 10.1093/rheumatology/keq182. url: <http://dx.doi.org/10.1093/rheumatology/keq182>.
- Porcher, R., SaidNahal, R., and, D'Agostino-M. A. (2005). "Two major spondylarthropathy phenotypes are distinguished by pattern analysis in multiplex families". *Arthritis Care &*. doi: 10.1002/art.21070. url: <http://dx.doi.org/10.1002/art.21070>.
- Proksch, E., Brandner, J. M., and Jensen, J. M. (2008). "The skin: an indispensable barrier". *Exp Dermatol* 17: pp. 1063–72. ISSN: 0906-6705.
- Raj, Towfique, Rothamel, Katie, Mostafavi, Sara, Ye, Chun, Lee, Mark N., Replogle, Joseph M., Feng, Ting, Lee, Michelle, Asinovski, Natasha, and Frohlich, Irene (2014). "Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes". *Science* 344: pp. 519–523. url: <http://www.sciencemag.org/content/344/6183/519.short>.
- Reich, K., Krger, K., and Dermatology, MSSNER-R. of (2009). "Epidemiology and clinical pattern of psoriatic arthritis in Germany: a prospective interdisciplinary epidemiological study of 1511 patients with plaque-type psoriasis". *Journal of Dermatology*. doi: 10.1111/j.1365-2133.2008.09023.x. url: <http://dx.doi.org/10.1111/j.1365-2133.2008.09023.x>.
- Schork, N. J., Murray, S. S., Frazer, K. A., and Topol, E. J. (2009). "Common vs. rare allele hypotheses for complex diseases". *Curr Opin Genet Dev* 19: pp. 212–9. ISSN: 0959-437X. doi: 10.1016/j.gde.2009.04.010.
- Stefan, M., Wei, C., Lombardi, A., and Li, C. W. (2014). "Genetic epigenetic dysregulation of thymic TSH receptor gene expression triggers thyroid autoimmunity". *Proceedings of the*. url: <http://www.pnas.org/content/111/34/12562.short>.

BIBLIOGRAPHY

- Strange, A., Capon, F., Spencer, C. C., Knight, J., Weale, M. E., Allen, M. H., Barton, A., Band, G., Bellenguez, C., Bergboer, J. G., et al. (2010). "A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1". *Nat Genet* 42: pp. 985–90. ISSN: 1061-4036. doi: 10.1038/ng.69410.1038/ng.694. Epub2010Oct17.. URL: <http://dx.doi.org/10.1038/ng.694>.
- Sun, Lian-Dang, Cheng, Hui, Xue-Qin, Yang, Jian-Zhong, Zhang, Ai, E. Xu, Ri-Na, Wu, Li-Min, Xu, Lin, Peng, Cynthia, A. Helms, Yun-Qing, Ren, et al. (2010). "Association analyses identify six new psoriasis susceptibility loci in the Chinese population". *Nature Genetics* 42: pp. 1005–1009. ISSN: 1061-4036. doi: doi: 10.1038/ng.690. URL: <http://www.nature.com/ng/journal/v42/n11/abs/ng.690.html>.
- Trynka, G. and Raychaudhuri, S. (2013). "Using chromatin marks to interpret and localize genetic associations to complex human traits and diseases". *Current opinion in genetics & development* 23: pp. 635–641. URL: <http://www.sciencedirect.com/science/article/pii/S0959437X13001445>.
- Tsoi, L. C., Spain, S. L., Knight, J., Ellinghaus, E., Stuart, P. E., Capon, F., Ding, J., Li, Y., Tejasvi, T., Gudjonsson, J. E., et al. (2012). "Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity". *Nat Genet* 44: pp. 1341–8. ISSN: 1061-4036. doi: 10.1038/ng.2467.
- Vence, L., Schmitt, A., and Medical, Meadows-C. E. (2015). "Recognizing guttate psoriasis and initiating appropriate treatment". *West Virginia Medical*. URL: <http://go.galegroup.com/ps/i.do?id=GALE%7CA422778063&sid=googleScholar&v=2.1&it=r&linkaccess=fulltext&issn=00433284&p=AONE&sw=w>.
- Villanova, Federica, Meglio, Paola, and Nestle, Frank O. (2013). "Biomarkers in psoriasis and psoriatic arthritis". *Annals of the rheumatic diseases* 72. doi: 10.1136/annrheumdis-2012-203037. URL: <http://dx.doi.org/10.1136/annrheumdis-2012-203037>.
- Wikramanayake, T. C., Stojadinovic, O., and Tomic-Canic, M. (2014). "Epidermal Differentiation in Barrier Maintenance and Wound Healing". *Adv Wound Care (New Rochelle)*. Vol. 3: pp. 272–280. ISBN: 2162-1918 (Print) 2162-1918 (Linking). doi: 10.1089/wound.2013.0503.
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., et al. (2010). "Common SNPs explain a large proportion of the heritability for human height". *Nat Genet* 42: pp. 565–9. ISSN: 1061-4036. doi: 10.1038/ng.608.
- Yin, Xianyong, Low, Hui Q., Wang, Ling, Li, Yonghong, Ellinghaus, Eva, Han, Jiali, Estivill, Xavier, Sun, Liangdan, Zuo, Xianbo, Shen, Changbing, et al. (2015). "Genome-wide meta-analysis identifies multiple novel associations and ethnic heterogeneity of psoriasis susceptibility". *Nature communications* 6: pp. 1–11. ISSN: 2041-1723. doi: 10.1038/ncomms7916. URL: <http://dx.doi.org/10.1038/ncomms7916>.

BIBLIOGRAPHY

Zhang, X. J., Huang, W., Yang, S., Sun, L. D., and Zhang, F. Y. (2009). "Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21". *Nature* 2: pp. 205–210. doi: 10.1038/ng.310. URL: <http://dx.doi.org/10.1038/ng.310>.