



# 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**

<b>Title</b>	<b>Year</b>
Authors	

# **Associated Publications**

## **Title**

Journal

Authors

# **Other Publications**

## **Title**

Journal

Authors

# Contents

<b>Abstract</b>	i
<b>Acknowledgements</b>	ii
<b>Declarations</b>	iii
<b>Submitted Abstracts</b>	iv
<b>Associated Publications</b>	v
<b>Contents</b>	vi
<b>List of Figures</b>	vii
<b>List of Tables</b>	viii
<b>Abbreviations</b>	ix
<b>1 Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis</b>	1
1.1 Introduction . . . . .	1
1.1.1 The systemic and skin-specific components in psoriasis . .	1
1.1.2 The personalised epigenome in disease . . . . .	2
1.1.3 Transcriptional profiles in psoriasis . . . . .	4
1.1.4 Chromatin accessibility, gene expression and genetic variability . . . . .	6
1.1.5 Fine-mapping using summary stats . . . . .	7
1.1.6 Aims . . . . .	8
1.2 Results . . . . .	8
1.2.1 Psoriasis and healthy controls: cohort description and datasets . . . . .	8
1.2.2 Exploring differences in the chromatin accessibility landscape of primary immune cell types in psoriasis . . . . .	12
1.2.3 Differential gene expression in psoriasis circulating immune cells . . . . .	28
1.2.4 RNA-seq in epidermis from psoriasis patients . . . . .	42
1.2.5 Comparison of the systemic and tissue-specific gene expression signatures in psoriasis . . . . .	52
1.2.6 Integration of chromatin accessibility and expression data in circulating immune cells . . . . .	54
1.2.7 Fine-mapping . . . . .	56
1.3 Discussion . . . . .	56

## **CONTENTS**

---

<b>Appendices</b>	<b>57</b>
<b>A Appendices</b>	<b>59</b>
A.1 Additional tables . . . . .	59
A.1.1 Chapter 5 Tables . . . . .	59
A.1.2 Chapter 4 Tables . . . . .	65
A.1.3 Chapter 5 Tables . . . . .	65
A.2 Additional figures . . . . .	65
A.2.1 Chapter 3 Figures . . . . .	65
A.2.2 Chapter 4 Figures . . . . .	65
A.2.3 Chapter 5 Figures . . . . .	67

# List of Figures

1.1	Quality control evaluation of the H3K27ac ChIPm libraries in immune cells isolated from psoriasis and control samples. . . . .	14
1.2	PCA analysis and chromatin annotation states of the H3K27ac enriched sites in four immune primary cell types from psoriasis and healthy control samples. . . . .	15
1.3	Epigenetic landscape at the chr3:121,675,048-121,677,505 enhancer in circulating CD14 <sup>+</sup> monocytes from psoriasis patients and healthy controls. . . . .	17
1.4	Quality control assessment of the ATAC-seq libraries generated from immune cells in psoriasis and control samples. . . . .	20
1.5	K-means clustered heatmap and conserved TFBS enrichment analysis in the consensus ATAC-seq regions identified in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> cells from the patients and controls cohort. . . . .	22
1.6	Epigenetic landscape at two ATAC-seq differential accessible regions between patients and controls in CD8 <sup>+</sup> cells. . . . .	25
1.7	Epigenetic landscape at the chr3:121,675,048-121,677,505 enhancer in circulating CD14 <sup>+</sup> monocytes from psoriasis patients and healthy controls. . . . .	27
1.8	Mapping rate and total reads after filtering (million) mapping to Ensembl genes in all the RNA-seq samples from psoriasis patients and controls in four cell types. . . . .	29
1.9	PCA analysis and sample distance heatmap with k-means clustering illustrating the sample variability based on the gene expression profiles. . . . .	30
1.10	Magnitude and significance of the gene expression changes between psoriasis patients and healthy controls in four immune cell types. . . . .	32
1.11	RNA-seq expression levels of the lncRNA <i>HOTAIRM1</i> and its putative target UPF1 in psoriasis and healthy controls CD14 <sup>+</sup> monocytes. . . . .	35
1.12	Mapping of the DEGs identified in CD8 <sup>+</sup> cells between psoriasis patients and healthy controls onto the TNF- $\alpha$ and the chemokine signalling pathways. . . . .	40
1.13	Mapping quality control and PCA analysis for the RNA-seq data in the uninvolved and lesional epidermis from psoriasis patients. .	43
1.14	Mapping of the DEGs between lesional and uninvolved epidermis from psoriasis patients onto the HIF-I signalling pathway. . . . .	49
1.15	Mapping of the DEGs between lesional and uninvolved epidermis from psoriasis patients onto the NOD-like signalling pathway. . . . .	51

## LIST OF FIGURES

---

1.16 Epigenetic landscape at the chr3:121,675,048-121,677,505 enhancer in circulating CD14 <sup>+</sup> monocytes from psoriasis patients and healthy controls. . . . .	55
A.1 FAST-ATAC and Omni-ATAC NHEK tapestation profiles. . . . .	61
A.2 Assessment of TSS enrichment from ATAC-seq and FAST-ATAC in healthy and psoriasis skin biopsies samples. . . . .	62
A.3 Percentage of MT reads in the ATAC-seq libraries generated in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> isolated from psoriasis patients and healthy controls. . . . .	65
A.4 Genomic annotation of the consensus master list of ATAC-seq enriched sites built for downstream differential chromatin accessibility analysis in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> . . . . .	66
A.5 PCA analysis illustrating batch effect in ATAC-seq and RNA-seq samples. . . . .	66
A.6 Permutation analysis SF vs PB in CD14 <sup>+</sup> ,CD4m <sup>+</sup> ,CD8m <sup>+</sup> and NK. . . . .	67
A.7 Heatmap for the top 20 marker genes of the CC-mixed and CC-IL7R CD14 <sup>+</sup> monocytes subpopulations. . . . .	68
A.8 Identification of the CD14 <sup>+</sup> monocytes populations from bulk SFMCs and PBMCs using scRNA-seq transcriptomes. . . . .	68

# List of Tables

1.1	Summary table of the most comprehensive transcriptional studies in psoriasis skin biopsies and their key findings. . . . .	5
1.2	Description and metadata of the psoriasis patients cohort. . . . .	9
1.3	Description of the healthy control cohort. . . . .	11
1.4	Datasets generated for the psoriasis and control cohort samples. . . . .	12
1.5	Summary results from the differential H3K27ac analysis between psoriasis patients and healthy controls in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> cells. . . . .	16
1.6	Summary results from the differential chromatin accessibility analysis between psoriasis patients and healthy controls in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> cells. . . . .	24
1.7	Summary results from the DGE analysis between psoriasis patients and healthy controls in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> cells. . . . .	31
1.8	Overlap between putative psoriasis GWAS genes and the reported significantly DEGs in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> cells. . . . .	34
1.9	Most relevant pathways enriched for DEGs between psoriasis patients and healthy controls in CD14 <sup>+</sup> monocytes and CD8 <sup>+</sup> cells. . . . .	38
1.10	Summary results of the DGE analysis between uninvolved and lesional psoriatic epidermal biopsies. . . . .	44
1.11	Most relevant pathways enriched for DEGs between lesional and uninvolved epidermis isolated from psoriasis patients skin biopsies. . . . .	48
1.12	Overlap of the dysregulated genes in the four circulating cell types (psoriasis patients versus controls) and the DEGs in psoriasis patients skin biopsies (lesional versus uninvolved). . . . .	53
A.1	Enrichment of ATAC-seq reads across the TSS for the CD14 <sup>+</sup> monocytes and CD4 <sup>+</sup> samples fresh, frozen and fixed. . . . .	59
A.2	Evaluation of ChiPm library complexity for the psoriasis and control chort 1B ChiPm assay. . . . .	60
A.3	Additional enriched pathways for DEGs between psoriasis and healthy controls in CD14 <sup>+</sup> monocytes and CD8 <sup>+</sup> cells. . . . .	62
A.4	Additional enriched pathways for DEGs between lesional and uninvolved epidermis isolated from psoriasis patients skin biopsies. . . . .	63
A.5	Additional enriched pathways for the DEGs between SF and PB CD14 <sup>+</sup> monocytes from the CC-mixed and CC-IL7R subpopulations. . . . .	64

# Abbreviations

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

# Chapter 1

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

### 1.1 Introduction

#### 1.1.1 The systemic and skin-specific components in psoriasis

The kin lesions are the main disease manifestation reflecting dysregulation of the innate and immune response in psoriasis as the result of genetic and environmental factors. In addition to KCs, the relevance of other circulating immune cells that are recruited to the site of inflammation have been previously reviewed in the Chapter ???. In this lines, some report cases studies in humans have demonstrated the ability of bone-marrow transplants of inducing and terminating psoriasis (Gardembas1990; Eedy et al. 1990). Moreover, the importance of T cell recruitment and activation in psoriasis has been demonstrated in xenotransplantation models of psoriasis and studies in human patients (Wrone-Smith1996; Baker1984; Nickoloff and Wrone-Smith 1999). It is well accepted that recruitment of T cells to the inflammed skin is mediated by activation of endothelial cells amongst others (Meglio et al. 2014). Moreover, psoriasis patients have demonstrated increased risk to develop joint inflammation following skin disease in the form of PsA as well as other co-

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

morbidity including CVD (Ibrahim et al. 2009; Shapiro et al. 2007). All these findings highlights the systemic component of psoriasis and reinforces the importance of cell type specific characterisation of relevant circulating immune cells to better understand the disease pathophysiology.

### **1.1.2 The personalised epigenome in disease**

The technical revolution in the epigenetic field has opened an avenue to ascertain the epigenome profile of clinical samples from different diseases and cell types, helping in the interpretation and understanding of non-coding GWAS variants. In terms of mapping the chromatin landscape, ATAC-seq and ChIPm have enabled the interrogation of chromatin accessibility, histone modifications and TF binding using few thousand cells and have opened the door to map the regulatory landscape in a wide range of cell types and tissues obtained from valuable clinical samples (**Buenrostro2013; Schmidl2015**).

Profiling the chromatin accessibility in the cell populations isolated from the tissues of interest provides details about the molecular programming of cells and can also inform about the location and status of cis-regulatory elements. The various ATAC-seq available protocols have been used to conduct systematic studies to identify inter- and intra- individual differences and pathological changes in epigenetic marks. For example differential analysis in B cells isolated from SLE and healthy controls has revealed changes in chromatin accessibility nearby genes involved in B cells activation as well as enrichment for TFBS potentially regulating pathogenic processes (**Scharer2016**). Similarly a study in age-related macular degeneration (AMD) has identified the retina epithelium as the main tissue driving disease onset though global loss of chromatin accessibility compared to healthy tissue (**Wang2018**). Interestingly, Treatment of retinal pigmented epithelium cells simulating cigarette smoking recapitulated the chromatin accessibility changes found in AMD, highlighting the ability of

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

the chromatin landscape to reflect environmental stress in complex diseases. Mapping of the regulatory landscape in acute myeloid leukemia cells using Fast-ATAC allowed to reveal patient-specific alterations in a set of regulatory elements and tumor cells found across different hematopoietic stages (**Corces2016**).

The study of particular histone marks modifications in combination with the chromatin accessibility profile can reveal additional functional information to understand the cell type specific regulatory landscape. ChIPm has been successfully used to identify, for example, subtype-specific epigenome signatures in chronic lymphocytic leukaemia (**Rendeiro2015**). In terms of relevant histone modifications, acetylation increases chromatin accessibility and is associated with active transcription due to the added negative charges and reduced affinity of these proteins for the DNA (**Creyghton2010**).

A particularly relevant histone acetylation in the study of complex diseases is the active enhancer H3K27ac mark, as GWAS SNPs are mostly located in intergenic regions that may act as gene expression regulatory elements. Precisely, a study using 75 SNPs tagging all the psoriasis GWAS loci showed enrichment of those SNPs for H3K27ac in twenty-six relevant cells and/or tissues (**Lin2018**). Interestingly, almost half of all the SNPs in LD with psoriasis GWAS tag SNPs and overlapping at least one of five core histone mark in total CD4<sup>+</sup> T cells were annotated as enhancers. A relevant example of enhancer profiling through the characterisation of H3K27ac has been conducted in the autoimmune disease juvenile idiopathic arthritis showing a disease H3K27ac-specific super-enhancer (spanning up to 50Kb) signature in SF mCD4<sup>+</sup> cells (**Peeters2015**). Importantly, Hi-C combined with H3K27ac ChIP has allowed to identify specific enhancerpromoter interactions in primary human cells and revealed the ability of GWAS SNPs to interact with multiple distal gene promoters (**Mumbach2017**). In addition to this, inhibitors of the histone de-acetylases (HDACs), involved

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

in removal of acetyl groups from histone proteins, are being investigated as therapeutic agents for RA and SLE, amongst others ??.

In terms of personalised epigenome studies in psoriasis, a epigenomic-wide association study (EWAS) investigating DNA methylation has been conducted (**Zhou2016**). Interestingly, this study revealed nine disease-associated differentially methylated sites underlying disease status and environmental factors but no overlap with genetic effects. Regarding comparative analysis of chromatin accessibility landscape and histone marks modifications between psoriasis and healthy individuals, no studies have yet been conducted neither in circulating immune cells or skin biopsies.

### **1.1.3 Transcriptional profiles in psoriasis**

#### **Trancriptomics in psoriatic skin**

Characterisation of the patients transcriptional profile in complex diseases have been performed to better understand the disease pathophysiology and also to inform the role of genetic variation and susceptibility in regulating gene expression. In psoriasis, the majority of the transcriptional studies have been performed for inflamed skin (lesional) using pre-lesional (uninvolved) skin adjacent to the lesion as the internal control (Table ??). Other studies have also incorporated healthy control skin biopsies in order to determine the extent of dysregulation of the transcriptomic profile prior to lesion development in the uninvolved skin (Table ??).

Discrepancies in the results have been found between microarray and RNA-seq studies, which have demonstrated greater sensibility of RNA-seq as well as the ability to identify other transcript species including lncRNAs in an unbiased way (**Jaabari2011**; Li et al. 2014). Interestingly, lncRNAs have proven an important role in the dysregulation of psoriasis skin transcriptome and

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

Author and year	Sample type and size	Technology	Description
(Jaabari2011) (Li et al. 2014)	WB (L=3, U=3) WB (L=92, C=82)	RNA-seq and microarray RNA-seq and microarray	Technology co-regulates lncRNAs
(Keermann2015)	WB (L=12, U=12, C=12)	RNA-seq	Dormant psoriasis and IL-36 expression
(Tsoi2015) (Swindell et al. 2015)	WB (L=97, U=29, C=90) WB (L=14, U=14)	RNA-seq RNA-seq and mass-spectrometry	Psoriatic skin-specific 209 co-regulated genes
(Swindell2017)	CK (L=4, U=4, C=4)	RNA-seq	Decreased differentiation signature
(Tervaniemi2016)	EpB (L=6, U=6, C=9)	RNA-seq	NOD-like autoimmunity pathways in psoriasis

**Table 1.1: Summary table of the most comprehensive transcriptional studies in psoriasis skin biopsies and their key findings.** WB= whole biopsy; EpB=epidermal biopsy; CK=cultured KCs; C=control; L= psoriatic lesional skin; U=psoriatic uninvolved skin.

approximately 1,000 were found to be differentially expressed between lesional and uninvolved skin, including *de novo* characterised species with greater skin-specificity patterns of expression (Tsoi2015). Interestingly, discrepancies regarding the transcriptional similarities between normal and uninvolved skin have been identified. Some studies have found around 2,500 DEGs when compared to controls skin, whereas others have only found 3 transcripts differentially expressed between the same groups, likely due to different FC filtering criteria (Keermann2015; Tsoi2015). The correspondence between DGE and protein abundance have also been investigated in psoriatic skin, unveiling that only a 5% of the dysregulated transcripts presented a similar trend at the protein level (Swindell et al. 2015).

Regarding the type of biopsy, the majority of the transcriptional studies have been performed in whole skin biopsies containing a mix of cells from the epidermis, dermis, basal layer, muscle and adipose tissue (Table ??). Lately, psoriatic cultured KCs (from lesional and uninvolved biopsies) and epidermis

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

KCs through split-thickness skin graft sampling have revealed differences in gene expression and enriched functional pathways when compared to the previous studies (**Swindell2017; Tervaniemi2016**). These results have reinforced the importance of homogenous cell type samples in order to better dissect the altered biological processes in each of the relevant cell types at the site of inflammation contributing to the development of psoriasis.

Interestingly, (**Coda2012**) explored the overlap between the DEGs in PBMCs (psoriasis versus controls) and the dysregulated genes when comparing lesional and uninvolved skin. The results revealed a limited overlap with more than 50% of the common genes presenting opposite directions of modulation in the two tissues. At the cell type specific level, some studies have been performed in *in vitro* cultured and stimulated T cells and monocytes (**Palau2013; Jung2004**). For instance, Palau and colleagues conducted *in vitro* CD4<sup>+</sup> and CD8<sup>+</sup> T cell activation using anti-CD3 and anti-CD28 Abs and found 42 DEGs enriched for cytokine and IFN ( $\alpha$ ,  $\beta$  and  $\gamma$ ) signalling pathways.

Understanding of disease-specific systemic gene expression dysregulation has been approached through comparison with other chronic inflammatory diseases. For example, gene expression from psoriasis, IBD and RA PBMCs for 96 immune-relevant genes were compared and identified five DEGs unique for psoriasis as well as a common inflammatory signature consisting of five dysregulated genes in all three diseases (**Mesko2015**).

### **1.1.4 Chromatin accessibility, gene expression and genetic variability**

As previously explained, accessible chromatin involves a less compact form of DNA more likely to be bound by TFs and other co-regulatory proteins. As a result, chromatin accessibility is used as a proxy to tag genomic loci involved in gene expression regulation or undergoing transcription. Moreover, the overlap

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

with open chromatin is one of the markers used to infer putative functional relevance of GWAS SNPs. Cell type specific coordinated changes in the chromatin landscape and gene expression are required to orchestrate an appropriate immune response (**Goodnow2005**). Integration of ATAC-seq data and gene expression in pancreatic islets revealed chromatin accessibility to be a better proxy for gene expression in  $\alpha$ - compared to  $\beta$  cells. Furthermore, moderate correlation was found between the intensity in chromatin accessibility and gene expression in retina and pigmented epithelium retina from AMD patients (**Wang2018**). In the context of genetic variability, the correlation between chromatin accessibility and gene expression in homeostatic and stimulated conditions has been addressed by integrating eQTL and chromatin accessibility QTLs (ca-QTLs). For example, a study in human iPS derived macrophages demonstrated that many stimulus-dependent eQTLs drive changes in gene expression by indirect of chromatin accessibility in the naïve state through enhancer priming (**Alasoo2018**).

### **1.1.5 Fine-mapping using summary stats**

As commented in the Chapter ??, the generation of cell type specific epigenetic maps can be used to inform statistical fin-mapping in the effort to identify putative causal SNPs to undergo functional validation. Integration of Bayesian fine-mapping for X complex immune diseases performed by Farh and colleagues demonstrated the greatest enrichment of fine-mapped causal variants to be immune cell types enhancer elements, particularly in activated cell types (**Farh2015**). Particularly, in this study psoriasis PICS showed enrichment for Th0 naïve CD4 $^{+}$  T cells followed by Th-1, Th-2 and Th-17 CD4 $^{+}$  subsets.

Traditional Bayesian fine-mapping requires the use of genotyping data from the GWAS cohorts to perform genotypes phasing and imputation using reference

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

panels prior to the association analysis and the calculation of PP. Availability of this genotyping data, commonly due to ethical reasons, can represent a limitation when intending to perform this type of analysis. Since summary statistics data from GWAS studies are mostly available, methods like DIST have been developed to impute summary statistics instead of genotypes for the unmeasured SNPs in the study (Lee2013). DIST uses the z-score statistics from the genotyped SNPs in the study to then impute the z-scores of the missing variants based on the LD correlation from reference panels such as 1000 Genomes Project. In addition to this, Bayesian fine-mapping using summary statistics data and functional annotation as a prior in the model have also been developed. A relevant example for psoriasis is the development of risk variant inference using epigenomic reference annotation (RiVIERA), which has been applied to perform fine-mapping for the Immunochip GWAS associated loci incorporating the 43 most enriched annotation features in psoriasis from ENCODE and Epigenome Roadmap (Li2016).

### **1.1.6 Aims**

## **1.2 Results**

### **1.2.1 Psoriasis and healthy controls: cohort description and datasets**

PB samples were collected from a cohort of psoriasis patients and healthy individuals in order isolate four relevant immune cells types (CD14<sup>+</sup> monocytes, total CD4<sup>+</sup>, total CD8<sup>+</sup> and CD19<sup>+</sup>) and perform ATAC-seq, RNA-seq and ChIPm analysis. Additionally, the epidermis from paired uninvolved and lesional skin biopsies collected from three psoriasis patients were processed downstream for RNA-seq analysis.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

---

### Cohorts description

Psoriasis patients PB was collected as detailed in Chapter ?? from a total of eight psoriasis patients, six males and two females (Table 1.2).

Sample ID	Sex	Age at diagnosis	Disease duration (months)	PASI	Nails affected	Family history
<b>Cohort 1A</b>						
PS1011	Male	55	420	11	Yes	No
PS2014	Female	65	588	17	No	No
PS2015	Male	56	384	5	Yes	No
PS2016	Male	40	180	10	No	No
<b>Cohort 1B</b>						
PS2000	Male	61	156	10	No	Yes
PS2001	Male	56	432	10	Yes	No
PS2314	Male	42	120	6.5	Yes	No
PS2319	Female	64	372	10.2	No	Yes
Average	-	55	331.5	10	-	-

**Table 1.2: Description and metadata of the psoriasis patients cohort.** For each of the individuals information related to sex, age at the time of sampling, disease duration, PASI score, affection of nails and family history has been included. Patients are divided in cohort 1A and cohort 1B based on the batch of ATAC-seq and RNA-seq processing. PASI evaluates the percentage of affected area and the severity of redness, thickness and scaling for four body locations (as detailed in Table ??). For each of the locations the test quantifies the percentage of affected area and the severity of those three clinical signs (redness, thickness and scaling). The percentage of affected area is scored in a scale 1 to 6 (1=1-9%; 2=10-29%; 3=30-49%; 4=50-69%; 5=70-89%; 6=90-100%) and the severity of the three clinical signs in a scale from 0 to 4 (from none to maximum). A combined score for each of the body regions is calculated as the sum of the clinical signs severity scores for that region multiplied by score of that percentage affected area and the proportion of body surface represented by that body region (0.1 for head and neck, 0.2 for upper limbs, 0.3 for trunk and 0.4 for lower limbs. The final PASI score is the addition of each of those scores for each body region. PASI ranges from 0 (no disease) to 72 (maximal disease severity).

The average age of the cohort was 55 years old and the average of disease duration 331.5 months. All the patients presented active skin disease and none of them had reported joint affection at the time of sample collection. Disease severity was quantified using the PASI score, previously reviewed in Chapter

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

??, being the average cohort score 10. Currently, PASI thresholds to define mild and moderate-to-severe lack of consensus. A review study regarding the use of PASI as an instrument to determine disease severity of chronic-plaque psoriasis have suggested to consider psoriasis as moderate when PASI ranges between 7 to 12 and severe for PASI>12 (**Schmitt2005**). On the other hand, NICE and other studies had defined psoriasis as severe based on PASI $\geq$ 10 (**Woolacott2006**; Finlay 2005). In this cohort, six out of ten patients had PASI $\geq$ 10, being eligible to be categorised as severe psoriasis. Only two of them presented PASI<7 showing a mild psoriasis phenotype. All patients were naïve for biologics therapies. PS2319 was currently on MTX therapy and the remaining patients had only been treated occasionally with topical steroids or UVB therapy. Interestingly, PS2014 presented the most severe PASI score (17) and was a non-responder to MTX. Patients PS1011, PS2015, PS2001 and PS2314 presented nails pitting, which has been defined as one of the markers for increased risk of developing joint affection and PsA (**Moll1976; McGonagle; 2011; Griffiths and Lancet 2007**). Psoriasis family history was reported by PS2000 and PS2319, which could indicate those individuals are HLA-C\*06:02 positive. In addition to the psoriasis samples, PB blood was collected from ten sex and age-matched healthy individuals (Table 1.3)

For both, patients and controls, the subdivision in cohort 1A and cohort 1B relates to the batches in which ATAC-seq RNA-seq samples were processed and sequenced.

### **Datasets**

For both cohorts, ATAC-seq and RNA-seq data were generated from CD14 $^{+}$  monocytes, total CD4 $^{+}$ , total CD8 $^{+}$  and CD19 $^{+}$  cells (Table 1.4). For cohort 1A ATAC-seq data was generated using the standard ATAC-seq protocol from Buenrostro *et al.*, 2013, which was replaced by the FAST-ATAC method (**Corces2016**), according to the improvements of this protocol as explained in

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

---

Sample ID	Sex	Age
<b>Cohort 1A</b>		
CTL1	Male	36
CTL2	Male	53
CTL3	Male	34
CTL4	Female	46
CTL5	Male	42
<b>Cohort 1B</b>		
CTL6	Male	31
CTL7	Male	57
CTL8	Female	50
CTL9	Male	50
CTL10	Male	67
Average	-	46.6

**Table 1.3: Description of the healthy control cohort.** Controls are divided in cohort 1A and cohort 1B based on the batch of ATAC-seq and RNA-seq processing, similarly to the psoriasis patients samples.

Chapter ???. Additionally, samples from cohort 1B were also processed to assess differences in H3K27ac modification between patients and controls using ChIPm (Table 1.4).

For three of the psoriasis patients (PS2014, PS2015 and PS2016) paired biopsies from lesional and uninvolved skin were collected and the epidermal sheets were isolated to perform RNA-seq differential analysis (Table 1.4). This should be considered as a pilot study aiming to refine the previous RNA-seq studies performed in whole skin biopsies, with a more heterogeneous cell type composition compared to epidermis, which could not be expanded due to time and cost constructions.

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

<b>Technique</b>	<b>Cohort or samples ID</b>	<b>Sample size (patient/control)</b>
ATAC-seq	Cohort 1A and 1B	8/10
RNA-seq	Cohort 1A and 1B	8/10
ChIPm	Cohort 1B	4/4
Skin RNA-seq	PS1011, PS2015 and PS2016	3/0

**Table 1.4: Datasets generated for the psoriasis and control cohort samples.** Cohort identity includes both, patients and controls. The skin RNA-seq samples include lesional and uninvolved paired-skin biopsies from each of the three individuals.

### **1.2.2 Exploring differences in the chromatin accessibility landscape of primary immune cell types in psoriasis**

In order to assess the chromatin accessibility landscape psoriasis and control individuals, ChIPm and ATAC-seq were performed in PB isolated CD14<sup>+</sup> monocytes, total CD4<sup>+</sup>, total CD8<sup>+</sup> and CD19<sup>+</sup> B cells. ChIPm data was only generated in the patients and controls from cohort 1B whereas ATAC-seq was performed in all the patients and controls samples from cohorts 1A and 1B (Table 1.4).

#### **Assessment of changes in the enhancer mark H3K27ac in psoriasis immune cells**

A total of 32 ChIPm libraries from four patients and four controls in four cell types were sequences and reads were filtered as detailed in Chapter ???. After filtering, the median of total number of million reads ranging between 46.9 and 60.5 millions compliant with the approximately 40 million of total reads recommended by ENCODE (Figure 1.1 a). As part of the quality control, library complexity for each of the samples was measures by the NRF and the PCR bottlenecking coefficients PBC1 and PBC2. According to the ENCODE standards, most of the libraries had appropriate complexity and moderate to mild bottlenecking (Table A.2). Nonetheless, the CD8<sup>+</sup> CTL7 and the CD19<sup>+</sup>

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

PS2000 and PS2314 were the libraries failing the recommended complexity NRF values and also the ones with more severe PCR bottlenecking according to the PBC1 coefficient threshold. These observations were consistent with the greater number of duplicated reads identified in this libraries compared to the rest (>50% of the total sequenced reads) and consequently the lower number of reads after filtering (Figure 1.1 a).

Regarding the enrichment of ChIPm signal for each of the libraries, cross-correlation analysis was performed to determine the NSC and RSC coefficients which provide a measure for the signal-to-noise ratios in the samples. All the ChIPm libraries presented appropriate signal-to-noise following ENCODE standards (Landt2012), with NSC and RSC values equal or greater than 1.05 and 0.8, respectively (Figure 1.1 b and c). Interestingly, the CD14<sup>+</sup> monocytes and total CD4<sup>+</sup> ChIPm libraries had lower signal enrichment compared to the total CD8<sup>+</sup> and CD19<sup>+</sup> libraries, which correlates with the cell type grouping for the sample processing. Altogether, CD8<sup>+</sup> CTL7 and the CD19<sup>+</sup> PS2000 and PS2314 H3K27ac ChIPm libraries were removed for downstream analysis.

PCA analysis using a combined master list of the H3K27ac enriched sites in patients and controls for all four cell types (excluding the aforementioned low quality samples) confirmed the ability of this data to recapitulate the cell type specific epigenetic landscape of this enhancer mark and reinforced the appropriate quality of the data (Figure 1.2 a). When performing PCA analysis by cell type, the PS2314 CD8<sup>+</sup> library appeared as an outlier compared to the rest of CD8<sup>+</sup> H3K27ac ChIPm patients and control libraries and was also removed for downstream analysis (data not shown).

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

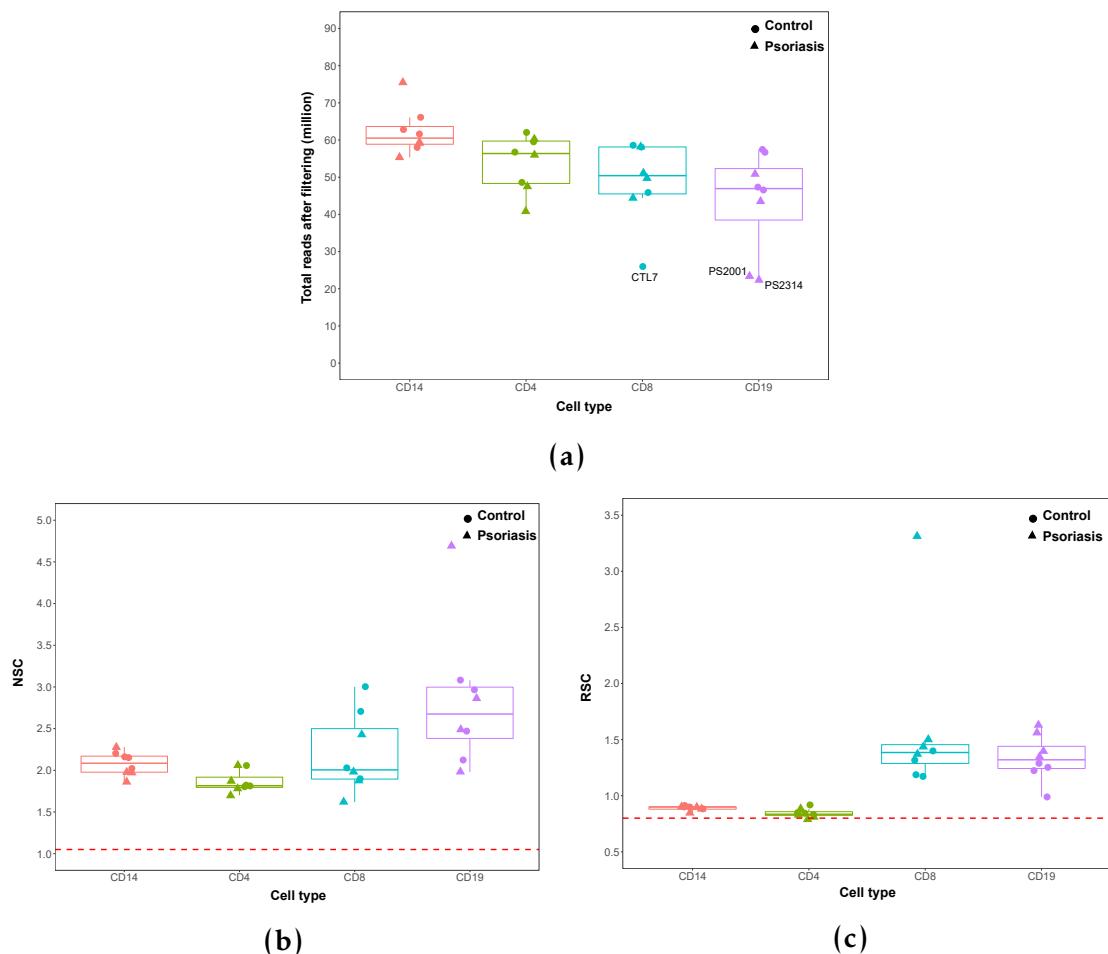
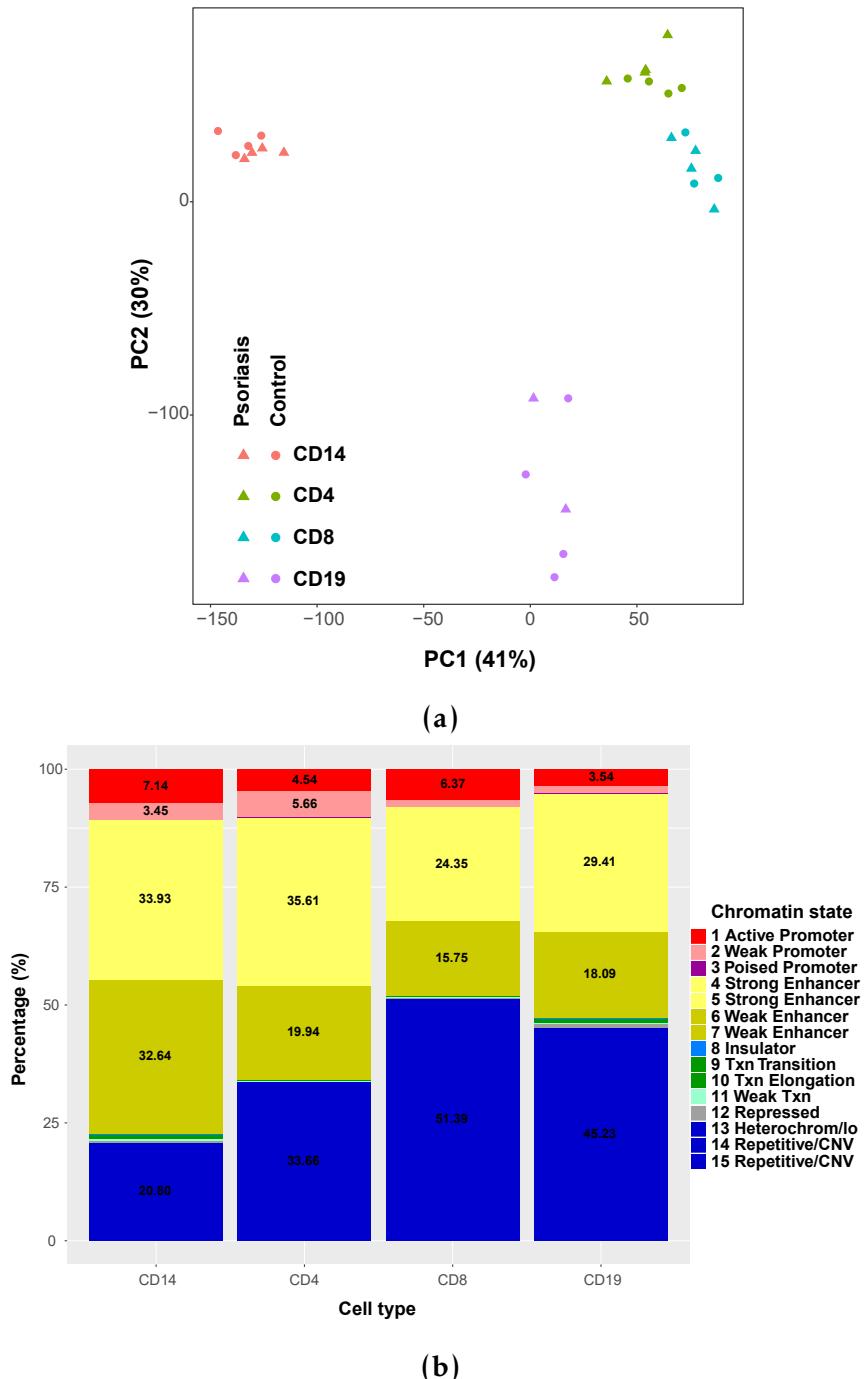


Figure 1.1: Quality control evaluation of the H3K27ac ChIPm libraries in immune cells isolated from psoriasis and control samples.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.2: PCA analysis and chromatin annotation states of the H3K27ac enriched sites in four immune primary cell types from psoriasis and healthy control samples.**

a) PCA analysis was performed using the normalised counts across a consensus master list of the combined H3K27ac enriched regions in psoriasis patients and healthy control samples across CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup> cells. b) Annotation of the H3K27ac list of consensus enriched sites built by DiffBind for each cell was performed using the appropriate cell type specific RoadMap chromatin segmentation maps. Results are expressed as the percentage of regions annotated with a particular chromatin state over the total number of H3K27ac enriched sites in each of the four cell type master lists.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

---

Differential analysis for H3K27ac was performed for each cell type between psoriasis and healthy control samples using DiffBind. The consensus merged list of H3K27ac sites assembled by this algorithm to perform the differential analysis (as explained in Chapter ??) showed a great percentage of those sites to be annotated as heterochromatin or repetitive (Figure 1.2 b). These percentages ranged from 20.8 to 51.39% in CD14<sup>+</sup> monocytes and total CD8<sup>+</sup> cells, respectively, which are less likely to be relevant since H3K27ac is a histone modification mainly enriched at enhancers. When restricting the differential analysis to those regions annotated as enhancers (weak and strong), CD14<sup>+</sup> monocytes appeared as the cell type with the greatest number of differentially modified enhancers (8 hits), followed by CD4<sup>+</sup> (4 hits) and CD8<sup>+</sup> (1 hit) (Table 1.5).

Cell type	Master list size genome-wide/enhancers	Differential regions genome-wide/enhancers
CD14 <sup>+</sup>	99,862/60,962	15/8
CD4 <sup>+</sup>	110,353/56,282	0/4
CD8 <sup>+</sup>	137,194/51,607	8/1
CD19 <sup>+</sup>	199,014/88,722	12/0

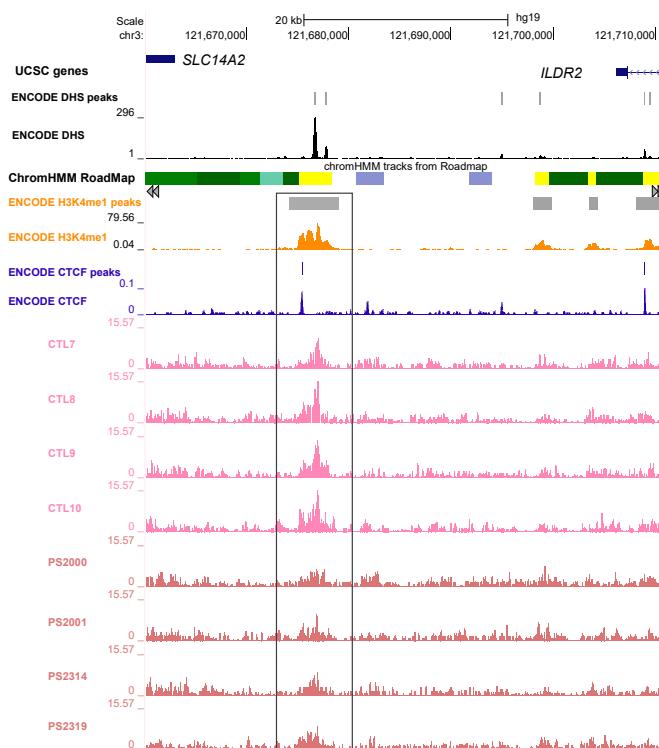
**Table 1.5: Summary results from the differential H3K27ac analysis between psoriasis patients and healthy controls in CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup> cells.** In the genome-wide analysis, the master list size refers to the number of H3K27ac enriched sites included in the consensus list built by DiffBind to perform the differential analysis. In the analysis restricted to enhancers, the size of the master list was reduced to only those sites from the genome-wide master list annotated as enhancers (weak and strong) according to the chromatin segmentation map for each particular cell type. Genome-wide hits in CD14<sup>+</sup> monocytes and CD8<sup>+</sup> also contained the hits identified in the enhancers restricted analysis. Significant differentially H3K27ac regions using FDR<0.05 and no FC threshold.

One of the differentially acetylated regions between patients and controls in CD14<sup>+</sup> monocytes is at an enhancer located between the 3' UTR of the genes *SLC15A2* and *ILDR2* (Figure 1.7). Consistently with its annotation as an enhancer, this region is overlapping a DHS site and H3Kme1 (enhancer mark) modifications. Although *ILDR2* is widely expressed in immune cells and has

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

recently been proven to act as a negative regulator for T cells, publicly available promoter capture and Hi-C data failed to show interactions of this regions with the promoter of any gene. Conversely, this region harbours SNPs that have been identified as *cis*-eQTL in whole blood and total unstimulated and stimulated CD14<sup>+</sup> monocytes for the calmodulin-binding motif-containing protein *IQCB1* gene, at 186.4Kb up-stream this peak (GTeX; Fairfax et al. 2014). MS risk LD SNPs (rs34543553, rs73855480 and rs73855480) in this peak are also *cis*-eQTLs in unstimulated total CD4<sup>+</sup> and CD8<sup>+</sup> cells for *IQCB1* (Kasela2017), which has been prioritised by Dr Fang in-house pipeline as a potential drug target for this autoimmune disease. Amongst the CD4<sup>+</sup> hits, all the regions presented very moderated changes with no evidence of regulating expression of any relevant gene.

\*CTCF peak may help to maintain, edit to K562 or remove



**Figure 1.3: Epigenetic accessibility landscape at the chr3:121,675,048-121,677,505 enhancer in circulating CD14<sup>+</sup> monocytes from psoriasis patients and healthy controls.** Describe the track

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

When performing genome-wide differential analysis, CD14<sup>+</sup> monocytes and CD8<sup>+</sup> revealed additional statistically significant differentially H3K27ac modified regions, which also included those already identified in the restricted enhancers analysis (Table 1.5). The newly identified regions in both cell types were mostly in regions lacking of ATAC-seq and H3K4me1 signal or interactions with nearby gene regulatory regions. Genome-wide analysis in CD4<sup>+</sup> cells did not identify significant differentially modified targets outside enhancers and also failed to retain significance for the four hits identified in the restricted analysis, most likely due to increase in multiple testing due to the increase in size of the master list (Table 1.5). Interestingly, the differential regions between psoriasis patients and controls identified in CD19<sup>+</sup> cells when performing genome-wide analysis presented considerable fold changes (absolute log<sub>2</sub>FC ranging from 2.47 to 6.37). However, most of them (10 out of 12) did not overlap open chromatin regions and none was found to be interacting with other enhancers or distal promoters. The absence of differentially H3K27ac sites between patients and controls at CD19<sup>+</sup> enhancers may be reinforcing the lack of relevance of circulating B cells in psoriasis when compared to the other three cell types.

Overall, restricting the differential analysis to enhancer annotated regions did not show a great increase in the number of significant differentially modified H3K27ac sites when compared to the genome-wide analysis in any of the four cell types. The results in this pilot cohort did not show relevant global epigenetic changes in H3K27ac sites between psoriasis patients and controls at the systemic level for these cell types and sample size.

## **Identifying global changes in immune cells chromatin accessibility between psoriasis patients and healthy controls**

As previously explained, the cell type specific chromatin accessibility landscape is determined by the combination of histone modifications and

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

DNA binding proteins (including TF and co-regulatory proteins) in a particular locus. The previous results showing only moderate changes in the H3K27ac landscape between patients and controls are not necessarily representative of the overall changes in the chromatin accessibility landscape in disease. In order to interrogate genome-wide changes in chromatin accessibility between patients and controls not restricted to changes in H3K27ac, ATAC-seq was performed in the same four cell types in 8 patients and 10 controls (Table 1.4)

ATAC-seq quality control was performed in 72 libraries generated in CD14<sup>+</sup> monocytes, total CD4<sup>+</sup>, total CD8<sup>+</sup> and CD19<sup>+</sup> cells from psoriasis patients and control individuals. The median of total reads after filtering ranged between 39.2 and 49.8 in CD4<sup>+</sup> and CD19<sup>+</sup>, respectively, and was over the 15 million reads determined as appropriate minimum (Chapter ??) in all the samples (Figure 1.4 a).

The differences in read depths across ATAC-seq samples are mostly due to the difficulties in appropriate determining the molarity of the libraries prior to the sequencing pooling (due to the fragment size heterogeneity) as well as the number of duplicates present in each library, which also correlated with MT reads, as previously mentioned. In this particular case, differences in the percentage of MT reads were observed between samples from cohort 1A generated with standard ATAC-seq protocol from Buenrostro *et al.*, 2013 and the FAST-ATAC libraries of cohort 1B using the later modified (**Corces2016**) protocol (Figure ??).

All the samples presented the characteristic ATAC-seq fragment size distribution recapitulating nucleosome periodicity, which is one of the measured to determine samples quality control, as previously detailed in Chapter ?? (data not shown). Analysis of ATAC-seq signal enrichment across gene TSS revealed that most of the samples had enrichment over 6, considered the lower threshold by ENCODE (Figure 1.4 b) and only PS2000 and PS2001 CD14+ monocytes were

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

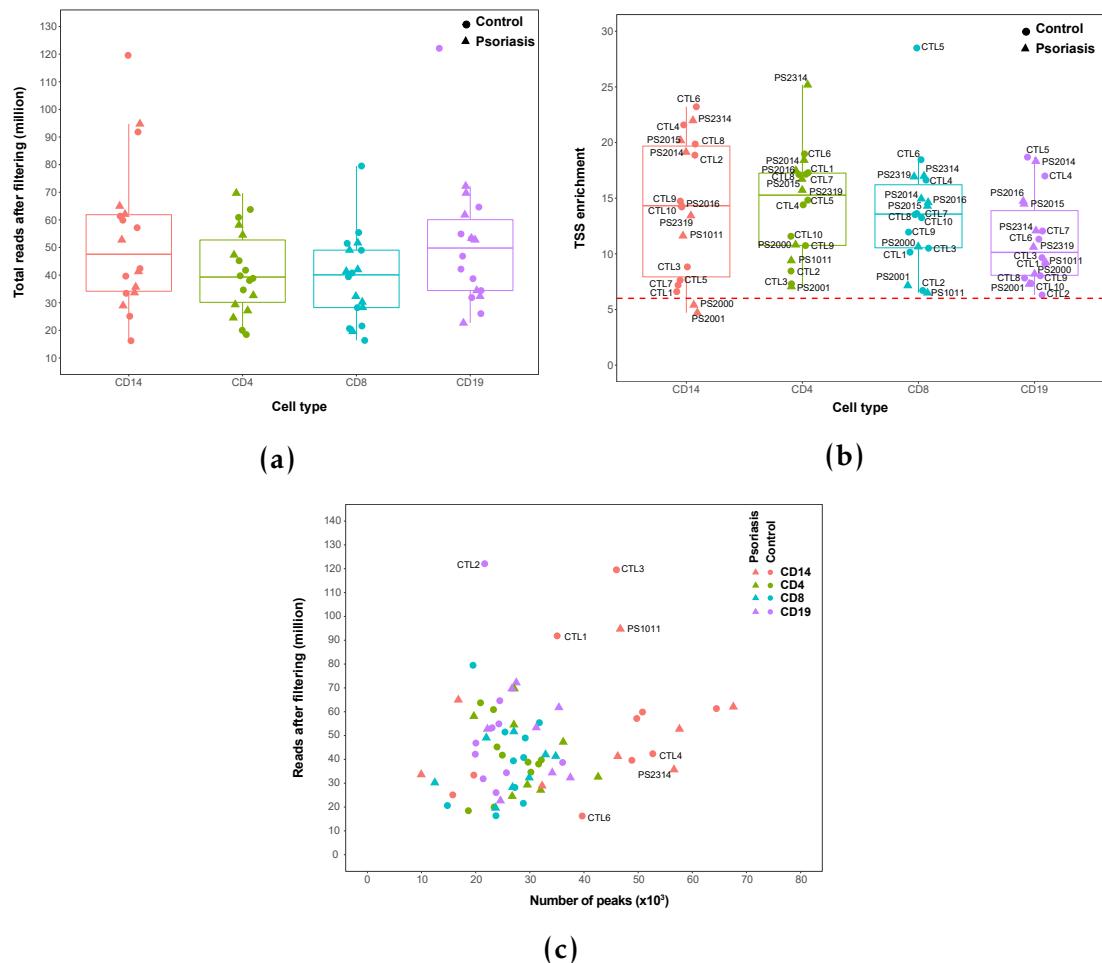


Figure 1.4: Quality control assessment of the ATAC-seq libraries generated from immune cells in psoriasis and control samples.

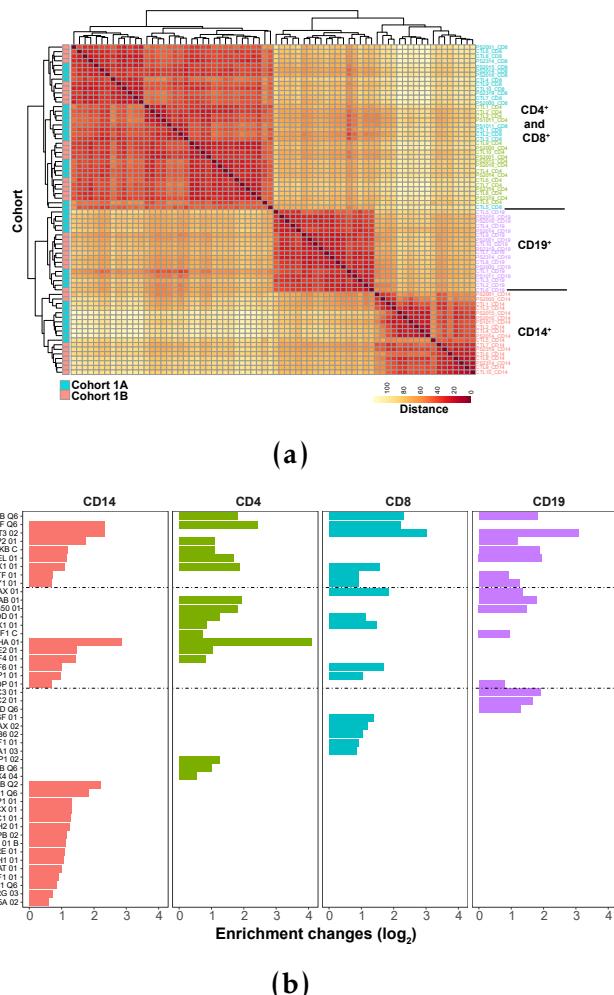
## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

removed from downstream analysis due to low signal-to-noise ratios. When comparing the number of peaks passing IDR filtering in each samples versus the number of reads after filtering, most of the samples presented between 10,000 to 35,000 peaks (Figure 1.4 c). Since sequencing depth of most of the samples was equal or greater than 15 million reads, most of the differences in number of called peaks were intrinsic to the cell type and the signal-to-noise differences in the samples, as previously studied in Chapter ???. For example, CD14<sup>+</sup> monocytes presented greater number of peaks when compared to the other three cell types, despite the median of total number of reads after filtering being similar to the other cell types (Figure 1.4 a). Also, within the CD14<sup>+</sup> monocytes samples, those with the greatest TSS fold-enrichment (e.g CTL4, CTL6 and PS2314) presented similar number of good quality called peaks to the samples with higher number of reads after filtering and lower signal-to-noise ratios (e.g PS1011, CTL1 and CTL3) (Figure 1.4 a and c). Interestingly, CD19<sup>+</sup> CTL2 appeared to be an outlier, presenting a noticeable lower number of peaks for its high sequencing depth (Figure 1.4 c). This observation together with its border line TSS enrichment supported removal of the CD19<sup>+</sup> CTL2 sample from downstream analysis.

A heatmap illustrating sample distance with additional k-means clustering using the consensus master list of ATAC-seq regions across the four cell types (ML\_all) showed successful separation of the samples according to the cell type into three main clusters corresponding to CD14<sup>+</sup> monocytes, CD19<sup>+</sup> and CD4+/CD8<sup>+</sup> T cells (Figure 1.5 a). Within each of the cell type clusters, samples did not separate based on disease condition, suggesting the absence of large global differences in the chromatin accessibility landscape between psoriasis patients and control individuals. Conversely, for some of the cell types such as x and y, samples grouped by batch within the cluster (fig of heatmap).

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.5: K-means clustered heatmap and conserved TFBS enrichment analysis in the consensus ATAC-seq regions identified in CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup> cells from the patients and controls cohort.** a) Distance matrix and k-means clustering for the 72 samples was performed based on the normalised read counts retrieved for each sample at the regions included in a consensus master list of ATAC-seq enriched sites built across all four cell types (MASTER\_ALL). Clusters have been additionally annotated using cohort identity. b) Enrichment analysis for the conserved TFBS was performed for each of the ATAC-seq cell type master lists of regions used for downstream differential analysis. Enrichment was tested for 258 human conserved TFBS identified by Transfac using a position-weight matrices based on experimental results in the scientific literature. Significant enrichment using FDR<0.01.

Each of the four cell type master list of ATAC-seq peaks (ML\_CD14, ML\_CD4, ML\_CD8, and ML\_CD19), used for the downstream differential chromatin accessibility analysis (explained in Chapter ??), presented the highest percentage of regions annotated as gene promoters, intronic and intergenic, as expected for ATAC-seq (Figure ??). *cis*-eQTL SNPs from a number of immune

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

cell types, including CD14<sup>+</sup> monocytes (unstimulated and stimulated), B cells, neutrophils, CD4<sup>+</sup> and CD8<sup>+</sup> cells (**Kasela2016**; Fairfax et al. 2012; Fairfax et al. 2014), were enriched within the ATAC-seq consensus peak list of each cell type. For example, eQTLs from unstimulated and stimulated (LPS or IFN- $\gamma$ ) monocytes presented greatest significant enrichment (FDR<0.01) in the ML\_CD14 (unstimulated fold-enrichment 5.1, LPS 2h fold-enrichment 4.7 and IFN- $\gamma$  fold-enrichment 5.0) when compared to the other eQTL datasets. Similarly, the *cis*-eQTLs from the CD4<sup>+</sup> and CD8<sup>+</sup> were the most enriched datasets (fold-change 8.3 and 8, respectively) in the ML\_CD8. The relevance of each cell type master list was further reinforced by the significant enrichment (FDR<0.01) of conserved TFBS within those ATAC-seq regions (Figure 1.5 b). For example, enrichment of conserved NF $\kappa$ B binding motifs(NFKB Q6, NFKB C and NFKAPPAB 01) was identified across the master lists from the different cell types. Similarly, conserved binding motifs fro TF involved in T cell biology , such as AREB6 (ZEB1), ATF6 and the heat-shock transcription factor HSF1 (**Guan2018**; **Yamazaki2009**; **Gandhapudi2013**). Overall, the enrichment of eQTL SNPs and conserved TFBS highlighted the potential of each cell type master list to harbour functional relevant differences in chromatin accessibility between psoriasis patients and controls.

Differential chromatin accessibility analysis between patients and controls was performed on the ATAC-seq normalised read counts for the regions of each cell type master lists using DESeq2. PCA analysis on the normalised counts of each cell type master list prior to the differential analysis revealed a batch effect correlating with the different ATAC-seq protocols used in cohort 1A and cohort 1B (standard ATAC-seq and FAST-ATAC, respectively) (Figure A.5 a). Therefore, the ATAC-seq protocol was included as a covariate in the differential analysis model. Moreover, CTL5 appeared as a cohort 1A outlier for all the cell types

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

(representative example Figure A.5 a) and was also removed from the differential analysis.

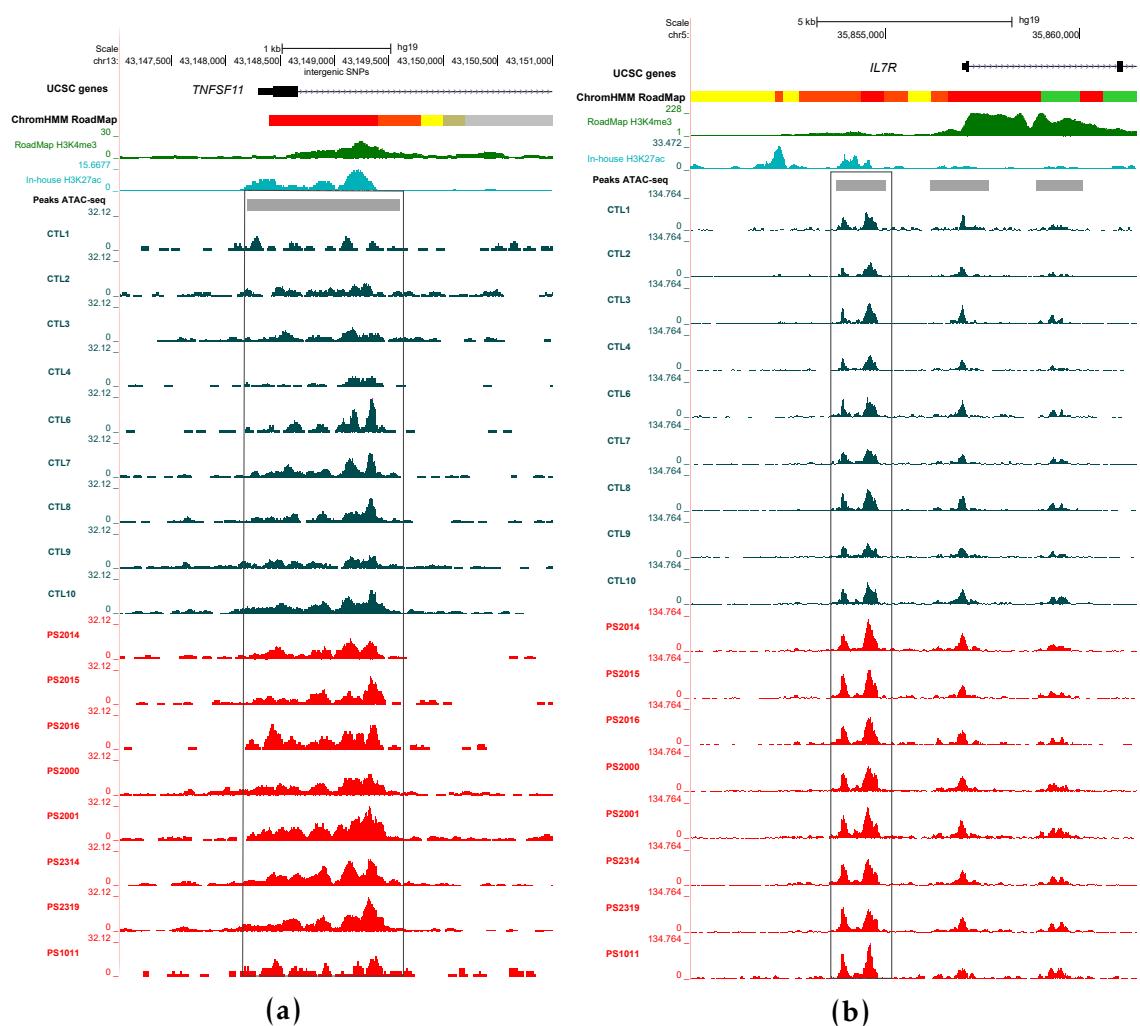
Genome-wide differential chromatin accessibility analysis revealed 55 significant (FDR<0.05) DARs between psoriasis patients and healthy controls in CD8<sup>+</sup> cells (Table 1.6). Conversely, CD14<sup>+</sup> monocytes, CD4<sup>+</sup> and CD19<sup>+</sup> cells only presented one or none DARs.

Cell type	Number of DARs FDR<0.05
CD14 <sup>+</sup>	1
CD4 <sup>+</sup>	0
CD8 <sup>+</sup>	55
CD19 <sup>+</sup>	1

**Table 1.6: Summary results from the differential chromatin accessibility analysis between psoriasis patients and healthy controls in CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup> cells.** The number of differentially accessible regions (DARs) refers to those statistically significant when using a cut-off for background reads of 80% (see Chapter ?? and an FDR<0.05. No threshold for the FC was applied in this analysis.

Annotation of the 55 CD8<sup>+</sup> DARs using cell type specific Roadmap Epigenomics chromatin segmentation map revealed the potential of some of those regions to be involve in regulation of gene expression, including 24 (44.4%) weak enhancers, 7 (12.9%) active promoters, 6 (11.1%) weak promoter and 2 (3.7%) strong enhancers. The functional relevance of the DARs in terms of regulation of gene expression was further investigated by integration of the CD8<sup>+</sup> cells eRNA data from the FANTOM5 project. Interestingly, only 8 of the CD8<sup>+</sup> DARs overlapped significantly expressed eRNAs. Amongst others, the DARs overlapping eRNA include a region at the TSS of the *TNSF11* gene and another at a distal promoter from *IL7R*, which were more accessible in the psoriasis patients compared to the healthy controls (Figure 1.6 a and b). The two DARs were also overlapping chromatin harbouring H3K4me3, a histone mark indicating an active promoter, and H3K27ac, mostly found at active enhancers, consistently with the overlap of those regions with FANTOM5 eRNAs in CD8<sup>+</sup>.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.6: Epigenetic landscape at two ATAC-seq differential accessible regions between patients and controls in CD8<sup>+</sup> cells.**xxx

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

Both of them are relevant genes in driving and maintaining the inflammatory response. For example, *TNFSF11* is a cytokine from the TNF family involved in the regulation of T cell-dependent immune response and osteoclast differentiation in RA and PsA (MirandaCar“s2006; Ritchlin2003). *TNFSF11* is also downstream the lead SNPs for a CD risk locus (ImmunoBase). Interestingly, the *TNFSF11* protein, RANKL was found to be overexpressed in epidermis from psoriasis patients compared to controls and cutaneous lupus erythematosus, highlighting the role of this gene in the pathophysiology of psoriasis (Toberer2011). On the other hand *IL7R* is a proximal gene to SLE and MS, amongst others, and the axis IL-7/IL-7R has been found to drive IL7 independent TNF- $\alpha$  inflammation in RA patients presenting iTNF resistance (van Roon2017). Other potentially interesting CD8 $^{+}$  DARs were found nearby genes such as the MAPK *MAP3K7CL* and *NFKB1*; However they were not at regions annotated as enhancers or overlaped with experimentally validated eRNAs. Additionally, none of the CD8 $^{+}$  DARs were found within an LD block ( $r^2 \geq 0.8$ ) of the psoriasis risk GWAS loci.

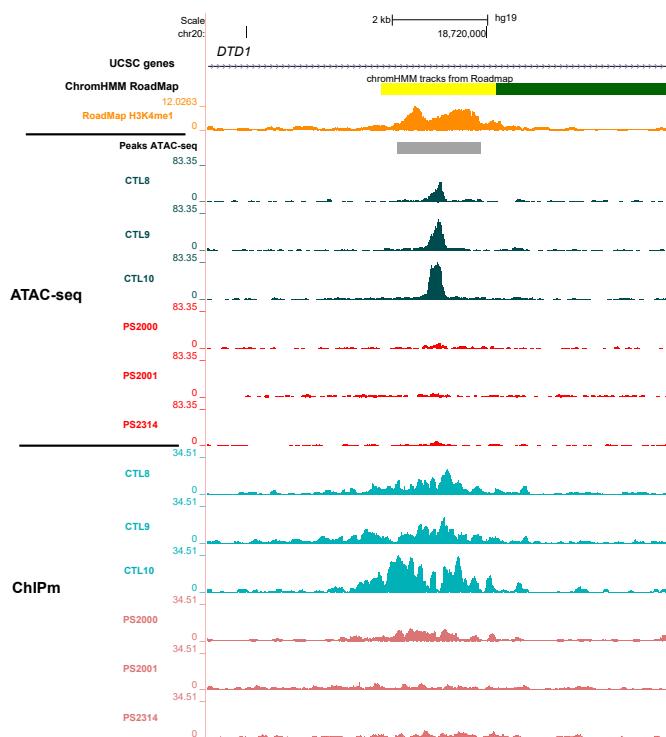
### **Integration of H3K27ac ChIPm and ATAC-seq chromatin accessibility profiles**

After defining the H3K27ac and chromatin accessibility profiles in four psoriasis circulating immune cell, the next step was to investigate the commonalities in the disease specific changes characterised by the two epigenetic functional approaches. The overlap between the H3K27ac ChIPm and ATAC-seq differential sites between patients and controls only showed one shared region in CD8 T cells within an intron of the D-tyrosyl-tRNA deacylase 1 (*DTD1*) gene (Figure). This region presented lower levels of H3K27ac (4 patients versus 4 controls) and was less accessible (8 patients and 9 controls) in the psoriasis patients when compared to the healthy controls (Figure). This differential region was annotated as an active enhancer according to the CD8 $^{+}$  ChormHMM

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

segmentation map and did not present 3D interaction with the promoter of any gene according to Hi-C and promoter Hi-C data in total CD8<sup>+</sup> cells (**Javierre2016**). Conversely, SNPs within this region appeared a significant eQTL for *DTD1* expression in whole (<https://gtexportal.org/home/eqtls>). The role of this gene has been described in the initiation of DNA replication and has been associated with aspirine-intolerance in asthmatics (**Pasaje2011**). However, no studies have yet highlighted a direct link of this gene with the pathophysiology of chronic inflammatory diseases.

Altogether, these results suggest that differences in H3K27ac are not driving the genome-wide changes in chromatin accessibility between psoriasis patients and healthy controls in total CD8<sup>+</sup> cells in this data.



**Figure 1.7: Epigenetic accessibility landscape at the chr3:121,675,048-121,677,505 enhancer in circulating CD14<sup>+</sup> monocytes from psoriasis patients and healthy controls. Describe the track**

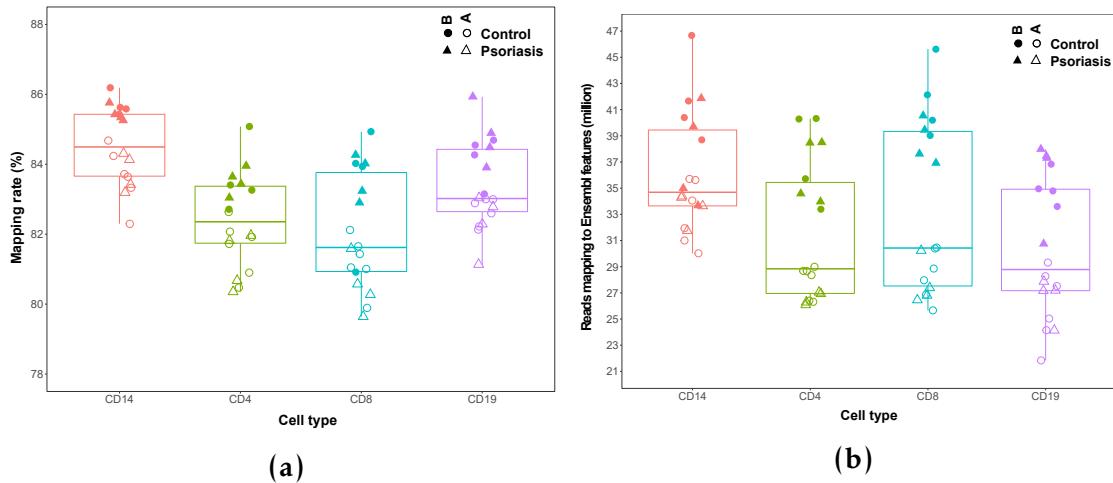
### **1.2.3 Differential gene expression in psoriasis circulating immune cells**

#### **Assessing quality control of the RNA-seq data**

In addition to characterising the chromatin accessibility landscape, gene expression profiles in psoriasis and healthy individuals were also analysed for the same four primary circulating immune cell types using RNA-seq. The percentage of RNA-seq mapping to a unique location in the genome using STAR algorithm (see Chapter ??) was appropriate (minimum recommended 70 to 80%), ranging between 79.64 and 86.19% ( CD8<sup>+</sup> CTL3 and CD14<sup>14</sup> monocytes CTL8, respectively) across all the 72 samples generated across the four cell types in psoriasis and control samples (Figure 1.8 a). After appropriate filtering of non-uniquely mapped and duplicated reads, all the samples presented a minimum of 20 total million reads (as required by ENCODE standards) mapping to a comprehensive list of Ensembl features, including protein coding genes and lncRNAs (Figure 1.8 b). The median of total million reads mapping to Ensembl features was greater for CD14 monocytes when compared to the other three cell types. However, the difference in medians across the four cell types did not exceed in more than 5 million reads. Interestingly, in all four cell types analysed, greater mapping rates and total million of reads mapping to Ensembl features were observed for cohort 1B samples when compared to cohort 1A. These differences were attributed to the library preparation and sequencing of each cohort in two different batches.

Similarly to ChIPm and ATAC-seq, the first and second PC from PCA analysis using the normalised number of reads mapping to each the 20,493 Ensembl genes passing quality control (see Chapter ??) showed that most of variability was driven by cell type differences (Figure 1.9 a). A heatmap illustrating sample distance based on the expression profile of each sample

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



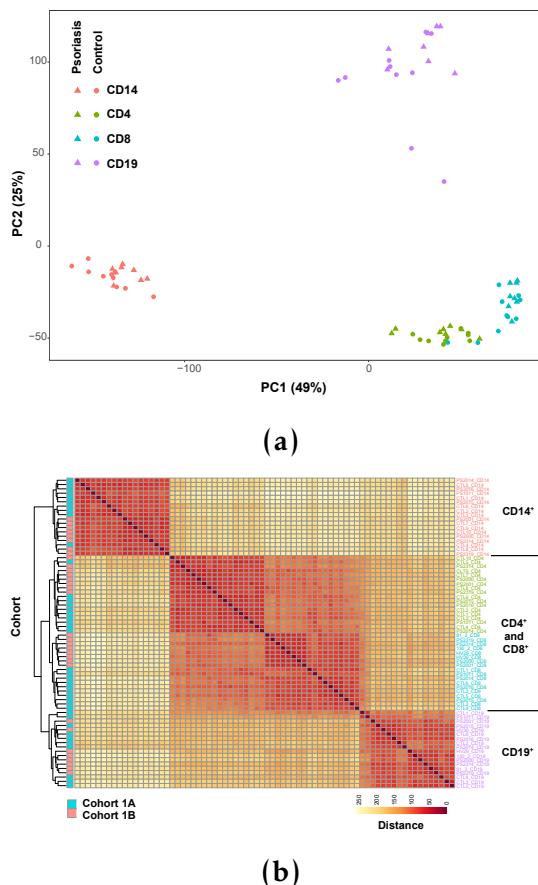
**Figure 1.8: Mapping rate and total reads after filtering (million) mapping to Ensembl genes in all the RNA-seq samples from psoriasis patients and controls in four cell types.** a) The mapping rate refers to the percentage of total sequenced reads from each sample that uniquely mapped to a particular site of the genome. b) The total number of reads after filtering for non-uniquely mapped and duplicated reads that mapped to Ensembl features, including coding protein genes and lncRNAs.

followed by k-means clustering revealed three main clusters corresponding to CD14<sup>+</sup> monocytes, CD4<sup>+</sup> and CD8<sup>+</sup> lymphocytes and CD19<sup>+</sup> cells (Figure 1.9 a). Within each of cell type clusters, samples were further grouped by cohort (1A and 1B) and not by condition (psoriasis and control), consistently with the previous differences in mapping rate and total million reads mapping to Ensembl genes observed across the two cohorts (Figure 1.8 a and b). This was reinforced by the clear correlation of sample batch with PC4 from the PCA analysis, which led to a very clear separation of the samples by cohort 1A and 1B, explaining 3% of the total variance (Figure A.5 b). Consequently, cohort identity was included in the differential gene expression model as a covariate.

### mRNA and lncRNA differential expression

DGE between 8 psoriasis patients and 10 healthy controls in CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup> was performed using DESeq2 and including the cohort identity as a covariate to account for the batch effect previously

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.9: PCA analysis and sample distance heatmap with k-means clustering illustrating the sample variability based on the gene expression profiles.** a) The first (x-axis) and second (y-axis) PCs from PCA analysis are used to reveal the main sources of variability across the 72 samples. b) Distance matrix and k-means clustering for the 72 samples was performed based on the normalised read counts mapping to 20,493 Ensembl featured remaining after appropriate filtering. Annotation of the clustering using cohort identity is included.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

---

mentioned. For each of the cell types a number of mRNAs were identified as differentially expressed for an FDR lower than 0.05 or 0.01 (Table 1.7).

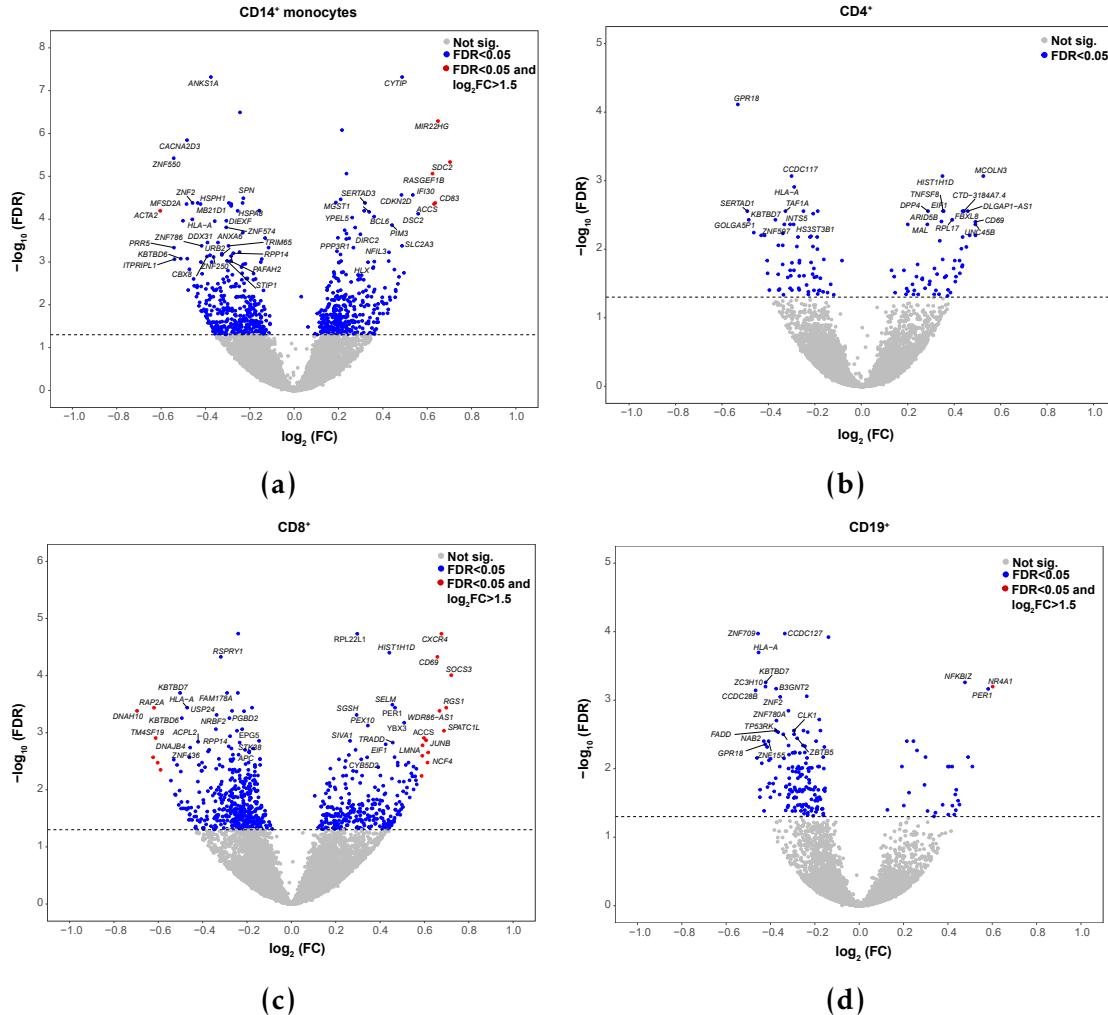
Cell type	mRNA	lncRNA
	FDR<0.05/0.01	FDR<0.05/0.01
CD14 <sup>+</sup>	671/229	28/8
CD4 <sup>+</sup>	108/40	12/4
CD8 <sup>+</sup>	656/175	31/5
CD19 <sup>+</sup>	167/71	6/2

**Table 1.7: Summary results from the DGE analysis between psoriasis patients and healthy controls in CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup> cells.** The number of statistically differentially expressed mRNAs and lncRNAs are listed for two FDR threshold (FDR<0.05 and FDR<0.01). No threshold for the FC was applied in this analysis.

CD14<sup>+</sup> monocytes and total CD8<sup>+</sup> were the two cell types presenting the largest number of mRNAs with modulated expression between psoriasis patients and controls. The more dysregulated gene expression response between patients and controls found in circulating psoriasis CD8<sup>+</sup> when compared to the CD4<sup>+</sup> may suggest the same hypothesis as in skin, where CD8<sup>+</sup> are considered the main effector cells undergoing activation upon the inflammatory stimuli ((Nickoloff1999)). Interestingly, CD19<sup>+</sup> presented greater number of differentially expressed mRNAs than CD4<sup>+</sup>, regardless their up to date absence of implication in disease (confirm).

The magnitud in the FC of gene expression between psoriasis patients and controls was moderate in all four cell types, showing the largest changes in CD14<sup>+</sup> monocytes and CD8<sup>+</sup> cells. Regarding the directionality of the statistically significant modulated genes (mRNA and lncRNAs) using FDR<0.05, CD14<sup>+</sup> monocytes (up 344, down 379) and CD4<sup>+</sup> (up 57, down 66) presented similar number of genes up-regulated and down-regulated in psoriasis patients when compared to the healthy controls. In contrast, CD8<sup>+</sup> (up 278, down 429) CD19<sup>+</sup> (up 29, down 148) a largest number of modulated genes were down-regulated in patients compared to controls.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.10: Magnitude and significance of the gene expression changes between psoriasis patients and healthy controls in four immune cell types.** Volcano plots for the results of the DGE analysis in a) CD14<sup>+</sup> monocytes, b) CD4<sup>+</sup>, c) CD8<sup>+</sup> and d) CD19<sup>+</sup> cells. For each gene, the log<sub>2</sub>(FC) represents the change in expression for that gene in the psoriasis group with reference to the healthy controls. Significant DEGs (FDR<0.05) in blue for absolute FC <1.5 and red for absolute FC >1.5. The volcano plots include mRNAs and lncRNAs

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

Some of the DEGs across the four cell types overlapped some of the DEGs from the two most comprehensive studies comparing expressin of PBMCs isolated from psoriasis patients an healthy controls (**Leo2009; Coda2012**). The greatest overlap (7 genes) was found between the DEGs in CD14<sup>+</sup> monocytes and those identified by Coda *et al.*, 2012. However, 5 out of 7 presented opposite direction. One of those genes dysregulated in the same direction was ubiquitin conjugating enzyme E2 D1 (*UBE2D1*), which mediates, for example, ubiquitination of the TNF receptor-associated factor 6 (TRAF6) protein (**Gru2008**). The greatest overlap with the Lee and colleagues DEGs was found for CD8<sup>+</sup> cells (3 genes) in the same direction. For example, CDC like kinase 1 (*CLK1*) involved in protein splicing was downregulated in this CD8<sup>+</sup> data set and in the Lee PBMCs in patients when compared to controls and its deficiency has been shown to lead to neuro-inflammation in mice (**Gu2017**). Similarly, only one overlap was found with the psoriasis DEGs in a study comparing PBMCs transcriptional profiles of three inflammatory diseases (IBD, RA and psoriasis)(**Mesko2010**). This was for the nicotinamide pPhosphoribosyltransferase (*NAMPT*) gene involved in metabolism and stress response, which was up-regulated in our CD14<sup>+</sup> monocytes as well as in PBMCs from psoriasis, IBD and RA patients, suggesting its role as a marker of inflammation rather than marker for psoriasis.

In order to incorporate the psoriasis genetic information from GWAS studies, an overlap between the significant DEGs (FDR<0.05) across the four cell types and a list with the putative genes associated with psoriasis GWAS hits from the NHGRI-EBI catalog (<https://www.ebi.ac.uk/gwas>) curated to include other genes from more recent studies was performed (Table 1.8). CD8<sup>+</sup> was the cell type with the largest number of DEGs (7 hits) overlapping with putative GWAS genes, followed by CD14<sup>+</sup> monocytes and CD4<sup>+</sup> (3 hits each). Some of those

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

---

genes were found in more than one cell type, including *NFKBIA*, *TNFAIP3* and *NFKBIZ*, amongst others.

Cell type	Number of GWAS overlaps	Up-regulated genes	Down-regulated genes
CD14 <sup>+</sup>	3	<i>NFKBIA</i>	<i>IL23A, FASLG</i>
CD4 <sup>+</sup>	3	<i>TNFAIP3, NFKBIZ</i>	<i>FASLG</i>
CD8 <sup>+</sup>	7	<i>TNFAIP3, NFKBIA, ETS1, SOCS1, NFKBIZ</i>	<i>B3GNT2, FASLG</i>
CD19 <sup>+</sup>	2	<i>NFKBIZ</i>	<i>B3GNT2</i>

**Table 1.8:** Overlap between putative psoriasis GWAS genes and the reported significantly DEGs in CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup> cells. DEGs list based on FDR<0.05.

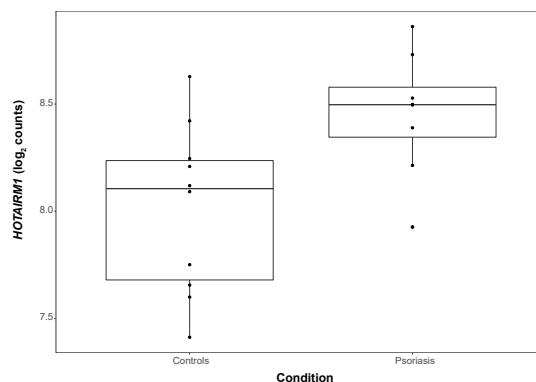
### The role of lncRNAs in psoriasis circulating immune cells

In addition to protein coding genes, some of the DEGs identified were classified as lncRNAs. Total CD8<sup>+</sup> and CD14<sup>+</sup> monocytes were the two cell types presenting the largest number of dysregulated lncRNAs between psoriasis patients and controls (Table 1.7. In contrast, CD19<sup>+</sup> was the cell type showing the lower number of lncRNAs differentially expressed. The vast majority of the dysregulated lncRNAs between patients and controls at FDR<0.05 were functionally uncharacterised, difficulty the interpretation of these results.

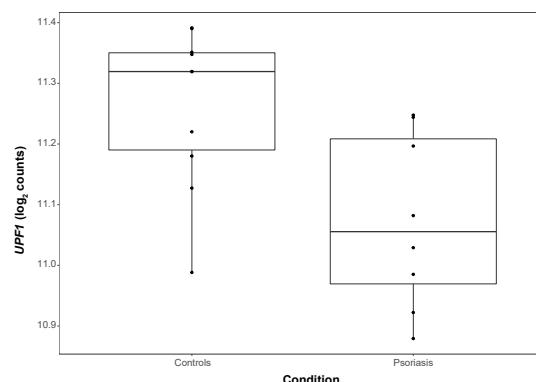
Amongst the known ones in CD14<sup>+</sup> monocytes, the negative regulator of antiviral response *DYNLL1-AS1* (or NAV) which has been shown to affect the histones modifications of some critical IFN-stimulated genes (ISGs), such as *IFITM3* and *MxA* leading to down-regulation of their expression (Ouyang2014). In this data, *DYNLL1-AS1* was down-regulated indicating lack of one of the negative regulators of the IFN response, regardless the reverse IFN- $\gamma$  signature observed in the pathway enrichment analysis. Another interesting dysregulated lncRNA in CD14<sup>+</sup> monocytes was the HOXA transcript antisense RNA myeloid-specific 1*HOTAIRM1*. In a study using PsA PBMCs *HOTAIRM1* was found to be

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

down-regulated and connected to the expression of the RNA helicase and ATPase *UPF1* ???. *UPF1* is involved in the nonsensemediated decay and in partnership with the monocyte chemotactic protein-1-induced protein-1 (*MCPIP1*) gene driven degradation of inflammation-related mRNAs to ensure maintenance of the homeostasis (Mino2015). In this present study, *HOTAIRM1* appeared to be up-regulated in the CD14+ monocytes from psoriasis patients (Figure 1.11 a) and this was consistent with significant down-regulation of *UPF1* in the same cell type (Figure 1.11 b), suggesting impairment of this homeostatic mechanisms in the psoriasis patients.



(a)



(b)

**Figure 1.11: RNA-seq expression levels of the lncRNA *HOTAIRM1* and its putative target *UPF1* in psoriasis and healthy controls CD14<sup>+</sup> monocytes.** Expression is illustrated as the log<sub>2</sub> of the normalised read counts mapping to a) the lncRNA *HOTAIRM1* and b)*UPF1*, which as been identified as one of the putative genes regulated by this lncRNA in the study conducted by Dolcino *et al.*, 2018.

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

The last relevant lncRNA differentially expressed in monocytes was *NEAT1*, which was up-regulated in the patients group when compared to the controls. *NEAT1* has been reported to also be up-regulated in SLE CD14<sup>+</sup> monocytes and knocking it down has revealed impairment of the TLR-4 signalling and down-regulation of inflammatory genes including IL-6 and CXCL10, amongst others (Zhang2016).

For total CD8<sup>+</sup> cells, the most relevant non-coding RNA was the miR *MIR146A*, which was captured in the standard RNA library preparation . Interestingly, *MIR146A* expression was down-regulated in psoriasis patients when compared to controls. This is in line with findings in serum from SLE and early RA patients (Wang2012; Filkov“-a”). In contrast, transcriptomic studies using PBMCs from plaque psoriasis and also similar studies in RA (including PBMCs, SFMCs and CD4<sup>+</sup> isolated from both tissues, amongst others) have reported increased levels of miR-146a in patients when compared to controls (Ele-Refaei2015; Churov2015). Molecular studies have suggested the role on miR-146a as a negative regulator of the TLR-4 pathway (Taganov2006). This study demonstrated regulation of *MIR146A* expression by NF-κB upon inflammation as well as *in vitro* targeting of the 3'UTR from *TRAF6* and IL-1 receptor-associated kinase 1 (*IRAK1*) genes. These are adaptor molecules that lead to activation of kinases from the TLR-4 pathway that eventually lead to translocation of NF-κB and AP-1 into the nucleus and transcriptional up-regulation of inflammatory genes.

Other lncRNAs were found to be dysregulated in more than one cell type. For example, *KCNQ1OT1* was downregulated in both, CD4<sup>+</sup> and CD8<sup>+</sup> cells. Dysregulated expression of this lncRNA has been reported in the Beckwith-Wiedemann syndrome consisting of a loss-of-imprinting pediatric overgrowth disorder with some skin features such as creases or pits in the skin near the ears (Pandei2008). *KCNQ1OT1* has been reported to interact with DNA

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

methyltransferases and also to facilitate the interaction of these enzymes with the chromatin, leading to misregulation of imprinted loci.

### **Pathway enrichment analysis for the DEGs**

In order to better understands the biological role of the significantly modulated genes, pathway enrichment analysis was performed for each cell type. The ATAC-seq results and the moderated FCs in the DGE analysis illustrated in the volcano plots suggest that the differences between patients and controls in these circulating immune cells are moderate. However, moderate differences may have an important impact in their phenotype for infiltration and activation in the skin. Therefore, to avoid losing biological meaningful effects and have a better understanding of the affected pathways, genes with FDR<0.05 and no FC cut off were included in this analysis. Functionally relevant significantly enriched pathways (FDR<0.01) were revealed for CD14<sup>+</sup> monocytes and CD8<sup>+</sup> (Table ?? and A.3). In CD19<sup>+</sup> cells, only one pathway named as generic transcription appeared to be significantly enriched for DEGs in this cell type. In contrast, in CD4<sup>+</sup> cell no pathway was enriched for modulated genes, consistently with the fact that this cell type presented the most moderate signature of disease modulated genes.

Amongst the relevant pathways, MAPK signalling enrichment was shared between CD14<sup>+</sup> monocytes and CD8<sup>+</sup> cells. Some of the DEGs contributing to the enrichment of this pathway in both cell types included MAPK gene members such as *MAP3K4* and *MAPK14*, both down-regulated in psoriasis when compared to controls. Interestingly, the *MAP3K4* is a member of the MAPKKK family which expression is down-regulated in LPS-stimulated PBMCs from CD patients, leading to reduced expression of the cytokine *IL-1A* and a relative immune deficiency in TLR-mediated cytokine production. Contributing to the enrichment (**Kraan2012**). Moreover, DGE of members of the dual-specificity

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

---

Cell type	Pathways
CD14 <sup>+</sup> monocytes	MAPK signalling IL-12 mediated signaling events Th-1 and Th-2 cell differentiation Th-17 cell differentiation TCR signalling Platelet-derived growth factor (PDGF- $\beta$ ) signalling Forkhead box O (FoxO) signalling
CD8 <sup>+</sup>	Osteoclast differentiation MAPK signalling TNF signalling IL-12 mediated signalling events NF- $\kappa$ B signalling Chemokine signalling

**Table 1.9: Most relevant pathways enriched for DEGs between psoriasis patients and healthy controls in CD14<sup>+</sup> monocytes and CD8<sup>+</sup> cells.** The enrichment analysis was conducted using significantly DEGs FDR<0.05 and no FC threshold. Enriched pathways had FDR<0.01 and a minimum of ten gene members overlapping with DEGs for that particular cell type.

phosphatases (DUSP) family involved in the immune response fine-tuning (Qian2009) in the enrichment of the MAPK pathway in CD14<sup>+</sup> monocytes and CD8<sup>+</sup> cells. *DUSP10* was down-regulated in the psoriasis CD14<sup>+</sup> monocytes and could be modulating reactive oxygen species production according to a knock-out mice phenotype presenting enhanced inflammation Qian2009. Conversely, *DUSP4* presented up-regulation in patients when compared to healthy controls and it could represent a pro-inflammatory feature as this gene has been demonstrated to have a role in driving inflammation in a sepsis mice model (Cornell2010).

Other interesting pathways enriched for both cell types DEGs IL-12 mediated signalling (Table 1.9). IL-12 signalling leads to T cells proliferation and IFN- $\gamma$  production through activation of TFs from the STAT family, importantly STAT4and is also a well known drug target for psoriasis treatment. Interestingly, CD14<sup>+</sup> monocytes from psoriasis presented down-regulation of *STAT4* and *STAT5A* in patients compared to controls. Likewise, *IFNG* expression in psoriasis

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

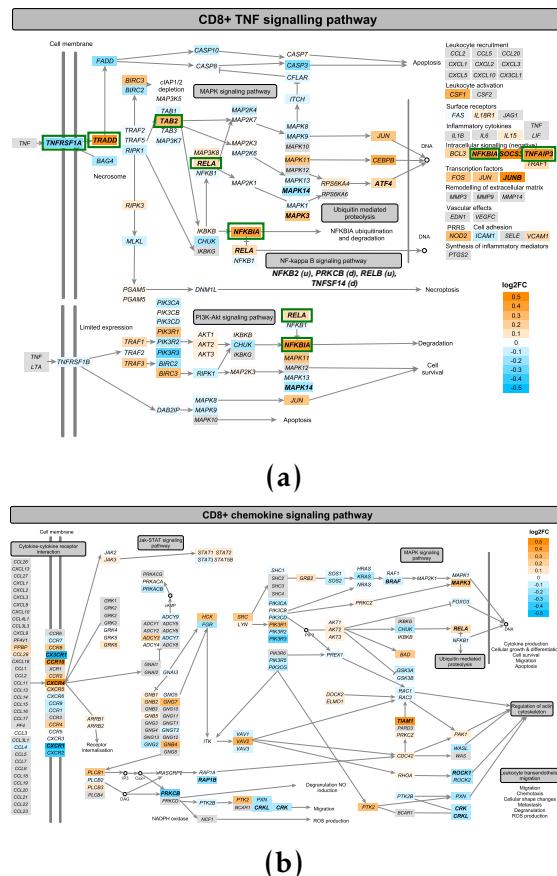
---

patients was lower than in healthy controls in CD8<sup>+</sup> cells. This phenomenon has been previously observed in macrophages derived from AS patients as well as in an SpA rat model (**Smith2008; Fert2014**). Additionally, *IL2RA* was up-regulated in CD8<sup>+</sup> from psoriasis patients when compared to controls which may enhance formation of the IL2-R and the signalling by this cytokine involved in effector and regulatory T cell differentiation (**Malek2010**).

The platelet-derived growth factor (PDGF- $\beta$ ) signalling pathway was only enriched in CD14<sup>+</sup> monocytes (Table 1.9). Within this pathway the *SLA* gene appeared to be down-regulated in psoriasis patients compared to controls. A *SLA* knock-out mice model have shown impaired IL-12 and TNF- $\alpha$  and failure of T cell stimulation by GM-CSG treated bone marrow-derived DCs (**Liontos2011**).

A number of very relevant inflammatory pathways in psoriasis were identified to be enriched only in CD8<sup>+</sup> cells. These included TNF, NF- $\kappa$ B and chemokine signalling (Table 1.9). Due to the close relationship between these three pathways, some DEGs contributed to the enrichment of more than one of them. For example, the NF- $\kappa$ B inhibitor A (*NFKBIA*) gene, which was up-regulated in CD8<sup>+</sup> cells from psoriasis compared to healthy controls, was present in all three pathways (Figure ?? a). Another example was the TNF- $\alpha$  induced protein 3 (*TNFAIP3*), also up-regulated, and a member of the TNF and NF- $\kappa$ B pathways (fig). Interestingly, *NFKBIA* and *TNFAIP3* were also up-regulated in CD14<sup>+</sup> monocytes and CD4<sup>+</sup> cells. *NFKBIA* gene code for the NF- $\kappa$ B inhibitor alpha ( $I\kappa B\alpha$ ) which binds to the NF- $\kappa$ B subunits preventing them from translocation to the nucleus by masking a nuclear localization signal (NLS). Similarly, *TNFAIP3* gene code for the zinc finger protein and ubiquitin-editing enzyme A20, that inhibits both NF- $\kappa$ B signalling and TNF-mediated apoptosis. Unexpectedly, these two genes with an anti-inflammatory role appeared to be up-regulated in psoriasis patients when compared to controls in two of the studied cell types.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.12: Mapping of the DEGs identified in CD8<sup>+</sup> cells between psoriasis patients and healthy controls onto the TNF- $\alpha$  and the chemokine signalling pathways.** The a) TNF- $\alpha$  and b) chemokine pathways were sourced from KEGG, manually curated in a way that all member genes are maximised visually and then automatically color-coded by the log<sub>2</sub>FC expression between psoriasis patients and healthy controls CD8<sup>+</sup> cells isolated from PB. Significant DEGs (FDR<0.05) are highlighted in bold. In a), members of the TNF- $\alpha$  pathway shared with the NF- $\kappa$ B are highlighted with a green box. Additional members of the NF- $\kappa$ B pathway differentially regulated in CD8 cells have also been indicated in brackets. Enrichment for a) and b) was identified by using only the CD8<sup>+</sup> DEGs (FDR<0.05).

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

Other genes with a prominent pro-inflammatory role also appeared to be down-regulated in the NF- $\kappa$ B or TNF signalling pathways, such as the activating transcription factor 2 (*ATF2*) and 4 (*textit{ATF4}*) members of the TNF signalling cascade and the protein kinase C beta*PRKCB* from the NF- $\kappa$ B and chemokine signalling pathways. Notably, *ATF4* was found to be up-regulated in CD noninflamed ileal biopsies causing activation of autophagy genes, whilst its expression was down-regulated in biopsies with active CD, pointing towards dysregulation of this pathway in disease onset (Bretin2016). In contrast, up-regulation of pro-inflammatory genes members of these two pathways were also found. For example JunB proto-oncogene (*JUNB*) coding for one of the subunits of the TF AP-1 and three of the NF- $\kappa$ B subunits including *RELA*, *RELB* and *NFKB2*. Particularly, AP-1 undergoes activation following growth factors, cytokines, chemokines, hormones and multiple environmental stresses and acts as a negative regulator of cell proliferation and IL-6 production (Schonthaler2011) Although in psoriasis lesional skin AP-1 protein levels have been found to be down-regulated, JunB was found to be increased both, at the protein and mRNA levels, consistently with this observation in CD8 $^{+}$  circulating cells (Johansen2004).

Regarding dysregulation of chemokines, a mix of up-regulation and down-regulation of members of this pathway was found in CD8 $^{+}$  cells from psoriasis patients when compared to healthy controls (Figure ?? b) and no changes were found in CD4 $^{+}$  cells. One of the most relevant dysregulated cytokine genes was *CCR10*, the receptor for the chemotactic skin-associated chemokine CCL27. In this data CD8 $^{+}$  cells from psoriasis patients presented up-regulated expression of the *CCR10* receptor. Some studies have demonstrated an increased of CCR10 $^{+}$  infiltrated T lymphocytes in psoriasis (Homey2002). In circulation, expression of CCR10 is restricted to the a subset of circulating mCD4 $^{+}$  and mCD8 $^{+}$  T cells expressing the cutaneous lymphocyte-associated antigen (CLA), which are

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

preferentially recruited to cutaneous sites of inflammation where KCs express CCL27 (**Hudak2002**). A study in psoriasis circulating cells revealed a correlation between the frequency of CTLA<sup>+</sup> CD8<sup>+</sup> cells and disease severity measured by PASI score (**Sigmundsdóttir2001**). Other up-regulated chemokine receptors in CD<sup>+</sup> circulating psoriatic cells included CXCR4 gene (receptor for CXCL12) for which controversial finding have been reported about its role in skin inflammation and psoriasis (**Zgraggen2014; Takekoshi2013**).

TNF:

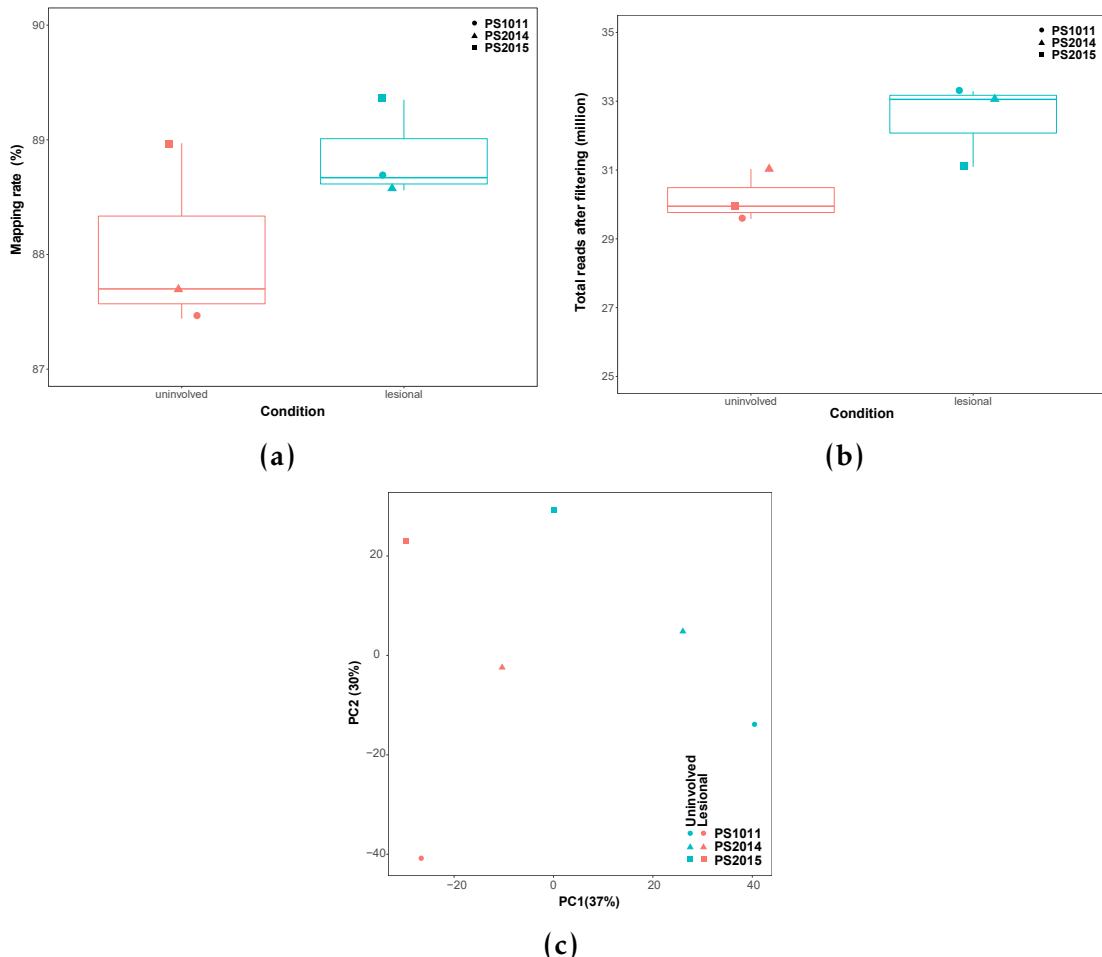
### **1.2.4 RNA-seq in epidermis from psoriasis patients**

#### **Quality control of the RNA-seq data**

For the three paired samples, both the uninvolved and the lesional samples presented a mapping rate greater than 80% for all the samples, being the rate moderately greater in the lesional compared to the uninvolved samples in all the three patients (Figure ?? a). The number of reads after filtering non-uniquely mapped and duplicated reads that were mapped to Ensembl genes ranged between 29.5 and 33.2 million in PS1011 uninvolved and PS1011 lesional, respectively (Figure 1.13 b). Similarly to the mapping rate, the final million of reads mapping to genes after filtering was also greater in the lesional samples compared to the controls, not differing in more than 4 million reads.

PCA analysis using the normalised number of reads mapping to the genes after filtering (see Chapter ??) revealed separation the lesional samples from the uninvolved by the first PC, which explained 37% of the variance (Figure ?? c). The second PC explained 30% of the variance and correlated with the patients ID. Overall, PCA analysis revealed substantial variation between the lesional and uninvolved samples and also showed biological variability between individuals, for which the paired design in the DGE analysis was accounting for.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.13: Mapping quality control and PCA analysis for the RNA-seq data in the uninvolved and lesional epidermis from psoriasis patients.** a) Mapping rate calculated as the proportion of sequencing reads mapping uniquely to a particular region of the genome. b) The total number of reads mapping to an Ensembl feature (including protein coding genes and lncRNAs) after removing the non-uniquely mapped and duplicated reads. c) First and second component of the PCA analysis performed on the normalised number of reads mapping to the Ensembl list of mRNAs and lncRNAs.

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

### **Summary of the DGE results**

DGE analysis revealed a total of 1,227 (FDR<0.05) and 702 (FDR<0.01) genes dysregulated between the uninvolved and lesional epidermis skin biopsies, including mRNAs and lncRNAs (Table ??). Amongst the 1,227 DEGs, a similar proportion of genes up- (559 genes) and down-regulated (629) in lesional skin when compared to uninvolved were identified (vulcano plot). Moreover, the FCs in gene expression were notably larger when compared to the changes in expression from analysis in circulating immune cells.

FDR threshold	mRNA	lncRNA	Overlap with GWAS genes
0.05	1181	46	up( <i>IFIH1, NOS2, STAT3, LCE3D</i> ), down( <i>TNFAIP3</i> )
0.01	677	25	<i>NOS2, STAT3, TNFAIP3, LCE3D</i>

**Table 1.10: Summary results of the DGE analysis between uninvolved and lesional psoriatic epidermal biopsies.** Number of differentially expressed mRNAs and lncRNAs are reported for two threshold of significance (FDR<0.05 and FDR<0.01). The DEGs overlapping putative psoriasis GWAS genes and the directionality in the change of expression are also specified.

Amongst the DEGs between the uninvolved and lesional skin in our study, four genes (FDR<0.05) overlapped with putative GWAS genes (Table ??). *IFIH1, NOS2, LCE3D* and *STAT3* were also found to be up-regulated in lesional compared to uninvolved skin biopsies from psoriasis patients in (Tsoi2015). In contrast, *TNFAIP3* was found to be up-regulated in (Jabbari2011), opposite to our finding.

### **Overall comparison with other skin transcriptomic studies**

As previously mentioned, the approach to perform the study of DGE in skin is different from most of the previously published studies using whole punch biopsies to compare lesional and uninvolved skin from psoriasis patients. During

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

the course of this project a study published by Tervaniemi and colleagues also aimed to characterise the transcriptional profiles of the epidermis from psoriasis patients lesional and uninvolved skin in a more elegant way than the previous studies using whole skin thickness biopsies. As previously detailed, Tevaniemi *et al.*, 2016 had a bigger sample size (six psoriasis patients) compared to this study (three psoriasis patients) and also included nine controls epidermis biopsies. In order to explore the similarities between the two studies, a comparison for the DEGs identified in both studies was conducted. Tervaniemi reported a total of 2,589 DEGs passing their filtering criteria ( $FC < 0.75$  or  $FC > 1.5$  and  $FDR < 0.05$ ). In contrast to our findings, the number of genes up-regulated in lesional epidermis compared to uninvolved (2,330) was larger than the number of down-regulated targets (261). Regarding overlap, a total of 359 out of the 1,227 DEGs (29.25%) identified in our study were shared with Tervaniemi results, of which 239 and 75 were up- and down-regulated, respectively. The direction of change in 45 out of the 359 shared genes appeared to be opposite across the two datasets. An example is the *SERPINB2* gene, a serine protease inhibitor of the serpin superfamily which presented down-regulation in our data and cultured lesional KCs from Swindell *et al.*, 2017 in contrast to the up-regulation in Tervaniemy results. Interestingly, an study demonstrated a defective stratum corneum in *SERPINB2* deficient mice as well as greater susceptibility to developing inflammatory lesions upon chemically induced atopic dermatitis compared to wild type controls (**Schroder2016**).

In addition to Tervaniemi study, our results were further contrasted to one of the most recent comprehensive RNA-seq study comparing lesional and uninvolved full thickness skin biopsies from psoriasis patients (**Tsoi2015**). Out of the 3,723 DEGs reported by Tsoi and colleagues, 507 genes were shared between the two studies (41% of the genes from this project), of which 272 were up-regulated and 228 appeared to be down-regulated. Interestingly, despite the

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

larger sample size, the DEGs between lesional and uninvolved skin from the Tsoi and colleagues study did not capture all the dysregulated genes from our study. This could suggest that our approach using more KCs enriched sample may be more sensitive to detect changes of gene expression that are masked when using whole punch biopsies. Amongst the 507 genes, only 7 (*ALOX15B*, *ARG2*, *LCE6A*,*MGST1*, *PNLIPRP3*, *TLDC1* and *UBL3*) presented opposite direction of change in the two studies. Amongst the genes showing opposite directions was *LCE6A*, one of the *LCE* gene family involved in the synthesis of the later cornified envelope layer, was down-regulated in lesional skin in our study, in contrast to the up-regulation found in the Tsoi analysis. Similarly, in the previous comparison, other genes from the *LCE* family, including *LCE1B*, *LCE1F* and *LCE2A* appeared also down-regulated in our data and up-in the Tervaniemi study. In contrast to discrepancies in other *LCE* genes, all three datasets presented up-regulation of the GWAS risk associated gene *LCE3B*. Notably, qPCR quantification of *LCE* genes from groups 1, 2, 5 and 6 demonstrated increased expression in psoriasis lesional skin (**Bergboer2011**).

Overall, the comparison of our results with these two studies suggested greater similarities with the full skin thickness biopsies from Tsoi *et al.*, 2015 in terms DEGs overlapping due to different technical reasons, as further detailed in the Discussion.

### **Dysregulated lncRNAs in the psoriatic lesional skin**

In addition to protein coding genes, a total of 46 lncRNAs were also significantly (FDR<0.05) differentially expressed between uninvolved and lesional skin from the three psoriasis patients. Similarly to the results presented for the circulating immune cells, the vast majority of modulated lncRNAs were functionally uncharacterised difficulting the interpretation of this results. 24 out of the 46 dysregulated lncRNAs were also reported as differentially expressed

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

by Tsoi and colleagues. An interesting example is H19 which was significantly down-regulated in the lesional skin when compared to uninvolved. H19 has been described to directly bind miR-130b-3p which down-regulates Desmoglein 1 (*DSG1*), a gene promoting KCs differentiation (Li2017). This finding was consistent with Tsoi *et al.*, 2015 and also with results from Li *et al.*, 2014 and *et al.*, 2016 where they compared lesional versus normal skin.

Interestingly, four miRNAs (*MIR146A*, *MIR22HG*, *MIR31HG* and *MIR205HG*) were also captured with the standard library preparation for mRNAs and lncRNAs implemented in our project. The relevance of miR-146A has been already commented in the DGE analysis from circulating immune cells. In lesional skin *MIR146A* was up-regulated when compared to uninvolved skin, consistently with other studies (Lerman2014; Tsoi2015), and also shown increase its expression when comparing lesional skin versus healthy biopsies (Li *et al.* 2014). Moreover, a polymorphisms in miR-146a has been associated with psoriasis in a small cohort study and a *MIR146A* knock-out mice with chemical induced psoriasis led to earlier disease onset and amplified epidermal activation (Srivastava2017). Another relevant finding was the up-regulation of *MIR31HG* in lesional skin, which has also been reported by (Tsoi2015). Notably, a functional study using the KCs immortal cell line HaCaT demonstrated that silencing miR-31hg induces cell cycle arrest and inhibits cell proliferation consistently with two characteristic functions dysregulated in psoriatic KCs (Gao2018).

### **Pathways enriched for the DEGs**

In order to better understand the functional role of the DEGs (FDR<0.05) between lesional and uninvolved epidermis from psoriasis patients skin biopsies, pathways enrichment analysis was performed. A considerable number of

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis

---

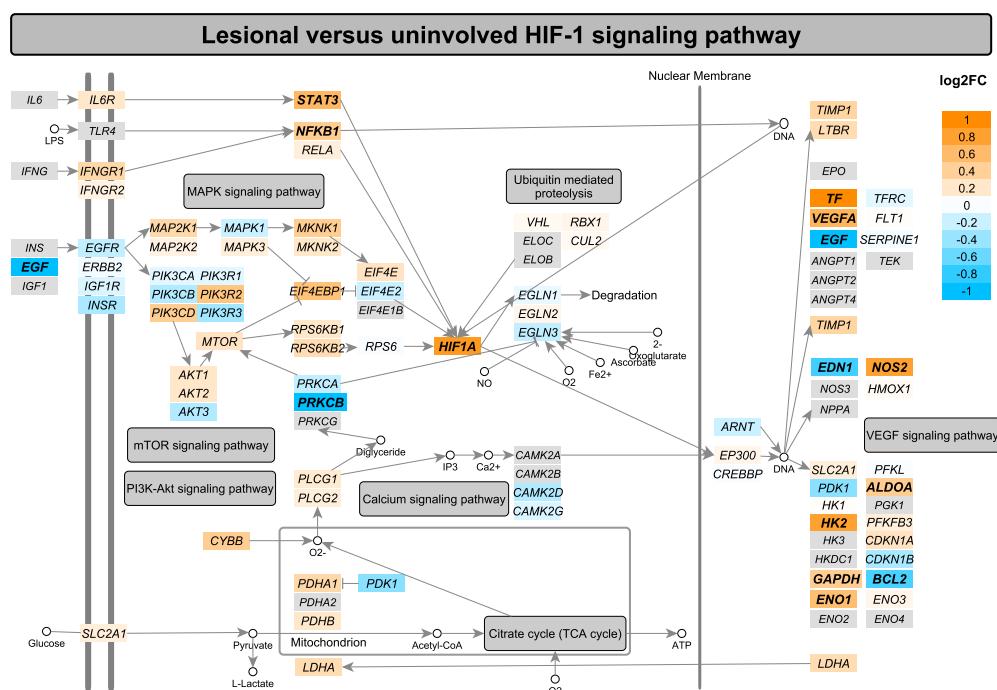
pathways were significantly enriched ( $FDR < 0.005$ ) for DEGs found in our analysis (Table ?? and A.4).

Lesional versus uninvolved epidermis enriched pathways
IFN- $\alpha/\beta$ /signalling
Peroxisome proliferator-activated receptors (PPAR) signalling
NOD-like receptor signaling pathway
IL-17 signalling
IL2-mediated signalling
G protein coupled receptor (GPCR) ligand binding
Hypoxia-inducible factor 1 (HIF-1) signalling
Cytokine signalling in immune system
Cell cycle
Apoptosis
Arginine and proline metabolism

**Table 1.11: Most relevant pathways enriched for DEGs between lesional and uninvolved epidermis isolated from psoriasis patients skin biopsies.** Significant pathways for  $FDR < 0.005$ . The analysis was performed using significantly DEGs  $FDR < 0.05$  and no FC threshold. Enriched pathways had a minimum of ten members overlapping with DEGs.

A number of pathways were related to alterations in cell cycle and metabolic processes, including hypoxia-inducible factor 1 (HIF-1) signalling, arginine and proline metabolism, glycolysis/gluconeogenesis and metabolism of amino acids and derivatives, amongst others. Dysregulation of similar functions have previously also been reported in other studies comparing lesional and uninvolved skin and genome-wide pathway analysis(Coda2012; Aterido2016; Tervaniemi2016). HIF-I signalling has been found to be up-regulated in psoriasis skin likely through hypoxia caused by increased cell proliferation rates and epidermal thickening. Our data revealed up-regulation of *HIF1A*, *VEGFA*, *ENO1* and *NOS2*, amongst others (Figure 1.15). Up-regulated expression of the hypoxia-inducible TFs HIF-1 $\alpha$  and HIF-2 $\alpha$  has been found in lesional skin and co-related with the increase in *VEGF* transcript levels, a target gene regulated by HIFs that mediates the pathological angiogenesis driving psoriasis (Rosenberg2007).

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



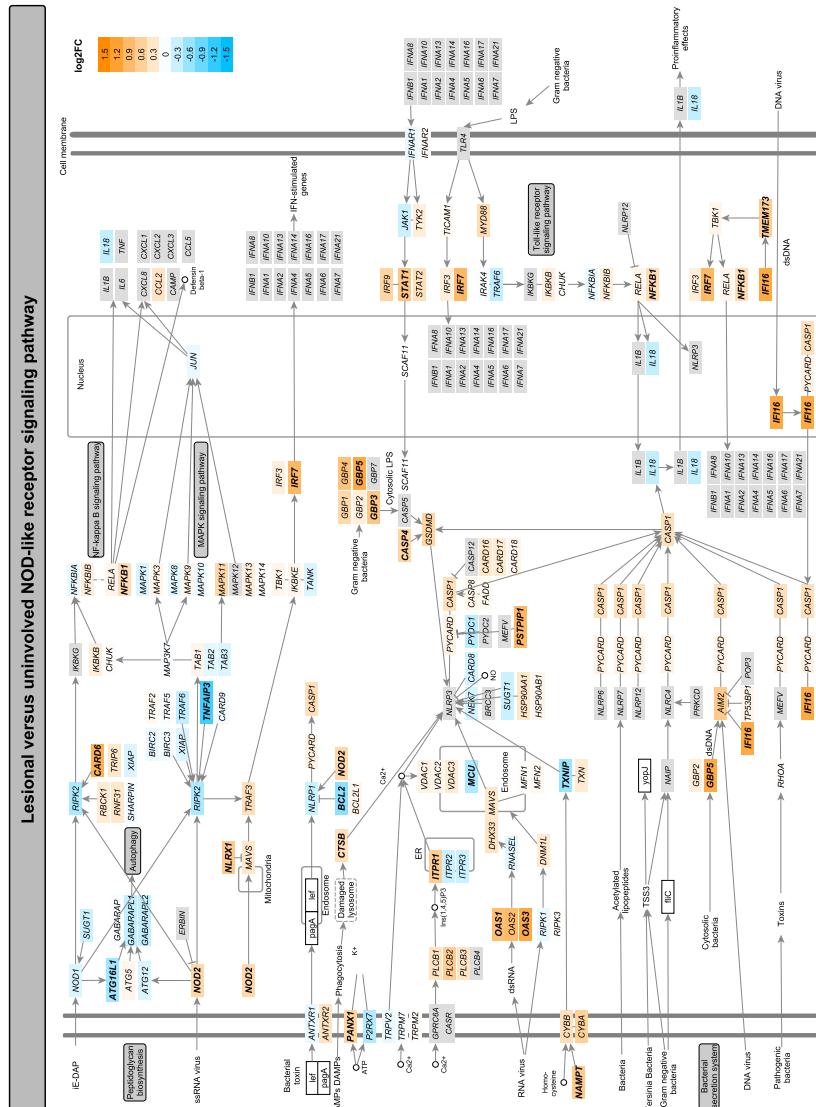
**Figure 1.14: Mapping of the DEGs between lesional and uninvolved epidermis from psoriasis patients onto the HIF-1 signalling pathway.** This pathway was sourced from KEGG, manually curated in a way that all member genes are maximised visually and then automatically color-coded by the log<sub>2</sub>FC expression between the lesional and uninvolved epidermis. Significant DEGs (FDR<0.05) are highlighted in bold. This pathway was identified by pathway enrichment analysis using only DEGs (FDR<0.05).

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

Immune relevant pathways including IFN, IL-17 and NOD-like receptor signalling were also identified in this analysis. The NOD-like receptor pathway responsible for detecting various pathogens and generating innate immune responses through NF- $\kappa$ B and MAPK activation, appeared enriched for 23 significantly DEGs (Figure 1.15 genes in bold). This pathway has also been identified as one of the most significantly enriched ones for DEGs in the epidermal study of Tervaniemi and colleagues. Some of the most up-regulated genes contributing to the enrichment included *NOD2*, *CARD6* or *IFI16*, amongst others, and they were also up-regulated in Tervaniemi's data. Interestingly, pathway enrichment analysis using the DEGs from Tsoi and colleagues failed to show significant enrichment for NOD-I signalling. Nevertheless, NOD-I signalling remained significantly enriched (10 DEGs mapping to this pathway) when using for the analysis only the DEGs from our study not overlapping the Tsoi and colleagues ones. These results reinforced the failure of whole skin biopsies transcriptomics to identify additional NOD-I signalling genes differentially regulated between lesional and uninvolved skin and the value of studying epidermal biopsies to unveil exacerbated dysregulation of functional pathways in KCs.

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.15: Mapping of the DEGs between lesional and uninvolved epidermis** from psoriasis patients onto the NOD-like signalling pathway. This pathway was sourced from KEGG, manually curated in a way that all member genes are maximised visually and then automatically color-coded by the  $\log_2FC$  expression between the lesional and uninvolved epidermis. Significant DEGs ( $FDR < 0.05$ ) are highlighted in bold. This pathway was identified by pathway enrichment analysis using only DEGs ( $FDR < 0.05$ ).

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

In addition to the NOD-I pathway, IL-17 signalling was another enriched pathway well known to be relevant in the development of psoriasis and the production of IL-17 by the Th-17 contributes to perpetuation of the innate host defense in the skin. Enrichment of the IL-17 signalling pathway in our data is driven by up-regulation of the S100 proteins family (*S100A7*, *S100A8* and *S100A9*) and the chemokine *CCL20*, which binds the CCR6 receptor and is involved in DCs and T cells chemotaxis. Moreover, enrichment of DEGs between lesional and uninvolved skin for the peroxisome proliferator-activated receptor (PPAR) signalling highlighted the link between metabolic dysregulation (particularly lipids) and innate immunity contributing to the psoriasis pathophysiology and increased risk of co-morbidities, such as metabolic syndrome and CVD. PPARs are ligand-dependent TF that have been shown to induce the inhibition of pro-inflammatory genes including IL-1 and TNF- $\alpha$  (Ji2001). In skin, PPARs have been demonstrated to contribute to homeostasis by inducing differentiation and preventing proliferation (Rivier1998). Moreover, inhibition of genes regulated by PPAR- $\gamma$  has been described in studies comparing lesional versus healthy skin biopsies (Li et al. 2014).

### **1.2.5 Comparison of the systemic and tissue-specific gene expression signatures in psoriasis**

In order to commonalities and differences in psoriasis gene expression at the affected tissue (skin) and the systemic level (circulating immune cells), overlap between the different list of DEGs was performed. Only modest overlap was found between dysregulated genes in lesional skin compared to uninvolved and the DEGs identified in circulating immune cells, showing the greatest overlap CD14 $^{+}$  monocytes and CD8 $^{+}$  cells. Interestingly, a notable proportion of the overlapping DEGs presented opposite direction of change in circulating immune cells and in the skin from psoriasis patients. A relevant case is the *TNFAIP3* gene,

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

which was up-regulated in psoriasis CD4+ and CD8<sup>+</sup> cells compared to controls and down-regulated in lesional epidermis when compared to uninvolved. In Tsoi *et al.*, 2015 this gene did not change expression between nor between lesional and uninvolved or lesional and healthy skin and Li *et al.*, reported its up-regulation in lesional skin. Another relevant transcript in the pathophysiology of psoriasis that presented opposite changes in CD<sup>+</sup> cells and in the skin analysis was the previously mentioned *MIR146A*. This miRNA was also up-regulated in lesional skin when compared to healthy controls and seems to be a clear feature of psoriasis pathogenesis in the involved tissue.

DEGs overlapping with skin	Total overlap	Same direction	Opposite direction
CD14 <sup>+</sup> monocytes	37	19	18
CD4 <sup>+</sup>	10	6	4
CD8 <sup>+</sup>	37	24	13
CD19 <sup>+</sup>	16	5	11

**Table 1.12: Overlap of the dysregulated genes in the four circulating cell types (psoriasis patients versus controls) and the DEGs in psoriasis patients skin biopsies (lesional versus uninvolved). DEGs FDR<0.05 for each of the comparisons**

The limited overlap between circulating and skin DEGs was also reflected in the different enriched pathways identified for each analysis. The pathways enriched for CD14<sup>+</sup> and CD8<sup>+</sup> DEGs were mostly related to immune-related pathways, including TCR , IL-12 , TNF and NF- $\kappa$ B signalling. Moreover, when compared to healthy controls, some of the genes in these circulating immune cells suggested certain immuno-suppression features that could be characteristic of these cells before or/and after having been exposed to the skin inflammatory *milieu*. The DEGs in lesional epidermis compared to uninvolved were enriched for pathways involved in metabolism, oxidative stress and cell cycle in addition to immune-related pathways. Moreover, the dysregulation of genes members of the immune-related pathways appeared to present an overall pro-inflammatory

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

signature, reinforcing the differences between circulating immune cells and skin in the study of psoriasis pathophysiology.

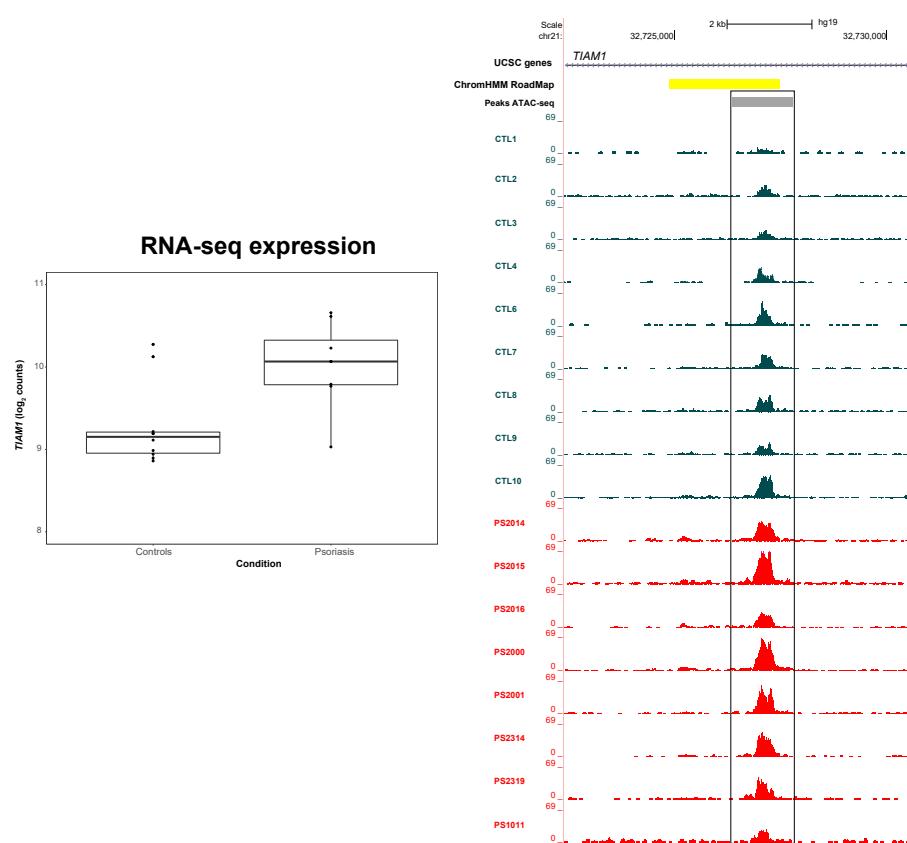
### **1.2.6 Integration of chromatin accessibility and expression data in circulating immune cells**

The characterisation of the chromatin accessibility landscape and the transcriptome in circulating immune cells from psoriasis patients, have revealed a greater effect of disease in gene expression than in chromatin accessibility. In order to assess to some extent the relationship between the two, overlap between DEGs and the genes proximal to a DARs ( $\leq 5\text{Kb}$ ) was performed. Limited overlap was only found CD8 $^{+}$  cells, where 6 out of the 53 DARs were annotated by proximity to an RNA-seq DEGs in the same cell type (*ARL4A, ASCL2, ENTPD1, TIAM1, TRAT1* and *ZNF276*).

An interesting example is the T Cell lymphoma invasion and metastasis 1 (*TIAM1*), which activates IL-17 expression and T cells transendothelial migration during inflammation (Kurdi2016; G“-e”rard2009). This gene showed an increased expression ( $\log_2\text{FC} 0.44$ ) in psoriasis patients CD8 $^{+}$  cells (Figure 1.16 left). Likewise, psoriasis CD8 $^{+}$  cells presented greater chromatin accessibility compared to healthy controls ( $\log_2\text{FC} 0.41$ ) in a region located at an intron of the *TIAM1* gene and annotated as an active enhancer according to RoadMAP chromatin segmentation map for this cell type (Figure 1.16 right). Common SNPs within this peak did not appear to be an eQTL regulating expression of any gene in CD8 $^{+}$  cells (Kasela2017) and chromatin conformation data did not reveal interaction of this particular region with the *TIAM1* promoter (**add**), difficulting the establishment of a mechanistic connection between chromatin accessibility and gene expression for this particular example in unstimulated conditions.

Another two relevant genes in the immune response for which ATAC-seq and RNA-seq presented overlap were the ectonucleoside triphosphate

## Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis



**Figure 1.16: Epigenetic accessibility landscape at the chr3:121,675,048-121,677,505 enhancer in circulating CD14<sup>+</sup> monocytes from psoriasis patients and healthy controls. Describe the track**

## **Defining the genome-wide chromatin accessibility landscape and gene expression profile in psoriasis**

---

diphosphohydrolase 1 (*ENTPD1*), which hydrolyses the pro-inflammatory mediator ATP attenuating the inflammation and acting as a modulator of the immune response, and the TCR-associated transmembrane adaptor 1 (*TRAT1*) gene, a positive regulator of the TCR signalling (Antonioli2013; Valk2006). Both genes presented up-regulated expression and an increased chromatin accessibility in psoriasis patients CD8<sup>+</sup> cells compared to healthy controls.

The changes in gene expression between psoriasis and healthy controls in the four circulating immune cell types included in this study have revealed a

### **1.2.7 Fine-mapping**

## **1.3 Discussion**

Trynka papers talking about relevant histone marks and similar studies. may be because is not one of the histones presenting phenotypic cell type specificity for complex disease GWAS SNPs and also in Lin did not appear as one of the histone marks that presented changes across the different cell types.

Skin other studies Swindell

skin and sf comparison <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4406155/>

# Appendices

<b>A Appendices</b>	<b>59</b>
A.1 Additional tables . . . . .	59
A.1.1 Chapter 5 Tables . . . . .	59
A.1.2 Chapter 4 Tables . . . . .	65
A.1.3 Chapter 5 Tables . . . . .	65
A.2 Additional figures . . . . .	65
A.2.1 Chapter 3 Figures . . . . .	65
A.2.2 Chapter 4 Figures . . . . .	65
A.2.3 Chapter 5 Figures . . . . .	67

# List of Figures

A.1 FAST-ATAC and Omni-ATAC NHEK tapestation profiles. . . . .	61
A.2 Assessment of TSS enrichment from ATAC-seq and FAST-ATAC in healthy and psoriasis skin biopsies samples. . . . .	62
A.3 Percentage of MT reads in the ATAC-seq libraries generated in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> isolated from psoriasis patients and healthy controls. . . . .	65
A.4 Genomic annotation of the consensus master list of ATAC-seq enriched sites built for downstream differential chromatin accessibility analysis in CD14 <sup>+</sup> monocytes, CD4 <sup>+</sup> , CD8 <sup>+</sup> and CD19 <sup>+</sup> . 66	66
A.5 PCA analysis illustrating batch effect in ATAC-seq and RNA-seq samples. . . . .	66
A.6 Permutation analysis SF vs PB in CD14 <sup>+</sup> ,CD4m <sup>+</sup> ,CD8m <sup>+</sup> and NK. 67	67
A.7 Heatmap for the top 20 marker genes of the CC-mixed and CC-IL7R CD14 <sup>+</sup> monocytes subpopulations. . . . .	68

A.8 Identification of the CD14 <sup>+</sup> monocytes populations from bulk SFMCs and PBMCs using scRNA-seq transcriptomes. . . . .	68
---	----

## List of Tables

A.1 Enrichment of ATAC-seq reads across the TSS for the CD14 <sup>+</sup> monocytes and CD4 <sup>+</sup> samples fresh, frozen and fixed. . . . .	59
A.2 Evaluation of ChiPm library complexity for the psoriasis and control chort 1B ChiPm assay. . . . .	60
A.3 Additional enriched pathways for DEGs between psoriasis and healthy controls in CD14 <sup>+</sup> monocytes and CD8 <sup>+</sup> cells. . . . .	62
A.4 Additional enriched pathways for DEGs between lesional and uninvolved epidermis isolated from psoriasis patients skin biopsies.	63
A.5 Additional enriched pathways for the DEGs between SF and PB CD14 <sup>+</sup> monocytes from the CC-mixed and CC-IL7R subpopulations.	64

# Appendix A

## Appendices

### A.1 Additional tables

#### A.1.1 Chapter 5 Tables

Cell type	Condition	TSS enrichment		
		CTL1	CTL2	CTL3
CD14	Fresh	17.4	19.6	14.11
	Frozen	26.3	25.2	27.1
	Fixed	2.5	16.5	22.4
CD4	Fresh	5.3	5.6	7.7
	Frozen	17.9	14.1	16.1
	Fixed	7.9	23.0	14.3

Table A.1: Enrichment of ATAC-seq reads across the TSS for the CD14<sup>+</sup> monocytes and CD4<sup>+</sup> samples fresh, frozen and fixed.

## Appendices

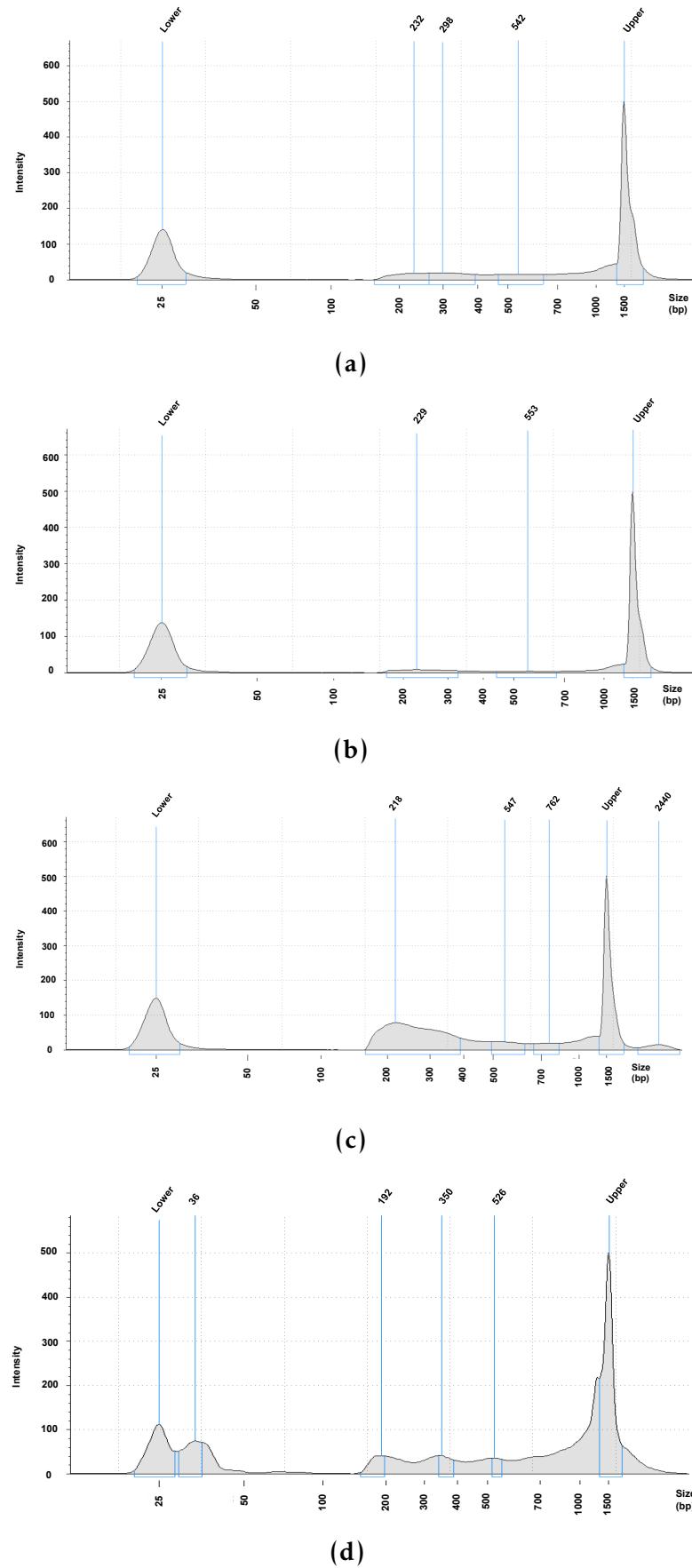
---

Sample ID	NRF	PBC1/PBC2
PS2000 CD14	77.6	0.60/2.5
PS2001 CD14	84.9	0.70/3.0
PS2314 CD14	81.1	0.60/1.8
PS2319 CD14	79.9	0.60/2.2
CTL7 CD14	81.1	0.65/2.2
CTL8 CD14	83.9	0.66/2.3
CTL9 CD14	80.7	0.60/2.3
CTL10 CD14	83.1	0.65/2.1
PS2000 CD4	84.8	0.75/3.4
PS2001 CD4	82.0	0.72/2.9
PS2314 CD4	82.9	0.71/2.8
PS2319 CD4	82.4	0.73/3.2
CTL7 CD4	78.6	0.68/2.5
CTL8 CD4	81.8	0.71/2.9
CTL9 CD4	81.6	0.74/3.3
CTL10 CD4	77.6	0.61/1.9
PS2000 CD8	77.0	0.76/4.5
PS2001 CD8	74.7	0.74/4.0
PS2314 CD8	74.2	0.75/4.1
PS2319 CD8	72.2	0.75/4.0
CTL7 CD8	32.7	0.32/1.5
CTL8 CD8	70.1	0.70/3.3
CTL9 CD8	73.9	0.73/3.7
CTL10 CD8	68.2	0.65/2.9
PS2000 CD19	38.0	0.42/1.9
PS2001 CD19	71.4	0.71/3.7
PS2314 CD19	29.4	0.34/1.8
PS2319 CD19	76.1	0.78/4.8
CTL7 CD19	74.2	0.69/3.1
CTL8 CD19	68.4	0.67/3.2
CTL9 CD19	75.1	0.76/4.6
CTL10 CD19	61.7	0.59/2.6

**Table A.2: Evaluation of ChiPm library complexity for the psoriasis and control cohort 1B ChiPm assay.** NRF, PBC1 and PBC2 are the three measures used according to the ENCODE standards as referred in Chapter ???.  $0.5 \leq \text{NRF} < 0.8$  acceptable;  $0.8 \leq \text{NRF} \leq 0.9$  compliant;  $\text{NRF} > 0.9$  ideal;  $0.5 \leq \text{PBC1} < 0.8$  and  $1 \leq \text{PBC2} < 3$  moderate bottlenecking;  $0.8 \leq \text{PBC1} < 0.9$  and  $3 \leq \text{PBC2} < 10$  mild bottlenecking.

## Appendices

---



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

## Appendices

---

### CD14<sup>+</sup> monocytes additional enriched pathways in psoriasis

---

Generic transcription  
RNA transport  
GnRH signalling  
Ribosome biogenesis in eukaryotes  
Neurotrophin signaling  
Spliceosome  
Autophagy  
Protein processing in endoplasmic reticulum

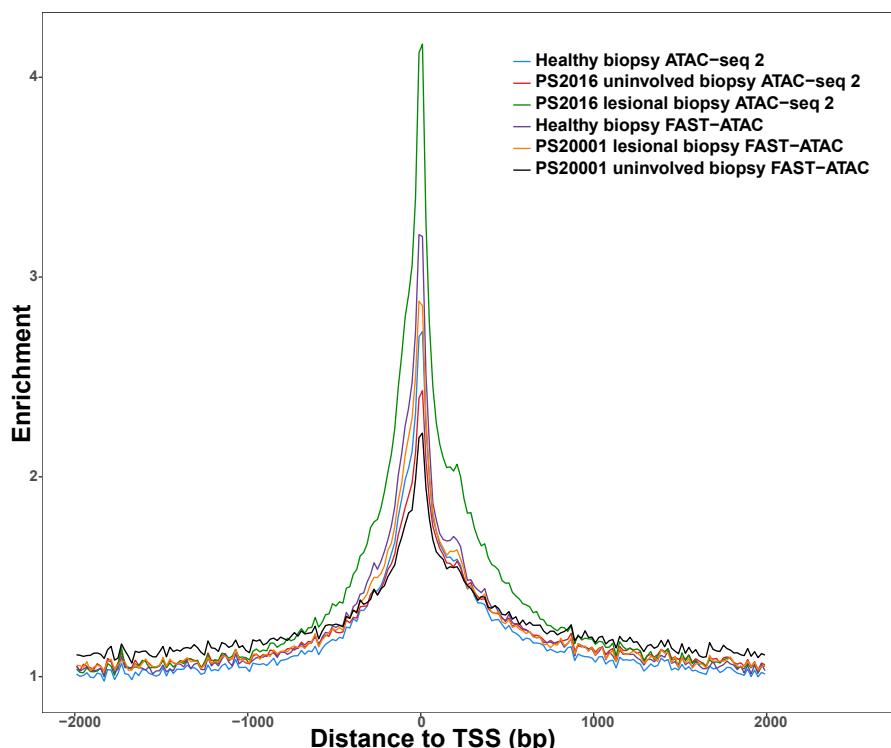
### CD8<sup>+</sup> additional enriched pathways in psoriasis

---

Epstein-Barr virus infection  
RNA Polymerase I and III, and mitochondrial transcription  
Apoptosis

---

**Table A.3: Additional enriched pathways DEGs between psoriasis and healthy controls in CD14<sup>+</sup> monocytes and CD8<sup>+</sup> cells.** Significant pathways for FDR<0.01. All the enriched pathways contained a minimum of ten DEGs FDR<0.05 from the analysis.



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

## Appendices

---

### **Lesional versus uninvolved epidermis additional enriched pathways**

---

Genes encoding extracellular matrix and extracellular matrix-associated proteins  
    Serine/threonine-protein kinase (PLK1) signalling  
        Genes encoding secreted soluble factors  
            Glycolysis/gluconeogenesis  
            FOXM1 transcription factor network  
        Phase 1 functionalization of compounds  
            Biological oxidations  
            G2/M Checkpoints  
            Biological oxidations  
            Aurora B signaling  
        Chemical carcinogenesis  
            Serotonergic synapse  
    Drug metabolism-cytochrome P450  
        Mitotic M-M/G1 phases  
            DNA Replication  
            MicroRNAs in cancer  
    Metabolism of amino acids and derivatives  
        Metabolism of carbohydrates  
            Glycosaminoglycan metabolism  
            E2F transcription factor network  
            p73 transcription factor network  
    Genes encoding structural ECM glycoproteins  
    Transmembrane transport of small molecules  
        Fc-epsilon receptor I signaling in mast cells  
            Tight junction  
    Origin recognition complex subunit 1 (Orc1) removal from chromatin

---

**Table A.4: Additional enriched pathways for DEGs between lesional and uninvolved epidermis isolated from psoriasis patients skin biopsies.** Significant pathways for FDR<0.005. All the enriched pathways contained a minimum of ten DEGs FDR>0.05 from the analysis.

## **Appendices**

---

---

### **CC-mixed CD14+ monocytes additional enriched pathways**

---

SLE  
Translation  
3'-UTR-mediated translational regulation  
Th-1 and Th-2 cell differentiation  
Peptide chain elongation  
Rheumatoid arthritis  
Metabolism of proteins  
Cell adhesion molecules (CAMs)  
Th-17 cell differentiation  
Nonsense mediated decay enhanced by the exon junction complex  
SRP-dependent co-translational protein targeting to membrane  
Hemostasis  
Metabolism of mRNA  
Platelet activation, signalling and aggregation  
HTLV-I infection  
Innate immune system  
Adaptive immune system

---

### **CC-IL7R CD14+ monocytes additional enriched pathways**

---

SLE  
Tuberculosis  
Epstein-Barr virus infection  
Immune System

---

**Table A.5: Additional enriched pathways for the DEGs between SF and PB CD14<sup>+</sup> monocytes from the CC-mixed and CC-IL7R subpopulations.** All the enriched pathways contained a minimum of ten DEGs from the analysis and were significant at an FDR<0.01.

## Appendices

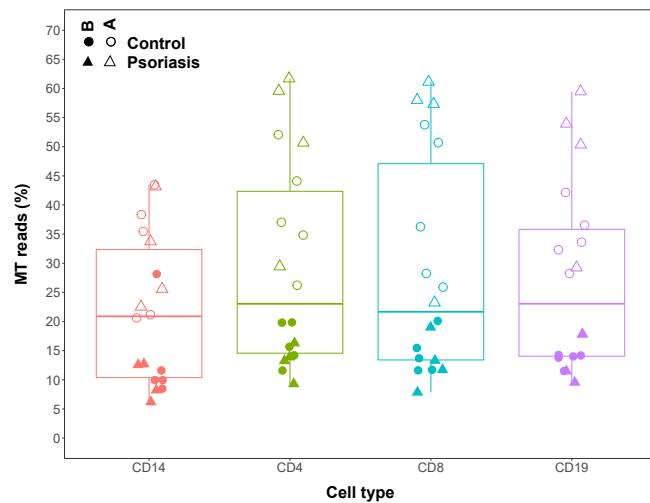
### A.1.2 Chapter 4 Tables

### A.1.3 Chapter 5 Tables

## A.2 Additional figures

### A.2.1 Chapter 3 Figures

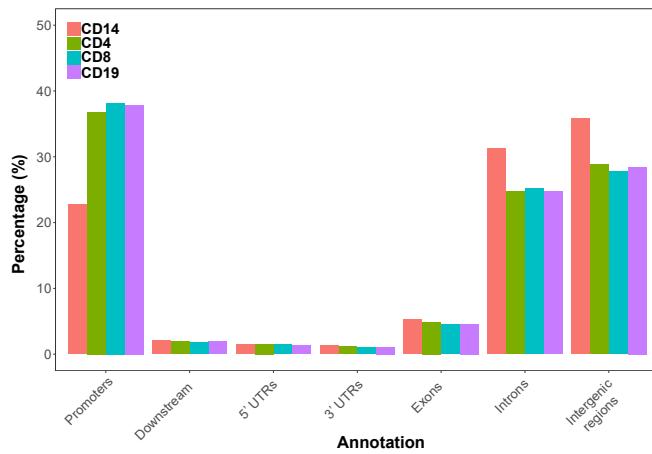
### A.2.2 Chapter 4 Figures



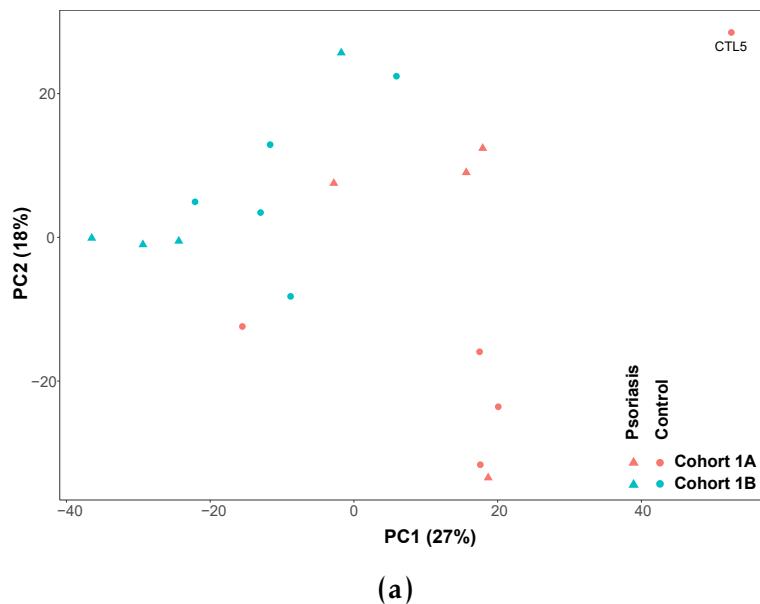
**Figure A.3: Percentage of MT reads in the ATAC-seq samples generated in CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup> isolated from psoriasis patients and healthy controls.** Samples from cohort 1A (open circles and triangles) were generated with the standard ATAC-seq protocol from Buenrostro *et al.*, 2013 whereas samples from cohort 1B (filled circles and triangles) were processed using FAST-ATAC (**Corces2016**).

## Appendices

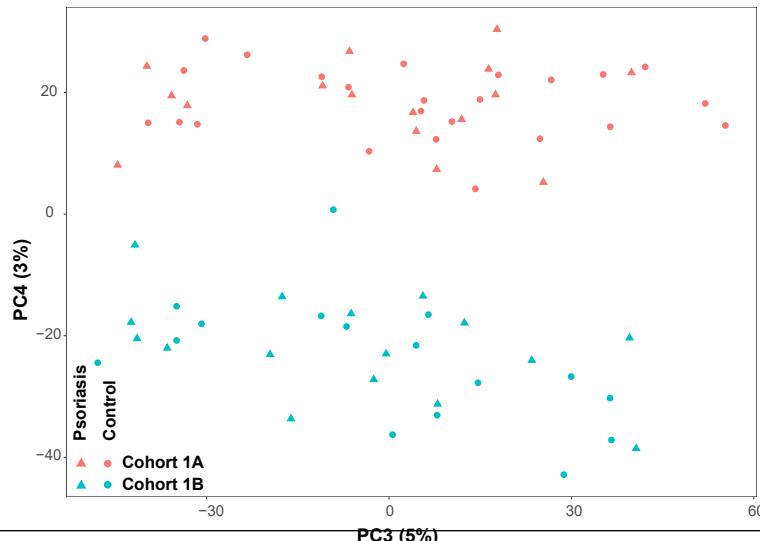
---



**Figure A.4: Genomic annotation of the consensus master list of ATAC-seq enriched sites built for downstream differential chromatin accessibility analysis in CD14<sup>+</sup> monocytes, CD4<sup>+</sup>, CD8<sup>+</sup> and CD19<sup>+</sup>. Annotation is expressed in percentage over the total number of ATAC-seq sites included in each particular cell type master list.**



(a)

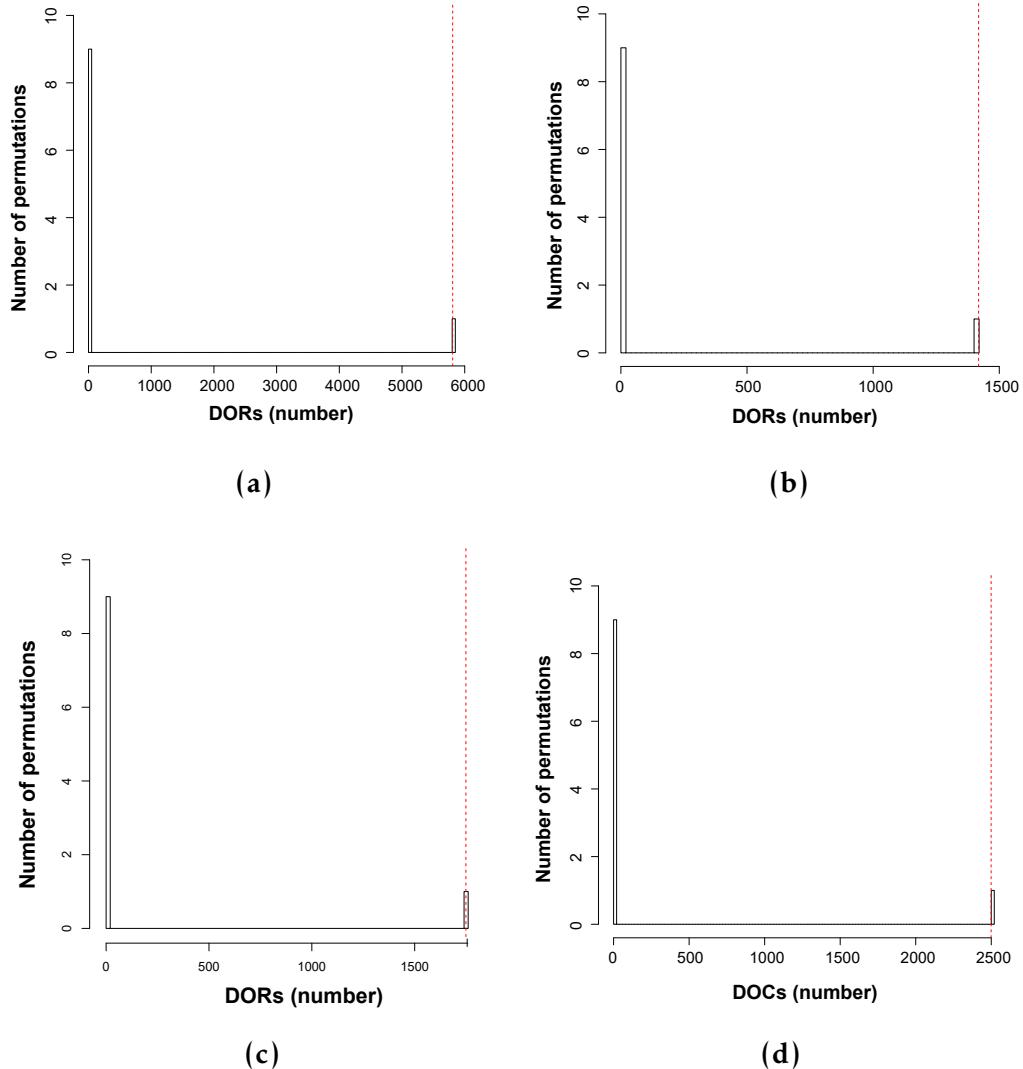


(b)

## Appendices

---

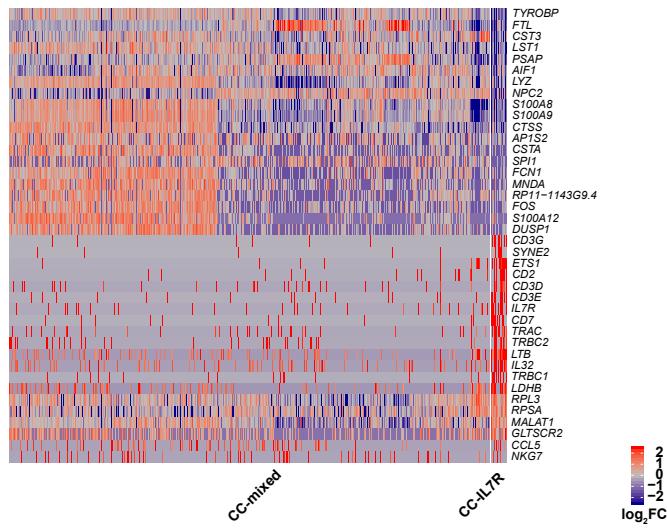
### A.2.3 Chapter 5 Figures



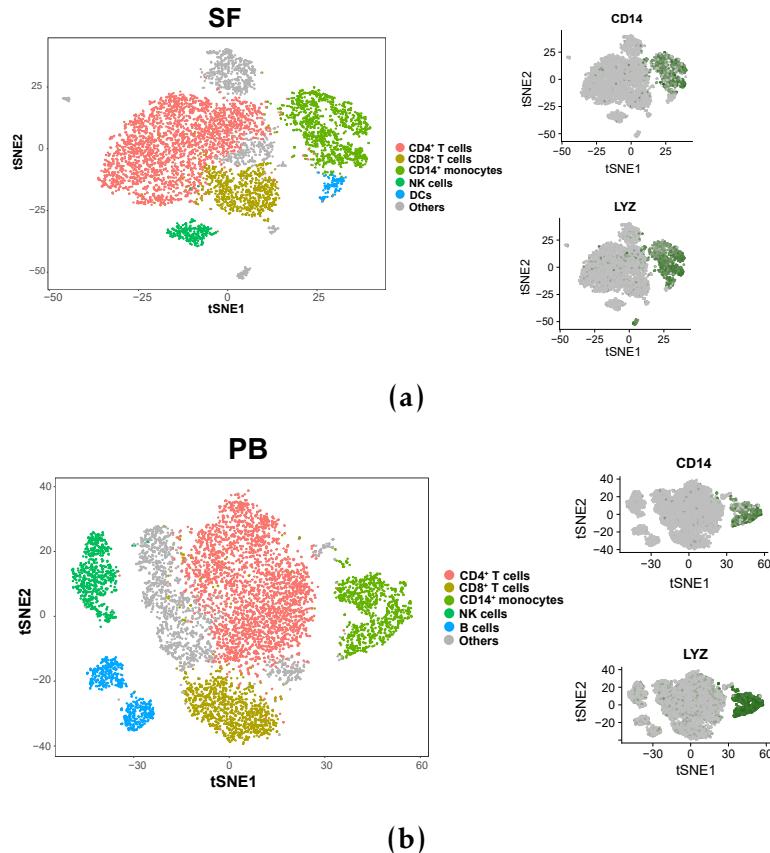
**Figure A.6:** Permutation analysis SF vs PB in CD14<sup>+</sup>,CD4m<sup>+</sup>,CD8m<sup>+</sup> and NK.

## Appendices

---



**Figure A.7: Heatmap for the top 20 marker genes of the CC-mixed and CC-IL7R CD14<sup>+</sup> monocytes subpopulations.** Rows are the top 20 marker genes for each of the two subpopulations (total of 40 genes). The columns represent each of the cells members of the CC-mixed (left) or CC-IL7R (right) clusters. The colour scale represents the log<sub>2</sub>FC in the expression of the marker gene in a particular cell of the cluster compared to the average expression of all the cells from the other cluster.



**Figure A.8: Identification of the CD14<sup>+</sup> monocytes populations from bulk SFMCs and PBMCs using scRNA-seq transcriptomes.** XXXX