Chapter 4

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

# Introduction

## The systemic and skin-specific manifestations of psoriasis

In psoriasis, skin lesions represent the main manifestation of the dysregulated innate immune response triggered by the interaction between genetic and environmental factors (reviewed in Chapter 1). In addition to keratinocytes, other circulating immune cells, such as T cells or DCs, are actively recruited to the site of inflammation contributing to disease initiation and progression (Leanne2012). A number of studies have identified systemic components of psoriasis, including an increase of circulating Th-17, Th-1 and Th-22 cells in patients’ blood and the impaired inhibitory function of circulating Tregs (Sugiyama2005; Kagami et al. 2010). Activated T cells isolated from psoriasis patients’ blood have demonstrated their ability to induce skin lesions in xenotransplantation models of psoriasis (Wrone-Smith and clinical 1996; Nickoloff and Wrone-Smith 1999). Psoriasis patients also present increased risk for PsA following skin lesions as well as other co-morbidities, such as CVD (Ibrahim et al. 2009; Shapiro et al. 2007). Overall, these findings reinforce

there being a systemic component in psoriasis and highlight the importance of investigating relevant circulating immune cells to better understand disease pathophysiology.

## The personalised epigenome in disease

The technical revolution in the epigenetics field has opened an avenue to profile the epigenome of individual cell type populations in clinical samples, contributing to the interpretation and understanding of GWAS non-coding variants. ATAC-seq and ChIPm have enabled the interrogation of chromatin accessibility, histone modifications and TF binding using a few thousand cells (Buenrostro2013; Schmidl2015). This has facilitated mapping the regulatory landscape in a wide range of cell types and tissues from clinical samples, providing details about the molecular programming of cells and the location and status of *cis*-regulatory elements in a disease-specific manner.

ATAC-seq has been used to identify inter- and intra- individual differences and pathological changes in chromatin accessibility (Qu et al. 2015). For example, differential analysis in B cells isolated from SLE patients and healthy controls has revealed changes in chromatin accessibility near genes involved in B cell activation and enriched for TFBS potentially regulating pathogenic processes (Scharer et al. 2016). Similarly, a study in age-related macular degeneration (AMD) has identified the retina epithelium as the main tissue driving disease onset through global loss of chromatin accessibility in comparison to healthy tissue (Wang et al. 2018).

In addition to the study of chromatin accessibility, the characterisation of histone modifications provides further functional information to understand the cell type specificity of the regulatory landscape. For example, in chronic lymphocytic leukaemia ChIPm has been used to identify subtype-specific epigenome signatures based on the interrogation of several histone marks (Rendeiro et al. 2016). As GWAS SNPs are mostly located in intergenic regions that may act as gene expression regulatory elements, assessing the active enhancer mark H3K27ac is of particular interest here. SNPs tagging all psoriasis GWAS loci showed enrichment for H3K27ac in twenty-six relevant cell types and/or tissues, with approximately 50% of variants in LD with those lead SNPs overlapping annotated enhancers in total CD4+ T cells (Lin et al. 2018). However, disease-specific data to use in this type of analysis is currently unavailable for most of the complex diseases. A relevant example of enhancer profiling through H3K27ac assay has been conducted in the autoimmune disease juvenile idiopathic arthritis, where a disease-specific H3K27ac super-enhancer (those spanning up to 50Kb) signature has been identified in SF mCD4+ cells (Peeters et al. 2015). In addition to this, inhibitors of histone de-acetylases (HDACs) are being investigated as potential therapeutic agents for RA and SLE, amongst others (Hsieh et al. 2014; Shu et al. 2017).

## Transcriptional profiles in psoriasis

Trancriptomics in psoriatic skin

Characterisation of transcriptional profiles in complex diseases has been performed to better understand disease pathophysiology and assess the role of genetic variability in regulating gene expression. In psoriasis, the majority of transcriptional studies have been performed for inflamed skin (lesional) using pre-lesional (uninvolved) skin, adjacent to the lesion, as the best internal control accounting for biological variability (Table 4.1).

Table 4.1: Summary table of the most comprehensive transcriptional studies in psoriasis skin and blood. SB= whole skin biopsy; EpB=epidermal biopsy; CK=cultured keratinocytes; C=control; L= psoriatic lesional skin; U=psoriatic uninvolved skin.

Author and year Sample type and size Technology Description

(Jabbari2011) SB (L=3, U=3) RNA-seq and microarray

(Li2014) SB (L=92, C=82) RNA-seq and

microarray

Technology discrepancies

Technology discrepancies and lncRNAs targets co-regulation

(Keermann et al. 2015)

SB (L=12, U=12, C=12) RNA-seq Dormant psoriasis signature and *IL36*

expression in psoriasis skin

(Tsoi et al. 2015) SB (L=97, U=29, C=90) RNA-seq Psoriatic skin-specific new lncRNAs

(Swindell and Genome 2015)

SB (L=14, U=14) RNA-seq and mass- spectometry

209 co-regulated mRNA-proteins

(Swindell et al. 2017)

(Tervaniemi et al. 2016)

CK (L=4, U=4, C=4) RNA-seq Decreased differentiation gene signature in

lesional skin

EpB (L=6, U=6, C=9) RNA-seq NOD-like and inflammasome pathways

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

131

(Coda et al. 2012) PBMCs (PS=6, C=5) and SB (L=5, U=5)

Microarray Partial overlap between PBMCs and skin DEGs

(Lee et al. 2009) PBMCs (PS=5, C=8) Microarray 202 DEGs, circulating gene expression signature

(Mesko2015) PBMCs(PS=15, IBD=12,RA=12,

C=18)

TaqMan customised array (96 genes)

6 psoriasis-specific DEGs

(Palau et al. 2013) Activated CD4++ and CD8+

(PS=17, C=7)

(Jung et al. 2004) IL-10 stimulated PBMCs and

CD14+ (C=5), IL-10 therapy PBMCs (PS=4)

Microarray 42 DEGs in T cell activation (*SPATS2L* and *KLF6*)

Microarray High correspondence between *in vitro* and *in*

*vivo* IL-10 driven DEGs

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

132

Other studies have also incorporated healthy control skin biopsies to ascertain the extent of dysregulation of the transcriptomic profile prior to lesion development in uninvolved skin (Table 4.1). Interestingly, discrepancies regarding the transcriptional similarities between normal and uninvolved skin have been identified, likely due to different filtering criteria for magnitude of effect (Keermann et al. 2015; Tsoi et al. 2015).

The latest transcriptomic studies in psoriasis using RNA-seq have demonstrated greater sensitivity as well as the ability to identify non-coding RNA species, such as lncRNAs, in an unbiased way (Jaabari2011; Li2014). LncRNAs expression has also been proven to have a role in psoriasis pathophysiology, showing approximately 1,000 species differentially expressed between lesional and uninvolved skin (Tsoi et al. 2015). Interestingly, comparison of protein abundance and DGE in psoriatic skin has revealed that only 5% of the dysregulated transcripts present a similar trend at the protein level (Swindell and Genome 2015).

The majority of the transcriptional studies have been performed in whole skin biopsies containing a mix of tissues from the epidermis, dermis, basal layer, muscle and adipose tissue (Table 4.1). Lately, studies in psoriatic cultured keratinocytes (from lesional and uninvolved biopsies) and epidermis from split-thickness skin grafts have identified differences in gene expression and functional pathway enrichment compared to the studies based on whole skin biopsies (Swindell et al. 2017; Tervaniemi et al. 2016). These results reinforce the importance of using homogenous tissue and cell type samples to better dissect the altered biological processes contributing to the development of psoriasis at the site of inflammation.

Transcriptomics in circulating immune cells

A limited number of comprehensive transcriptional studies comparing circulating immune cells between psoriasis patients and healthy controls have been conducted. The majority of these studies have investigated changes in gene expression between psoriasis and healthy controls in mixed PBMC populations using microarray technologies (Table 4.1).

A study conducted by Coda and colleagues explored the overlap between the differentially expressed genes (DEGs) in PBMCs (psoriasis versus controls) and comparing lesional and uninvolved skin biopsies (Coda et al. 2012). 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 performed *in vitro* culture and stimulation of T cells and monocytes (Palau et al. 2013; Jung et al. 2004). For instance, Palau and colleagues found forty-two DEGs enriched for cytokine and IFN (*α*,

*β* and *γ*) signalling pathways when comparing activated CD4+ and CD8+ T

cell from psoriasis patients and healthy controls. Further understanding of psoriasis-specific systemic gene dysregulation has also been approached through comparison with other chronic inflammatory diseases (Mesko2015).

## Chromatin accessibility, gene expression and genetic variability

As described in Chapter 1, accessible chromatin is more likely to be bound by TFs and other co-regulatory proteins, and so can be used as a proxy to tag genomic loci involved in regulation of gene expression and to infer the putative functional relevance of GWAS SNPs. The orchestration of cell type specific changes in the chromatin landscape and gene expression is pivotal for an appropriate immune response (Goodnow et al. 2005). For example,

integration of ATAC-seq data and gene expression in pancreatic islets has revealed chromatin accessibility to be a better predictor for gene activation in *α*- compared to *β* cells, which could be explained by the heterogeneity within each cell population or cell type intrinsic differences in gene regulation. In AMD clinical samples, integration of ATAC and gene expression found moderate correlation between the two in retina and pigmented epithelium retina (Wang et al. 2018). In the context of genetic variability, the relationship between chromatin accessibility and gene expression in homeostasis and stimulated conditions has been addressed by integrating eQTL and chromatin accessibility QTLs (ca- QTLs). For example, enhancer priming events have been described in human iPS derived macrophages, where the same genetic variants leads to changes in chromatin accessibility in the na¨ıve state prior to changes in gene expression upon stimulation (Alasoo et al. 2018).

## Fine-mapping using summary stats

The generation of cell type specific epigenetic maps can be used to inform statistical fine-mapping in the effort to identify putative causal SNPs to undergo functional validation (detailed in Chapter 1). Integration of Bayesian fine- mapping for twenty-one complex immune diseases performed by Farh and colleagues demonstrated greatest enrichment of fine-mapped causal variants in immune cell enhancer elements, particularly from activated conditions (Farh2015). In this study, psoriasis PICS showed the most significant enrichment for Th-1, Th-2 and Th-17 subsets.

Traditional Bayesian fine-mapping requires genotyping data from the GWAS cohorts to perform genotype phasing and imputation prior to association analysis and calculation of PP and credible sets of SNPs. Restricted access to GWAS genotyping data, commonly due to ethical reasons, can be a limitation when performing this type of analysis. Since summary statistics from GWAS

studies are widely available, methods like DIST have been developed to impute summary statistics instead of genotypes for the unmeasured SNPs in the study (Lee et al. 2013). In addition to this, summary statistics Bayesian fine-mapping methods using functional annotation as a prior in the model have also been developed. For example, the Risk Variant Inference using Epigenomic Reference Annotation (RiVIERA) method has been applied to perform fine-mapping for the Immunochip GWAS associated loci, incorporating in the model the forty- three ENCODE and Epigenome Roadmap annotation features showing greatest enrichment for psoriasis risk SNPs (Li and research 2016).

# Aims

The aim of this chapter is to determine chromatin accessibility, histone modification and gene expression differences between psoriasis patients and controls in four circulating immune cell types (CD14+ monocytes, CD4+ and CD8+ T cells, and CD19+ B cells), and to complement this with analysis of differential gene expression in lesional and uninvolved epidermis isolated from psoriatic skin biopsies. The long term goal is to identify disease and cell type specific changes in putative regulatory regions and integrate them with observed differences in gene expression to improve the understanding of systemic and skin inflammatory features of psoriasis and prioritise putative causal GWAS variants.

The specific aims for this chapter are:

* + 1. To identify differences in chromatin accessibility and the H3K27ac active enhancer mark between psoriasis patients and healthy controls in immune cells isolated from peripheral blood.
    2. To determine changes in genes expression between psoriasis patients and healthy controls in immune cells isolated from peripheral blood.
    3. To identify differentially expressed genes between lesional and uninvolved epidermis isolated from psoriatic skin biopsies.
    4. To compare the differences in the transcriptomic profile from circulating immune cells between patients and controls with the transcriptional differences from contrasting lesional and uninvolved epidermis.

# Results

## Psoriasis and healthy controls: cohort description and datasets

Peripheral blood samples were collected from a cohort of psoriasis patients and healthy individuals in order isolate four relevant immune cells types (CD14+ monocytes, CD4+, CD8+ and CD19+) 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.

A total of eight psoriasis patients, six males and two females (Table 4.2) were recruited following eligibility criteria detailed in Chapter 2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sample ID | Sex | Age at diagnosis | Disease duration (months) | PASI | Nails affected | Family history |
|  |  |  |  |  |  |  |
| Cohort 1A |  |  |  |  |  |  |
| 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 |
|  |  |  |  |  |  |  |
| Cohort 1B |  |  |  |  |  |  |
| 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 4.2: Description and metadata of the psoriasis patients cohort. For each of the individuals information relating to sex, age at the time of sampling, disease duration, PASI score, nail involvement and family history has been recorded. Patients are divided into cohort 1A and cohort 1B based on the timing (batch) of ATAC-seq and RNA-seq processing and type of ATAC-seq protocol applied. PASI evaluates the percentage of affected area and the severity of redness, thickness and scaling for four body locations (as detailed in Table ??). The skin RNA-seq samples include lesional and uninvolved paired-skin biopsies from each of the three individuals.

The average age of the cohort was 55 years old and the average disease duration 331.5 months. All the patients showed active skin disease and none of them had reported joint involvement at the time of sample collection. Disease severity was quantified using the PASI score, previously reviewed in Chapter 1, with the average cohort score being 10. Currently, there is no consensus on PASI thresholds to define mild and moderate-to-severe disease. A review study regarding the use of PASI as an instrument to determine disease severity of chronic-plaque psoriasis have suggested considering psoriasis as moderate when PASI ranges between 7 to 12 and severe for PASI*>*12 (Schmitt and Dermatology 2005). On the other hand, NICE and other studies had defined psoriasis as severe based on PASI≥10 (Woolacott et al. 2006; Finlay 2005). In this cohort, six out of ten patients had PASI≥10, and so were categorised as having severe psoriasis. Only two of them showed PASI*<*7 showing a mild psoriasis phenotype. All patients were na¨ıve for biologics therapies. PS2319 was currently on mitochondrialX therapy and the remaining patients had only been treated occasionally with topical steroids or UVB therapy. Interestingly, PS2014 showed the most severe PASI score (17) and was a non-responder to mitochondrialX. Patients PS1011, PS2015, PS2001 and PS2314 had nail pitting, which has been defined as one of the markers for increased risk of developing joint affection and PsA (Moll1976; McGonagle; Griffiths and Lancet 2007; medicine 2011). A family history of psoriasis was reported by PS2000 and PS2319. In addition to the psoriasis samples, PB blood was collected from ten concurrently recruited sex and age-matched healthy individuals (Table 4.3)

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

Average − 46.6

Table 4.3: Description of the healthy control cohort. Controls are divided in cohort 1A and cohort 1B based on the timing (batch) of ATAC-seq and RNA-seq processing and ATAC-seq protocol applied, similarly to the psoriasis patient samples.

For both cohorts, ATAC-seq and RNA-seq data were generated from CD14+ monocytes, CD4+, CD8+ and CD19+ cells (Table 4.2 and 4.3). 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 (Corces et al. 2016) in cohort 1B, due to the improvements of this protocol as explained in Chapter 3. Additionally, samples from cohort 1B were also processed to assess differences in H3K27ac modification between patients and controls using ChIPm. 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 4.3). 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 constraints.



## Investigation of psoriasis-specific changes in the enhancer mark H3K27ac in different peripheral blood immune cell populations

Data processing and quality control

A total of 32 ChIPm libraries from four patients and four controls in four peripheral blood immune cell types were sequenced, and reads filtered as detailed in Chapter 2. After filtering, the total number of reads ranged between 46.9 and 60.5 million, compliant with the 40 million total reads recommended by ENCODE (Figure 4.1 a). As part of the quality control, library complexity for each of the samples was measured based on the non-redundant fraction and 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). The CD8+ CTL7 and the CD19+ PS2000 and PS2314 libraries failed the recommended complexity non-redundant fraction values and also had more severe PCR bottlenecking (based on PBC1 coefficient threshold). These observations were consistent with the greater

number of duplicated reads identified in these libraries compared to the rest (*>*50% of the total sequenced reads) and consequently lower number of reads after filtering (Figure 4.1 a).

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 showed appropriate signal-to-noise following ENCODE standards (Landt et al. 2012), with NSC and RSC values equal or greater than 1.05 and 0.8, respectively (Figure 4.1 b and c). Interestingly, the CD14+ monocytes and CD4+ ChIPm libraries had lower signal enrichment compared to the CD8+ and CD19+ libraries, which correlates with the cell type grouping during sample processing. Following QC, the CD8+ CTL7 and the CD19+ PS2000 and PS2314 libraries were removed for downstream analysis.

PCA 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 4.2 a). When performing PCA analysis by cell type, the PS2314 CD8+ library appeared as an outlier compared to the rest of CD8+ H3K27ac ChIPm patients and control libraries and was also removed for downstream analysis (data not shown).

Differential H3K27ac enrichment analysis

Differential H3K27ac analysis was performed between psoriasis and healthy control samples for each cell type using DiffBind. The consensus merged list of H3K27ac sites assembled by this algorithm to perform differential analysis (as explained in Chapter 2) included a high percentage of sites annotated as heterochromatin or repetitive (Figure 4.2 b), ranging from 20.8% in CD14+ monocytes to 51.39% in CD8+ cells. Such sites are less likely to be relevant

90

PS2314

PS2001

CTL7

**Control**

**Psoriasis**

80

70

**Total reads after filtering (million)**

60

50

40

30

20

10

0

CD14 CD4 CD8 CD19

**Cell type**

(a)

5.0 3.5



**Control**

**Psoriasis**



**Control**

**Psoriasis**

4.5

3.0

4.0

3.5

2.5

3.0 2.0

**NSC**

**RSC**

2.5

1.5

2.0

1.0

1.5

1.0

CD14 CD4 CD8 CD19

**Cell type**

(b)

0.5

CD14 CD4 CD8 CD19

**Cell type**

(c)

Figure 4.1: Quality control evaluation of the H3K27ac ChIPm libraries in immune cells isolated from psoriasis and control samples. For each of the cell types boxplots representing a) million of reads after filtering, b) normalised strand cross-correlation coefficient (NSC) and c) relative strand cross-correlation coefficient (RSC). NSC and RSC are measures of signal enrichment independent of peak calling, where 1 and 0 indicate no enrichment, respectively. In c) and d) the dashed red line indicates the ENCODE threshold for low enrichment (NSC*<*1.05 and RSC*<*0.8). For each point, colour codes for cell type and shape for phenotype (psoriasis or control).

**PC2 (30%)**

0

−100

−150 −100 −50 0 50

**PC1 (41%)**

**CD14 CD4 CD8**

**CD19**

**Psoriasis**

**Control**

(a)

**20.80**

**33.66**

**45.23**

**51.39**

**32.64**

**19.94**

**18.09**

**15.75**

**35.61**

**33.93**

**29.41**

**24.35**

**3.54**

**6.37**

**4.54**

**5.66**

**7.14**

**3.45**

100

75

**Chromatin state**

1. **Active Promoter**
2. **Weak Promoter**

**Percentage (%)**

1. **Poised Promoter**
2. **Strong Enhancer**

50 **5 Strong Enhancer**

1. **Weak Enhancer**
2. **Weak Enhancer**
3. **Insulator**
4. **Txn Transition**
5. **Txn Elongation**
6. **Weak Txn**
7. **Repressed**

25 **13 Heterochrom/lo**

1. **Repetitive/CNV**
2. **Repetitive/CNV**

0

CD14 CD4 CD8 CD19

**Cell type**

(b)

Figure 4.2: PCA and chromatin annotation states of the H3K27ac enriched sites in four immune primary cell types from psoriasis and healthy control samples.

a) PCA 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+ monocytes, CD4+, CD8+ and CD19+ cells. The first two PCs (x- axis and y-axis, respectively) for all the H3K27ac ChIPm peaks in the master list are plotted. 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 individual cell type master list.

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+ monocytes had the greatest number of differentially modified enhancers (8 significant sites), followed by CD4+ (4) and CD8+ (1) (Table 4.5).

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

Table 4.5: Summary results from the differential H3K27ac analysis between psoriasis patients and healthy controls in CD14+ monocytes, CD4+, CD8+ and CD19+ cells. In the genome-wide analysis, the master list size refers to the number of H3K27ac enriched sites included in the consensus list built using 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 significant sites in CD14+ monocytes and CD8+ also contain the sites identified in the enhancer restricted analysis. Significant differentially H3K27ac modified regions were determined using FDR*<*0.05 and no FC threshold.

An example of a differentially H3K27ac modified region when comparing patients versus controls in CD14+ monocytes was located between the *SLC15A2* and *ILDR2* genes (Figure 4.3). *ILDR2* has recently been identified as relevant for negative regulation of T cells response in RA (Hecht et al. 2018). This region showed lower H3K27ac levels in psoriasis patients compared to controls and was annotated as enhancer by the Epigenome Roadmap chromatin segmentation map. Additionally this site was overlapping a DHS and H3Kme1 (enhancer mark) modification and a CTCF-binding site identified by ChIP-seq in K562 cells.

**UCSC genes**

Scale chr3:



20 kb hg19

121,670,000  121,680,000 121,690,000 121,700,000  121,710,000 

*SLC14A2*

*ILDR2 *

**ENCODE DHS peaks **   

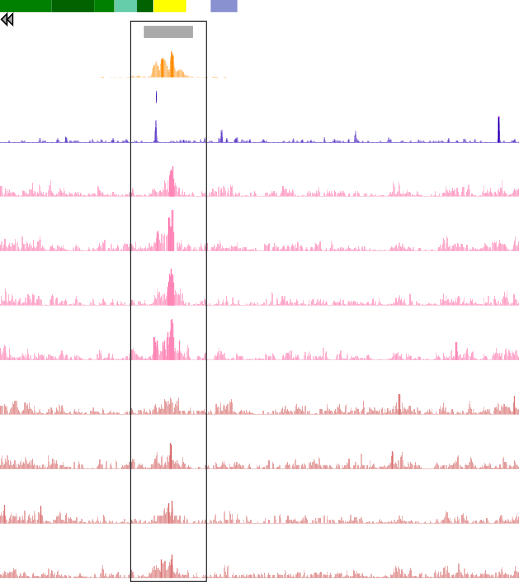
296 \_

**ENCODE DHS**

1 \_

chromHMM tracks from Roadmap

**ChromHMM RoadMap**



**ENCODE H3K4me1 peaks**

79.56 \_

**ENCODE H3K4me1**

0.04 \_

**ENCODE CTCF peaks**

0.1 \_

**ENCODE CTCF**

**CTL7**

**CTL8**

**CTL9**

**CTL10**

**PS2000**

**PS2001**

**PS2314**

**PS2319**

0 \_

15.57 \_

0 \_

15.57 \_

0 \_

15.57 \_

0 \_

15.57 \_

0 \_

15.57 \_

0 \_

15.57 \_

0 \_

15.57 \_

0 \_

15.57 \_

0 \_

Figure 4.3: Differential H3K27ac modification at a putative intergenic enhancer region in circulating CD14+ monocytes between psoriasis patients and healthy controls. UCSC Genome Browser view illustrating the normalised H3K27ac fold- enrichment (y-axis) at an intergenic differentially modified region located between *SLC14A2* and *ILDR2* genes (x-axis) in CD14+ monocytes (lower H3K27ac enrichment in psoriasis patients compared to healthy controls). CD14+ monocytes publicly available epigenetic data from ENCODE (including DHS, H3K4me1 and CTCF ChIP-seq) and the Epigenome Roadmap chromatin segmentation track are also shown. Differential H3K27ac modified regions were considered significant based on FDR*<*0.05 and no FC cut-off. H3K27ac tracks are colour-coded by condition: control(CTL)=pink and psoriasis (PS)=sienna.

Overall, restricting the differential analysis to enhancer annotated regions did not significant increase the number of observed 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 for these cell types and sample size.

## Identifying global changes in chromatin accessibility between psoriasis patients and healthy controls for different peripheral blood immune cell populations

In order to interrogate genome- wide changes in chromatin accessibility between patients and controls, ATAC-

seq was performed in the same four cell types in eight patients and ten controls (Table 4.4) giving a total of 72 libraries.

Data processing and quality control

ATAC-seq quality control showed that the median total reads after filtering ranged between 39.2 and 49.8 million in CD4+ and CD19+ cells respectively (> 15 million reads determined as appropriate minimum in Chapter

3) for all samples (Figure 4.4 a).

The differences in read depth across ATAC samples are mostly due to intrinsic difficulties in determining molarity of the ATAC libraries (due to the fragment size heterogeneity) as well as the number of duplicates present in each library, which also correlated with mitochondrial reads. Differences in the percentage of mitochondrial reads were noticeable between samples from cohort 1A generated with the standard ATAC-seq protocol from Buenrostro *et al.*, 2013 and the Fast-ATAC libraries from cohort 1B using the later modified (Corces et al. 2016) protocol (Figure ??).

All the samples showed the required characteristic ATAC-seq fragment size distribution recapitulating nucleosome periodicity (previously detailed in Chapter 3). Analysis of ATAC-seq signal enrichment across gene TSSs revealed that most of the samples had enrichment over 6 (Figure 4.4 b) and only PS2000 and PS2001 CD14+ monocytes were removed from downstream analysis due to low signal-to-noise ratios (*<*6). When comparing the number of peaks passing IDR filtering in each samples versus the number of reads after filtering,

most of the samples showed between 10,000 and 35,000 peaks (Figure 4.4 c). Since the sequencing depth of most samples was ≥15 million reads, the majority 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 3. For example, CD14+ monocytes had greater numbers of peaks when compared to the

130

120

110

**Total reads after filtering (million)**

100

90

80

70

60

50

40

30

20

10

CD14 CD4 CD8 CD19

**Cell type**

(a)

30

25

**Control**

**Psoriasis**

CTL5

**Control**

**Psoriasis**

PS2314

CTL6

PS2314

CTL4

PS2015 CTL8

PS2014 CTL2

CTL6

PS2319 PS2314

CTL5

PS2014 CTL4

CTL4

PS2016

CTL9

PS2016

6

PS2015

CTL1

CTL7 CTL1

CTL5

PS1011 TL2

PS2001

1

PS2001

CTL2

PS1011

CTL8 CTL9 PS2001 CTL10

CTL2

PS2000

PS2001

PS2000

CTL1

CTL3 PS101

PS2314 CTL7 CTL6

PS2319

PS2000 CTL3

CTL10 CTL9

CTL8 CTL7

CTL10

PS2319

PS1011

CTL3

PS2014 PS201 PS2015

|  |  |
| --- | --- |
| PS2014 PS2016 | CTL6 CTL1 |
| CTL8 CTL7  PS2015  PS2319 | |
| CTL4 CTL5  CTL10  PS2000 CTL9 | |
| C  CTL3 |  |

20

**TSS enrichment**

15

10

5

0

CD14 CD4 CD8 CD19

**Cell type**

(b)

140

CTL6

CTL4

PS2314

CTL1

PS1011

**CD14 CD4 CD8 CD19**

CTL3

CTL2

**Psoriasis Control**

130

120

110

**Reads after filtering (million)**

100

90

80

70

60

50

40

30

20

10

0

0 10 20 30 40 50 60 70 80

**Number of peaks (x103)**

(c)

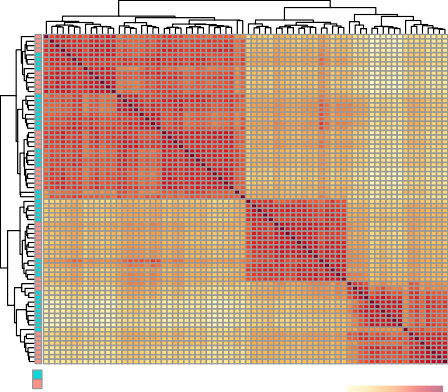
Figure 4.4: Quality control assessment of the ATAC libraries generated from circulating immune cells in psoriasis and control samples. For each of the cell types and samples, boxplots show a) million of reads after filtering, b) values for fold- enrichment of ATAC fragments across the Ensembl annotated TSS and c) representation of the number of significant peaks based on IDR optimal pval versus the total million reads after filtering for each of the samples. In b) the dashed red line indicates the recommended Encode threshold for TSS enrichment values. For each point, colour codes for cell type and shape for phenotype (psoriasis or control). In b) and c) sample IDs are included for all or some of the samples.

other three cell types, despite the median total number of reads after filtering being similar to the other cell types (Figure 4.4 a). CD19+ CTL2 appeared to be an outlier, with a noticeably lower number of peaks for its high sequencing depth (Figure 4.4 c). This observation together with its border line TSS enrichment supported removal of the CD19+ CTL2 sample from downstream analysis.

Functional relevance of the ATAC regions included in the differential analysis

A heatmap illustrating sample distance 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+ monocytes, CD19+ B cells and CD4+/CD8+ T cells (Figure

* 1. 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, within the cell types there was some grouping of samples by batch (Figure 4.5 a).

PS2001\_CD8 CTL6\_CD8 CTL8\_CD8 PS2314\_CD8 PS2015\_CD8 PS2014\_CD8 PS2016\_CD8 CTL4\_CD8 CTL9\_CD8 CTL10\_CD8 PS2319\_CD8 CTL7\_CD8 PS2000\_CD8 CTL1\_CD4 CTL2\_CD4 CTL3\_CD4 PS1011\_CD4 PS1011\_CD8 CTL1\_CD8 CTL2\_CD8 CTL3\_CD8 CTL9\_CD4 PS2000\_CD4 CTL10\_CD4 PS2001\_CD4 PS2015\_CD4 PS2016\_CD4 CTL4\_CD4 PS2014\_CD4 CTL6\_CD4 CTL7\_CD4 PS2314\_CD4 CTL8\_CD4 PS2319\_CD4 CTL5\_CD4

**Cohort**

**CD4+**

**and CD8+**

**Cohort 1A Cohort 1B**

(a)

**Distance**

80

60

40

CTL5\_CD8 CTL5\_CD19

PS2015\_CD19 PS2016\_CD19 CTL4\_CD19 PS2014\_CD19 CTL9\_CD19 PS2001\_CD19 CTL10\_CD19

**CD19+**

PS2319\_CD19 CTL7\_CD19 PS2314\_CD19

CTL8\_CD19 PS2000\_CD19 CTL1\_CD19 PS1011\_CD19 CTL3\_CD19 CTL2\_CD19

CTL6\_CD19 PS2001\_CD14

PS2000\_CD14 CTL1\_CD14 CTL3\_CD14 PS2015\_CD14 PS2016\_CD14 PS1011\_CD14 CTL2\_CD14 CTL4\_CD14 PS2014\_CD14 CTL5\_CD14

**CD14+**

CTL7\_CD14 PS2319\_CD14

CTL6\_CD14 CTL8\_CD14 PS2314\_CD14 CTL9\_CD14 CTL10\_CD14

100

20

0

NFKB Q6 USF Q6 STAT3 02

CP2 01 NFKB C CREL 01

HMX1 01

HTF 01

YY1 01

MYCMAX 01

NFKAPPAB 01

NFKAPPAB50 01

MYOD 01

RFX1 01 SEF1 C AP2ALPHA 01

NFE2 01

HNF4 01

ATF6 01

ARP1 01

CHOP 01

ZIC3 01

ZIC2 01 MYOD Q6 USF 01

MYCMAX 02

AREB6 02

HSF1 01

GATA1 03

SREBP1 02 MYB Q6 PAX4 04 CREB Q2 AP1 Q6 AP1 01

NCX 01

ZIC1 01

BACH2 01

CEBPB 02

HNF4 01 B

ISRE 01

BACH1 01

STAT 01

LYF1 01 NF1 Q6 PPARG 03

STAT5A 02

**CD14 CD4 CD8 CD19**



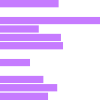
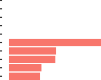








0 1 2 3 4 0 1 2 3 4 0 1 2 3 4 0 1 2 3 4

**Enrichment changes (log2)**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

(b)

Figure 4.5: Clustered heatmap and conserved TFBS enrichment analysis in the consensus ATAC regions identified in CD14+ monocytes, CD4+, CD8+ and CD19+ cells from the patients and controls cohort. a) Distance matrix and hierarchical 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 (ML 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 position-weight matrices based on experimental results in the scientific literature. Significant enrichment using FDR*<*0.01.

Each of the four cell type master lists of ATAC-seq peaks (ML CD14, ML CD4, ML CD8, and ML CD19) used for the downstream differential chromatin accessibility analysis (method detailed in Chapter 2), showed the highest percentage of regions annotated as gene promoters, intronic and intergenic, as expected for ATAC-seq (Figure ??).

The specificity and functional 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 4.5 b). For example, enrichment of conserved NF*κ*B binding motifs was identified across the cell type-specific differential ATAC peaks. Conserved binding motifs for TF involved in T cell biology, such as AREB6 (ZEB1), ATF6 and the heat-shock transcription factor HSF1 (Guan et al. 2018; Yamazaki et al. 2009; Gandhapudi et al. 2013), were enriched in ATAC-seq peaks differentially expressed in CD8+ T cells.

Differential chromatin accessibility analysis

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 list using DESeq2. PCA analysis on these data 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 B.6 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 (representative example Figure B.6 a) and was also removed from the differential analysis.

Genome-wide differential chromatin accessibility analysis revealed 55 significant (FDR*<*0.05) differentially accessible regions (DARs) between psoriasis patients and healthy controls in CD8+ cells (Table 4.6), of which 17 showed FDR*<*0.01. Conversely, CD14+ monocytes, CD4+ and CD19+ cells only showed one or no DARs.

Cell type Number of DARs

FDR*<*0.05

CD14+ 1

CD4+ 0

CD8+ 55

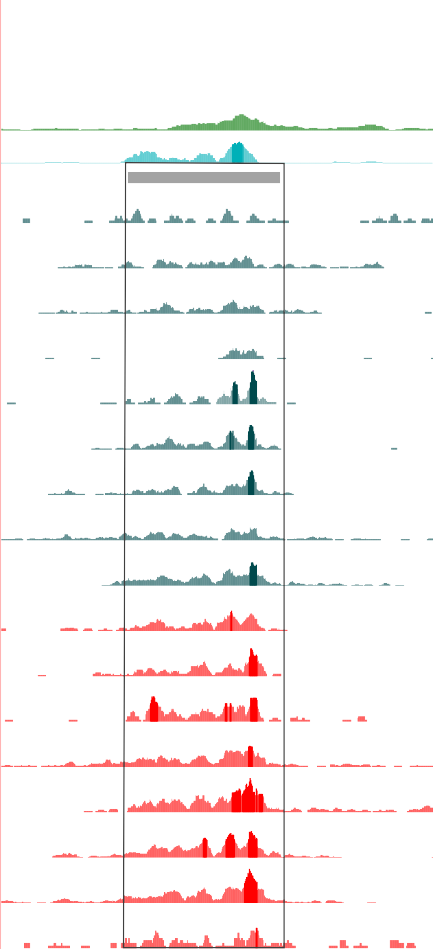
CD19+ 1

Table 4.6: Summary results from the differential chromatin accessibility analysis between psoriasis patients and healthy controls in CD14+ monocytes, CD4+, CD8+ and CD19+ cells. The number of DARs refers to those statistically significant when using a cut-off for background reads of 80% (see Chapter 3) and a FDR*<*0.05. No threshold for fold change was applied.

Annotation of the 55 CD8+ DARs using cell type specific Roadmap Epigenomics chromatin segmentation maps revealed the potential for some of the regions to be involved 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 CD8+ T cell eRNA data from the FANTOM5 project. 8 of the CD8+ DARs overlapped significantly expressed eRNAs. These included a region at the TSS of the *TNSF11* gene and another upstream the *IL7R* promoter, which were more accessible in the psoriasis patients compared to the healthy controls (Figure

* 1. a and b). The two DARs also overlap chromatin harbouring H3K4me3, a histone mark indicating an active promoter, and H3K27ac consistent with the transcription of those regions as eRNAs in CD8+ cells according to FANTOM5.

Scale



1 kb hg19

Scale chr5:

1. kb hg19

35,855,000 35,860,000



chr13: 43,147,500 43,148,000 43,148,500 43,149,000 43,149,500 43,150,000 43,150,500 43,151,000

intergenic SNPs

**UCSC genes** *TNFSF11*

**ChromHMM RoadMap**

30 \_

RoadMap H3K4me3

0 \_

15.6677 \_

In-house H3K27ac 0 \_

**Peaks ATAC-seq**

32.12 \_

**CTL1**

0 \_

32.12 \_

**CTL2**

**UCSC genes ChromHMM RoadMap**

228 \_

RoadMap H3K4me3

1 \_

33.472 \_

In-house H3K27ac

0 \_

**Peaks ATAC-seq**

134.764 \_

**CTL1**

0 \_

134.764 \_

**CTL2**

0 \_

*IL7R*

**CTL3**

**CTL4**

**CTL6**

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

**CTL3**

**CTL4**

**CTL6**

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

0 \_

**CTL7**

**CTL8**

**CTL9**

**CTL10**

**PS2014**

**PS2015**

**PS2016**

**PS2000**

**PS2001**

**PS2314**

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

0 \_

32.12 \_

**CTL7**

**CTL8**

**CTL9**

**CTL10**

**PS2014**

**PS2015**

**PS2016**

**PS2000**

**PS2001**

**PS2314**

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

0 \_

134.764 \_

**PS2319**

0 \_

32.12 \_

**PS2319**

0 \_

134.764 \_

**PS1011**

0 \_

32.12 \_

0 \_

(a)

**PS1011**

0 \_

134.764 \_

0 \_

(b)

Figure 4.6: Epigenetic landscape at two ATAC differentially accessible regions between patients and controls in CD8+ T cells. UCSC Genome Browser view illustrating the normalised ATAC read density (y-axis) in DARs located at a) the promoter of *TNFSF11* gene and b) up-stream the *IL7R* gene (x-axis). Both DARs were more open in CD8+ cells from psoriasis compared to controls. Tracks are colour-coded by condition: control(CTL)=dark turquoise and psoriasis (PS)=red. The Epigenome Roadmap chromatin segmentation map and H3K4me3 for CD8+ cells are also shown, together with a representative track from the in-house ChIPm H3K27ac in this cell type. All DARs were significant based on FDR*<*0.05 and no FC cut-off.

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 overlapping with experimentally validated eRNAs.

Integration of H3K27ac ChIPm and ATAC-seq chromatin accessibility profiles

One region showed co-localisation of observed differential H3K27ac ChIPm and ATAC between psoriasis patients and contols, and this was seen in CD8+ cells within an intron of the D-tyrosyl-tRNA deacylase 1 (*DTD1*) gene (Figure 4.7). Lower levels of H3K27ac and chromatin accessibility were found in the psoriasis patients when compared to healthy controls (Figure). This differential region was annotated as an active enhancer according to the CD8+ ChromHMM segmentation map and did not interact with the promoter of any gene according to Hi-C and promoter Hi-C data in CD8+ cells (Javierre2016). SNPs within this region are reported to be eQTL for *DTD1* in whole blood (https://gtexportal.org/home/eqtls).

## Differential gene expression in circulating immune cells in psoriasis

Data processing and quality control

In addition to characterising the chromatin accessibility landscape, gene expression profiles in psoriasis and healthy individuals were analysed for the same four primary circulating immune cell types using RNA-seq. The percentage of RNA-seq reads mapping to a unique location in the genome using STAR (see Chapter 2) was appropriate (minimum recommended 70 to 80%),

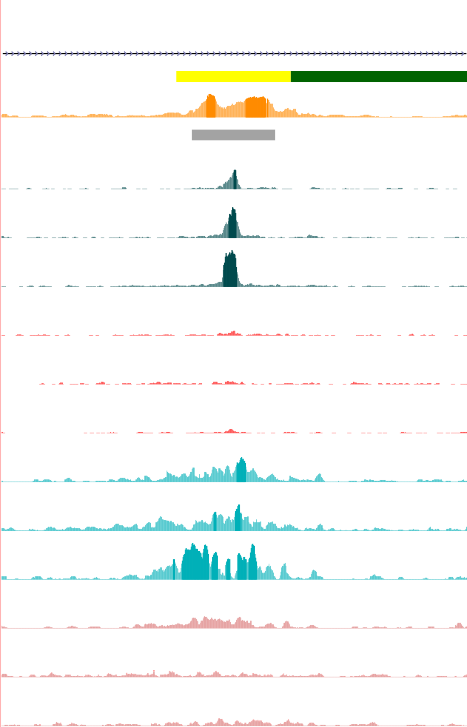
**UCSC genes**

Scale chr20:

*DTD1*

2 kb hg19

18,720,000



**ChromHMM RoadMap**

12.0263 \_

**RoadMap H3K4me1**

chromHMM tracks from Roadmap

0 \_



**Peaks ATAC-seq**

83.35 \_

**CTL8**

0 \_

83.35 \_

**ATAC-seq**

**CTL9**

**CTL10**

**PS2000**

**PS2001**

**PS2314**

0 \_

83.35 \_

0 \_

83.35 \_

0 \_

83.35 \_

0 \_

83.35 \_

0 \_

34.51 \_

**CTL8**

0 \_

34.51 \_

**ChIPm**

**CTL9**

**CTL10**

**PS2000**

**PS2001**

**PS2314**

0 \_

34.51 \_

0 \_

34.51 \_

0 \_

34.51 \_

0 \_

34.51 \_

0 \_

Figure 4.7: Epigenetic landscape at *DTD1* locus showing differential H3K27ac and chromatin accessibility between psoriasis patients and controls in CD8+ cells. UCSC Genome Browser view illustrating the normalised ATAC read density and H3K27ac normalised fold-enrichment (y-axis). Tracks are colour-coded by condition and assay: control(CTL)=dark and light turquoise and psoriasis (PS)=light and dark red, for ATAC and ChIPm respectively. The Epigenome Roadmap chromatin segmentation map and H3K4me1 for CD8+ cells are also shown.

ranging between 79.64 and 86.19% across the 72 samples (Figure 4.8 a). After appropriate filtering, all the samples had at least 20 million reads (as required by ENCODE standards) mapping to a comprehensive list of Ensembl features, including protein coding genes and lncRNAs (Figure 4.8 b). The median total reads mapping to Ensembl features was greater for CD14+ monocytes when compared to the other three cell types. For all four cell types, higher mapping rates and total 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.

88 47

**Control Psoriasis**

**Control Psoriasis**

**B A**

**B A**

45

43

**Reads mapping to Ensembl features (million)**

86

41

39

**Mapping rate (%)**

84 37

35

33

82 31

29

27

80

25

78

CD14 CD4 CD8 CD19

**Cell type**

(a)

23

21

CD14 CD4 CD8 CD19

**Cell type**

(b)

Figure 4.8: 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.

PCA using the normalised number of reads mapping to each of the 20,493 Ensembl genes passing quality control (see Chapter 2) showed that most variability was driven by cell type differences (Figure 4.9 a). A heatmap illustrating sample distance based on the expression profile of each sample followed by hierarchical clustering revealed three main clusters corresponding

to CD14+ monocytes, CD4+ and CD8+ lymphocytes, and CD19+ cells (Figure 4.9 a). Within each cell type cluster, samples were further grouped by cohort (1A and 1B) but not by condition (psoriasis and control), consistent with the differences in mapping rate and total reads mapping to Ensembl genes observed across the two cohorts (Figure 4.8 a and b). Clear correlation of sample batch with PC4 from the PCA led to a very clear separation of the samples into cohort 1A and 1B, explaining 3% of the total variance (Figure B.6 b). Consequently, batch was included in the DGE model as a covariate.

100

**CD14 CD4 CD8 CD19**

**Psoriasis**

**Control**

50

0

**PC2 (25%)**

−50

−100 0

**PC1 (49%)**

(a)

PS2014\_CD14 CTL5\_CD14 PS2016\_CD14 PS1011\_CD14 CTL1\_CD14 PS2015\_CD14 CTL6\_CD14 CTL2\_CD14 CTL3\_CD14 PS2001\_CD14 CTL7\_CD14 CTL9\_CD14 CTL10\_CD14 PS2000\_CD14 PS2314\_CD14 CTL4\_CD14 CTL8\_CD14

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**CD14+**

PS2319\_CD14 CTL10\_CD4

CTL5\_CD4 PS2314\_CD4 CLT9\_CD4 CTL7\_CD4 PS2000\_CD4 PS2001\_CD4 CTL8\_CD4 PS2319\_CD4 CTL6\_CD4 PS2015\_CD4 PS2016\_CD4 CTL2\_CD4 CTL1\_CD4

**Cohort**

CTL3\_CD4 PS1011\_CD4 CTL4\_CD4 PS2014\_CD4 81\_2\_CD8 PS2319\_CD8 PS2314\_CD8 130\_2\_CD8 HV29\_CD8 HV30\_CD8 PS2000\_CD8 PS2001\_CD8 CTL1\_CD8 PS1011\_CD8 PS2014\_CD8 CTL6\_CD8 PS2016\_CD8 CTL3\_CD8 CTL5\_CD8 PS2015\_CD8 CTL2\_CD8

**CD4+**

**and CD8+**

**Cohort 1A Cohort 1B**

(b)

**Distance**

CTL4\_CD8 CTL1\_CD19

PS1011\_CD19 PS2001\_CD19 PS2015\_CD19 HV30\_CD19 CTL6\_CD19 PS2016\_CD19 CTL2\_CD19 PS2014\_CD19 HV29\_CD19 130\_2\_CD19 PS2000\_CD19 PS2314\_CD19 81\_2\_CD19 PS2319\_CD19 CTL4\_CD19 CTL3\_CD19 CTL5\_CD19

250

200

150

100

50

0

**CD19+**

Figure 4.9: PCA and sample distance heatmap with hierarchical clustering illustrating the sample variability based on the gene expression profiles for all 72 samples. a) The first and second PCs (x-axis and y-axis, respectively), each point represents a sample, and colour codes for cell type and the shape for condition. The proportion of variation explained by each principal component is indicated.b) Distance matrix clustering based on normalised read counts mapping to 20,493 Ensembl featured remaining after appropriate filtering. Annotation of the clustering using cohort identity is included.

mRNA and lncRNA differential expression

DGE analysis between 8 psoriasis patients and 10 healthy controls in CD14+ monocytes, CD4+ and CD8+ T cells and CD19+ B cells was performed using DESeq2 including cohort identity as a covariate. For each of the cell types a number of mRNAs were identified as differentially expressed at an FDR *<*0.05 or 0.01 (Table 4.7).

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

Table 4.7: Summary results from the DGE analysis between psoriasis patients and healthy controls in CD14+ monocytes, CD4+, CD8+ and CD19+ 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 fold change was applied in this analysis.

CD14+ monocytes and CD8+ showed the largest number of differentially expressed mRNAs between psoriasis patients and controls, with and the highest magnitude of fold change in expression (Fig. 4.10). [add 1-2 sentences stating examples of genes with biggest fold change/highest significance]

**CD14+ monocytes**

8

*ANKS1A*

*RGS1*

*HIST1H1D*



**Not sig. FDR<0.05 FDR<0.05 and**

**FC>1.5**



**PS1011 PS2014 PS2015**



829



**CD4+**

5

*CYTIP*

*SELM*

*RSPRY1*

***NFKB2 (u), PRKCB (d), RELB (u), TNFSF14 (d)***

681

239

75

**≠** 45

R=-0.981

pval=5.3x10-4

**PS1011 PS2014 PS2015**

7

4

6

5 3

**−log10 (FDR)**

4

**−log10 (FDR)**

2

3

2

1

1

0

−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0

**log2 (FC)**

(a)

0

−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0

**log2 (FC)**

(b)

**CD8+**

6

5

4

**−log10 (FDR)**

3

2

1

0

−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0

**log2 (FC)**

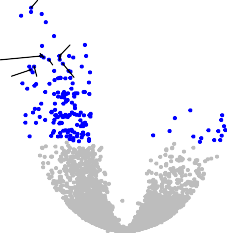
(c)

**CD19+**

5

|  |  |
| --- | --- |
|  | **Not sig.** |
|  | **FDR<0.05** |
|  | **FDR<0.05 and** |
|  | **FC>1.5** |
| *ZNF709 CCDC127* |  |
| *HLA−A* |  |
| *KBTBD7* | *NFKBIZ* |
| *ZC3H10 B3GNT2* | *NR4A1* |
| *CCDC28B* | *PER1* |
| *ZNF2* |  |
| *ZNF780A CLK1* |  |
| *TP53RK* |  |
| *FADD NAB2* |  |
| *GPR18 ZNF155 ZBTB5* |  |
|  | |

4



**−log10 (FDR)**

3

2

1

0

−1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0

**log2 (FC)**

(d)

Figure 4.10: 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+ monocytes, b) CD4+, c) CD8+ and d) CD19+ cells. For each gene, the log2(FC) represents the fold change in expression for that gene in the psoriasis group with reference to the healthy controls. Significant DEGs (FDR*<*0.05) in blue for FC *<*1.5 and red for FC *>*1.5. The volcano plots include mRNAs and lncRNAs species.

Enrichment of significant DEGs (FDR*<*0.05) across the four cell types and genes associated with psoriasis from reported GWAS was analysed using the NHGRI-EBI catalog [(https://ww](http://www.ebi.ac.uk/gwas))w[.ebi.ac.uk/gw](http://www.ebi.ac.uk/gwas))as) (Table 4.8). CD8+ was the cell type with the largest number of DEGs (7 hits) overlapping with putative GWAS genes, followed by CD14+ monocytes and CD4+ (3 hits each). Some of those genes were found in more than one cell type, including *NFKBIA*, *TNFAIP3* and

*NFKBIZ*, amongst others.

The role of lncRNAs in psoriasis circulating immune cells

In addition to protein coding genes, some of the DEGs identified were classified as lncRNAs. CD8+ and CD14+ monocytes were the two cell types

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

Table 4.8: Overlap between putative psoriasis GWAS genes and the reported significantly DEGs in CD14+ monocytes, CD4+, CD8+ and CD19+ cells. DEGs list based on FDR*<*0.05.

showing the largest number of dysregulated lncRNAs between psoriasis patients and controls (Table 4.7). [state some cell type-specific examples] Only one lncRNA, *RP11- 218M22.1* appeared to be dysregulated between psoriasis and healthy controls in all the cell types. The majority of the observed differentially expressed lncRNAs (FDR*<*0.05) were found to have a functional interacting partner based on

The lncRNAs differentially expressed between psoriasis and controls in CD14+ monocytes included the negative regulator of antiviral response



(*DYNLL1-AS1* or *NAV*), the HOXA transcript antisense RNA myeloid-specific 1 (*HOTAIRM1*) and the nuclear paraspeckle assembly transcript 1 (*NEAT1*). *DYNLL1-AS1* has been shown to affect the histone modifications of some critical IFN-stimulated genes (ISGs) including *IFITM3* and *MxA* leading to down- regulation of their expression (Ouyang et al. 2014). *DYNLL1-AS1* was down-regulated in CD14+ monocytes from psoriasis patients when compared to controls but no up-regulation of *IFITM3* and *MxA* was found. Conversely, *HOTAIRM1* was up-regulated in CD14+ monocytes from psoriasis patients (Figure 4.11 a). The experimentally validated target for *HOTAIRM1* reported by NPInter database was the RNA helicase and ATPase *UPF1* (Hao et al. 2016), which was found to be down-regulated in CD14+ monocytes in psoriasis versus control samples in this data (Figure 4.11 b). *NEAT1* was also up-regulated in psoriasis patients compared to controls and had also been found to be up-regulated in a study in SLE CD14+ monocytes (Zhang et al. 2016).

11.4

11.3

***HOTAIRM1* (log2 normalised counts)**

8.5

11.2

***UPF1* (log2 counts)**

8.0

11.1

11.0

7.5

10.9

Controls Psoriasis

**Condition**

(a)

Controls Psoriasis

**Condition**

(b)

Figure 4.11: RNA-seq expression levels of the lncRNA *HOTAIRM1* and its experimentally validated target *UPF1* in psoriasis and healthy controls CD14+ monocytes. Expression is illustrated as the log2 of the normalised read counts mapping to a) the lncRNA *HOTAIRM1* and b) *UPF1*, which has been experimentally identified as one of the genes regulated by this lncRNA according to NPInter database.

innate immunity, inflammatory response and antiviral pathway

Pathway enrichment analysis for the DEGs

To investigate the biological role of the significantly modulated genes, pathway enrichment analysis was performed for each cell type using DEG with FDR*<*0.05 and no FC cut off. Biologically relevant pathways were significantly enriched (FDR*<*0.01) for CD14+ monocytes and CD8+ cells (Table ?? and A.3). In CD19+ cells, only one pathway (generic transcription) was significantly enriched while in CD4+ cells no enrichment was seen for any pathways.

Cell type Pathways

MAPK signalling

IL-12 mediated signaling events Th-1 and Th-2 cell differentiation

CD14+ monocytes Th-17 cell differentiation

TCR signalling

Platelet-derived growth factor (PDGF-*β*) signalling Forkhead box O (FoxO) signalling

CD8+ Osteoclast differentiation MAPK signalling TNF signalling

IL-12 mediated signalling events NF-*κ*B signalling

Chemokine signalling

Table 4.10: Pathways enriched for DEGs between psoriasis patients and healthy controls in CD14+ monocytes and CD8+ 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.

Two of the significant enriched pathways, MAPK signalling and IL-12 mediated signalling, were found to be enriched in both CD14+ monocytes and CD8+ cells (Table ??). MAPK signaling was of particular interest as DEGs contributing to the enrichment of this pathway in both cell types included MAPK gene members *MAP3K4* and *MAPK14*, both down-regulated in psoriasis vs controls. MAP3K4 is a member of the MAPKKK family, whose expression is down- regulated in LPS-stimulated PBMCs from Crohn's patients leading to a relative immune deficiency in TLR-mediated cytokine production. Moreover, DGE of

members of the dual-specificity phosphatases (DUSP) family, involved in fine- tuning the immune response (Qian et al. 2009), contributed to the enrichment of the MAPK pathway in CD14+ monocytes and CD8+ cells. *DUSP10* was down-regulated in psoriasis CD14+ monocytes and knock- out in mice has been associated with enhanced inflammation (Qian et al. 2009). Conversely, *DUSP4* was up-regulated in psoriasis CD8+ cells vs healthy controls and has been demonstrated to have a pro-inflammatory role in a sepsis mice model (Cornell et al. 2010).

Regarding enrichment of the IL-12 signalling, CD14+ monocytes from

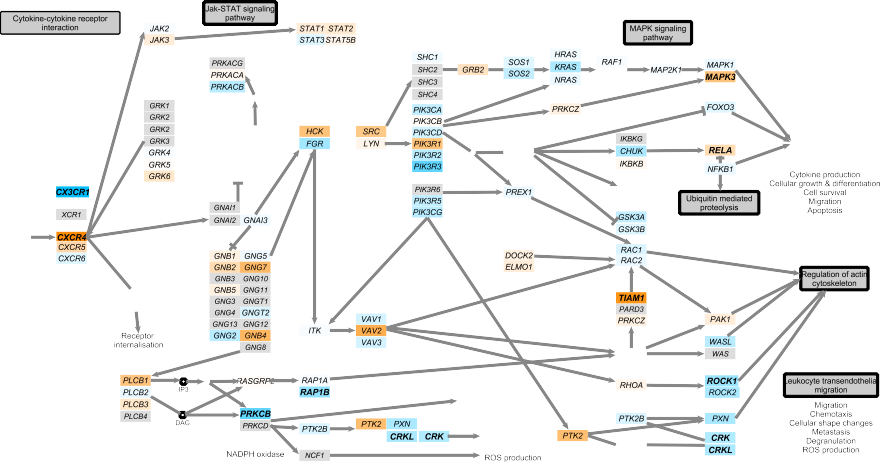
psoriasis showed down-regulation of *STAT4* and *STAT5A* in patients compared to controls. Neither *STAT4* and *STAT5A* were differentially expressed in CD8+ cells. By contrast, *IFNG* expression in psoriasis patients was lower than in healthy controls in CD8+ cells but showed no difference in CD14+ monocytes; and *IL2RA* was up-regulated in CD8+ T cells from psoriasis patients, which may enhance formation of the IL2-R*α* and the signalling by this cytokine involved in effector and regulatory T cell differentiation (Malek and Immunity 2010).

The platelet-derived growth factor (PDGF-*β*) signalling pathway was only enriched in CD14+ monocytes (Table 4.10). Within this pathway the *SLA* gene was down-regulated in psoriasis patients. A *SLA* knock-out mouse model has shown impaired IL-12 and TNF-*α* and failure of T cell stimulation by GM-CSG treated bone marrow-derived DCs (Liontos et al. 2011).

A number of very relevant inflammatory pathways in psoriasis were enriched only in CD8+ cells. These included TNF, NF-*κ*B and chemokine signalling (Table 4.10). Up- and down- regulation of specific chemokines was seen in CD8+ cells from psoriasis patients when compared to healthy controls (Figure 4.12 b), including upregulation of *CCR10*, the receptor for the chemotactic skin-associated chemokine CCL27. Other up-regulated chemokine receptors included *CXCR4* gene, the receptor for the chemokine SDF-1, highly expressed in skin (Zgraggen et al. 2014). Of note, none of the genes coding for well-known psoriasis drug target genes, including TNF-*α*, IL-17 and IL-6,



(a)



(b)

Figure 4.12: Mapping of the DEGs identified in CD8+ cells between psoriasis patients and healthy controls onto the TNF-*α* and the chemokine signalling pathways. The a) TNF-*α* 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 log2FC expression between psoriasis patients and healthy controls CD8+ cells isolated from PB. Significant DEGs (FDR*<*0.05) are highlighted in bold. In a), members of the TNF-*α* pathway shared with the NF-*κ*B are highlighted with a green box. Additional members of the NF-*κ*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+ DEGs (FDR*<*0.05).

were up-regulated in any of the four cells types from psoriasis patients compared to healthy controls.

## RNA-seq in epidermis from psoriasis patients

Data processing and quality control

For the three paired uninvolved-lesional samples (Table 4.4), all had a mapping rate greater than 80%. The number of reads after filtering mapped to Ensembl genes ranged between 29.5 and 33.2 million . The mapping rate and final number of reads mapping to genes was greater in the lesional samples compared to the controls (Fig 4.13).

PCA using the normalised number of reads mapping to the genes after filtering revealed separation of the lesional samples from the uninvolved by the first PC, which explained 37% of the variance (Figure 4.13 c). Overall, PCA revealed substantial variation between the lesional and uninvolved samples and biological variability across individuals, for which the paired design in the DGE analysis accounted.

Summary of the DGE results

DGE analysis revealed a total of 1,227 (FDR*<*0.05) and 702 (FDR*<*0.01) genes differentially expressed between 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 (Figure 4.14) and 46 were annotated as lncRNAs (Table ??). The magnitude of change in gene

90 35

33

**Total reads after filtering (million)**

89

**Mapping rate (%)**

31

29

88

27

87

uninvolved lesional

**Condition**

(a)

25

uninvolved lesional

**Condition**

(b)

20

0

**PC2 (30%)**

−20

**Uninvolved Lesional**

−40

−20 0 20 40

**PC1(37%)**

(c)

Figure 4.13: 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 Ensmbl list of mRNAs and lncRNAs detected in this study. Dots colour corresponds to condition (lesional or uninvolved) and shape refers to the patient ID.

expression between lesional and uninvolved skin were notably larger when compared to the changes in expression from analysis in peripheral blood immune cells, with 874 out of 1,227 genes showing FC¿1.5.

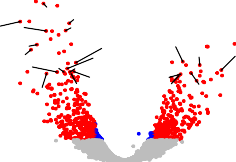
30

|  |  |
| --- | --- |
|  | **Not sig.** |
|  | **FDR<0.05** |
|  | **FDR<0.05 and** |
| ADRB2 | **FC>1.5** |
| SLC8A1 |  |
| GJB4 |  |
| PAMR1 |  |
| F3 |  |
| LSP1 |  |
| NR4A1 ID4 |  |
| RP11−215G15.5 IL34 | ZC3H12A LCE3A |
|  |  |
| PLXNA4 GAN | AQP5 HRH2 |
| RP11−217B1.2 SERPINA12 | CD36 DSC2 |
| RAB3B RAI14 KRT77 PSAPL1 | AKR1B10 |
| RORC POSTN | ANGPTL4 LTF |
| ADAMTSL3 SYNE1 C1orf68 CLDN1 | SAMD9  GPR68 |
| CACNA2D1 TMEM99 ELMOD1 PTPN21 | S100A7 OLFM1 IGFL1 |
| C14orf132 TPPP3 APOE | S100A2 TNFRSF21 |
| BTC SLC46A2 LCE1B | FSCN1 KRT9 SLC16A9 |
| CRY2 | CCNB1 CARHSP1 |
|  | |

25

20

15



**−log10 (FDR)**

10

5

0

−4 −3 −2 −1 0 1 2 3 4

**log2 (FC)**

Figure 4.14: Magnitude and significance of the gene expression changes between matched lesional and uninvolved epidermal biopsies from three psoriasis patients. The volcano plot represents for each gene the significance (-log10*FDR*) of the log2(FC) in expression for that gene in lesional skin vs uninvolved skin. Significant DEGs (FDR*<*0.05) in blue for FC *<*1.5 and red for FC *>*1.5. The volcano plot includes mRNAs and lncRNAs species.

Amongst the DEGs between the uninvolved and lesional skin, five genes (FDR*<*0.05) overlapped with putative GWAS genes (Table ??). *IFIH1*, *NOS2*, *LCE3D* and *STAT3* were up-regulated in lesional compared to uninvolved skin, whereas *TNFAIP3* was down-regulated.

|  |  |  |  |
| --- | --- | --- | --- |
| FDR threshold | mRNA | lncRNA | Overlap with GWAS genes |
|  |  |  |  |
| 0.05 | 1181 | 46 | up(*IFIH1*, *NOS2*, *STAT3*, |
| 0.01 | 677 | 25 | *LCE3D*), down(*TNFAIP3*)  *NOS2*, *STAT3*,*TNFAIP3*,*LCE3D* |

Table 4.11: 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.

Overall comparison with other skin transcriptomic studies

As detailed in Chapter 2, the approach to study DGE in skin is different from most previously published studies using whole punch biopsies to compare lesional and uninvolved skin from psoriasis patients. During the course of this project a study was published by Tervaniemi and colleagues that also aimed to characterise the transcriptional profiles of the epidermis from psoriasis patients lesional and uninvolved skin (Tervaniemi et al. 2016). In order to explore the similarities between the two studies, a comparison for the DEGs identified between lesional and uninvolved matched samples 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) and showing overall a larger number of differentially expressed genes between the two types of biopsies compared this study. The number of genes up-regulated in lesional epidermis compared to uninvolved (2,330) was larger than the number of down-regulated targets (261), contrasting to the in-house results where similar numbers of up- and down- regulated genes were found (Table ?? and Figure 4.15 bottom panel). Regarding overlap, a total of 359 out of the 1,227 DEGs (29.25%) identified by the in-house study were shared with the Tervaniemi results, of which 239 and 75 were up- and down-regulated, respectively. Amongst the up-regulated genes in both studies TFs such as *STAT1*, genes from the *S100* family (e.g *S100A9* and *S100A12*) and genes nearby psoriasis GWAS loci such as *STAT3* and *IFIH1*. The direction of change in 45 out of the 359 shared genes appeared to be opposite across the two datasets. For example, *SERPINB2* gene, a serine protease inhibitor of the serpin superfamily, presented down-regulation in the in-house data and up-regulation in the Tervaniemi results. Interestingly, a 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 (Schroder et al. 2016).

In house Tsoi

534

295

2,540

212

147 294

1,936

Tervaniemi

**Not sig. FDR<0.05 FDR<0.05 and**

**FC>1.5**

*KBTBD7*

*RAP2A HLA−A FAM178A*

*USP24*

*DNAH10 KBTBD6 NRBF2 PGBD2*

Figure 4.15: Overlap of the significantly differentially expressed genes between lesional and uninvolved epidermal sheets, split epidermis and full-thickness skin biopsies. The central venn diagram illustrated the DEGs overlapping between this study (in house), Tervaniemi *et al.*, 2016 split epidermis biopsies and Tsoi *et al.*,2015 full thickness skin biopsies. Overlap is considered regardless the direction of the change. Two additional venn diagrams provide more detail about the total overlap and directionality in the change of gene expression between the in-house data and the Tsoi *et al.* (top) or Tervaniemi *et al.* (bottom).

In addition to the Tervaniemi study, our results were further contrasted to one of the most recent comprehensive RNA-seq studies comparing lesional and uninvolved full thickness skin biopsies from psoriasis patients (Tsoi et al. 2015). Out of the 3,725 DEGs reported by Tsoi and colleagues, 507 genes were shared between the two studies (41% of the in-house DEGs) and 24 corresponded to dysregulated lncRNAs. Out of the 507 commonly dysregulated genes in

the two datasets, 272 were up-regulated, 228 down-regulated and 7 showed opposite direction of change (Figure 4.15 top panel). Overlap across the three studies only identified 212 DEGs shared by the three datasets.

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 in the three psoriasis patients from this study. Out of the 46 differentially regulated lncRNAs, 37 had a functional experimental partner functionally validated according to NPInter database (Hao et al. 2016). An interesting example was *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 keratinocyte differentiation (Li et al. 2017). Nevertheless, *DSG1* did not appear as one of the DEGs between lesional and uninvolved 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 been noted 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; Tsoi et al. 2015), and was also shown to have increased expression when comparing lesional skin versus healthy biopsies (Li2014). One of the predicted miR-146a targets by Target Rank software in a study conducted by Jazdzewski and colleagues revealed *NFAT5*, also down- regulated in lesional skin compared to uninvolved in the in-house data, as the

11*th* most confidently predicted target (Jazdzewski et al. 2009). Interestingly, a

negative correlation (R=-0.981, pval=5.3x10−4) between the normalised counts of the two genes was found in the three lesional-uninvolved paired samples (Figure 4.16 a).

12.5 12.5

12.0 12.0

11.5 11.5

***NFAT5* (log2 counts)**

***HIF1A* (log2 counts)**

11.0 11.0

10.5 10.5

**Uninvolved Lesional**

**Uninvolved Lesional**

10.0

3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0

***MIR146A* (log2 counts)**

(a)

10.0

0 1 2 3 4 5 6 7 8

***MIR31HG* (log2 counts)**

(b)

Figure 4.16: Correlation in gene expression between two dyregulated miRs in lesional skin and their putative target genes. Plots showing the correlation in log2 normalised read counts for a) *MIR146A* and its putative target genes *NFAT5* and b) *MIR31HG* and its putative target genes *HIF-1A*. Pearson correlation values (R) and significance (pval) are included. Each of the dots represents one samples, where colour represents condition (lesional or uninvolved) and shape corresponds to the patient ID.

Pathway enrichment analysis

In order to better understand the functional role of the DEGs (FDR*<*0.05) between lesional and uninvolved epidermis from psoriasis patient skin biopsies, pathways enrichment analysis was performed. Many pathways were significantly enriched (FDR*<*0.005) for DEGs found in our analysis (Table 4.12 and A.4).

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. 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. In this data up-regulation of *HIF1A*, *VEGFA*, *ENO1* and the GWAS gene *NOS2*, amongst others, contributed to the enrichment of this pathway (Figure 4.18).

Lesional versus uninvolved epidermis enriched pathways

IFN-*α*/*β*/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 4.12: 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.





Figure 4.17: Mapping of the DEGs between lesional and uninvolved epidermis from psoriasis patients onto the HIF-I 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 log2FC 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).

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-*κ*B and MAPK activation, was enriched with 23 significantly DEGs (Figure 4.18 in orange and bold) including *NOD2*, *CARD6* and *IFI16* (highly up-regulated) and *TNFAIP3* and *BCL-2 (down-regulated)* (Figure 4.18 in blue and bold).

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

179

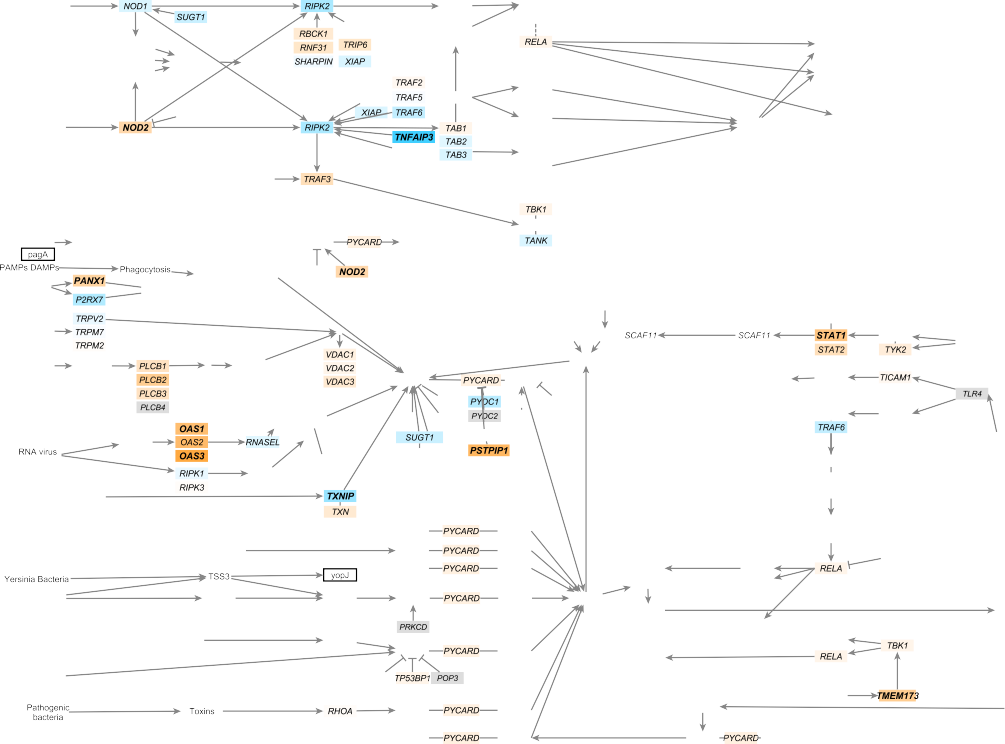
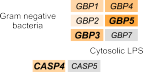
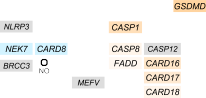
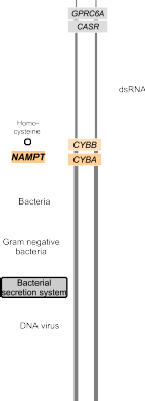
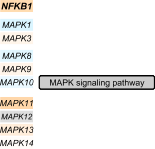
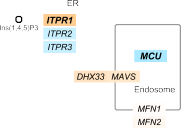


Figure 4.18: Mapping of the DEGs between lesional and uninvolved epidermis from psoriasis patients onto the NOD-like signalling pathway. This pathway was sourced from KEGG, manually curated in a way that all member genes are maximised visually and then automatically color-coded by the log2FC expression 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).

In addition to the NOD-I signalling, IL-17 signalling was another enriched pathway well known to be relevant in the development of psoriasis. Enrichment of the IL-17 signalling pathway in our data is driven by up-regulation of the S100 protein family (*S100A7*, *S100A8* and *S100A9*) and chemokines such as *CCL20*, which binds the CCR6 receptor and is involved in DCs and T cell chemotaxis. *IL-17RE* which together with IL-17RA forms the receptor for IL-17C was down- regulated in lesional skin. 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. This pathway included up-regulation of the PPAR receptor *δ* (*PPARD*), stearoyl-CoA desaturases such as *SCD* and *SCD5* involved in fatty acid synthesis and *CD36* which mediates fatty acid transport, also dysregulated in the Tsoi and/or Tervaniemi studies.

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

In order to investigate commonalities and differences in psoriasis gene expression at the affected tissue (skin) and the systemic level (circulating immune cells), overlap between the lists of DEGs was performed. Only modest overlap was found between differentially expressed genes in lesional skin compared to uninvolved and the DEGs identified in circulating immune cells, with CD14+ monocytes and CD8+ cells showing the greatest overlap although many of these genes did not show a consistent direction of change in differential expression (Table 4.13). Examples included *TNFAIP3*, which was up-regulated in psoriasis CD4+ and CD8+ cells

compared to controls and down-regulated in lesional epidermis when compared to uninvolved. Early growth response genes were up-regulated in both CD14+ monocytes and CD4+ cells, and in lesional skin but the magnitude fo change was much greater in the latter (*EGR2* and *EGR3* log2FC -0.74 and -0.53 respectively).

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

Table 4.13: Overlap between the DEGs in the four circulating immune cell types (psoriasis patients versus controls) and the DEGs in psoriasis patients skin biopsies (lesional versus uninvolved). DEGs based on 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+ and CD8+ DEGs were mostly immune-related pathways, including TCR , IL-12 , TNF and NF-*κ*B signalling. In skin, the DEGs were not only enriched in immune-related pathways but also for pathways involved in metabolism,

oxidative stress and cell cycle, and the genes contributing to the enrichment of immune-related pathways had a more pronounced pro-inflammatory signature, consistent with the skin being a site of more active inflammation compared to circulating immune cells in psoriasis.

## Integration of chromatin accessibility and expression data for peripheral blood immune cells in psoriasis

The characterisation of the chromatin accessibility landscape and the transcriptome in circulating immune cells from psoriasis patients has revealed a greater effect of disease status on gene expression than chromatin accessibility. To investigate whether an integrated approach combining evidence of differtnail accessibility and expression, overlap between DEGs and the genes proximal to DARs (≤5Kb) was investigated. Overlap was only

found in CD8+ cells, where 6 out of the 53 DARs were annotated by proximity to

an RNA-seq DEG in the same cell type (*ARL4A*, *ASCL2*, *ENTPD1*, *TIAM1*, *TRAT1*

and *ZNF276*).

For example, T Cell lymphoma invasion and metastasis 1 (*TIAM1*), activates IL-17 expression and T cell transendothelial migration during inflammation (Kurdi et al. 2016; Grard et al. 2009). This gene showed an increased expression (log2FC 0.44) in psoriasis patients CD8+ T cells (Figure 4.19 left) and greater chromatin accessibility compared to healthy controls (log2FC 0.41) in a region located at an intron of the *TIAM1* gene and annotated as an active enhancer according to the Roadmap chromatin segmentation data in this cell type (Figure 4.19 right). Common SNPs within this peak did not appear to be an eQTL regulating expression of any gene in CD8+ cells (Kasela et al. 2017) and chromatin conformation data did not reveal interaction of this particular region with the *TIAM1* promoter (Javiere2016), at least in unstimulated conditions, complicating

the establishment of a mechanistic connection between chromatin accessibility and gene expression.

Another two relevant genes in the immune response for which ATAC and RNA-seq showed overlap were the ectonucleoside triphosphate 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 TCR signalling (Antonioli et al. 2013; Valk et al. 2006). Both genes showed up-regulated expression and increased chromatin

accessibility in psoriasis patients CD8+ cells compared to healthy controls.

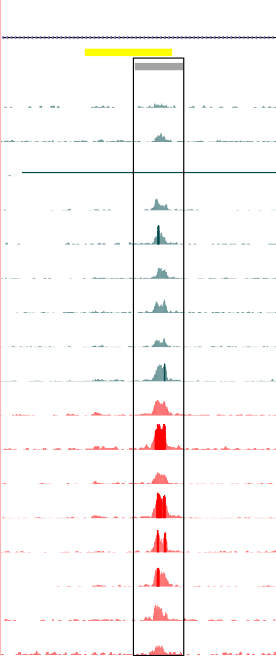
Scale chr21:

**UCSC genes**

*TIAM1*

2 kb hg19

32,725,000 32,730,000



**ChromHMM RoadMap**

**Peaks ATAC-seq**

69 \_

**CTL1**

**CTL2**

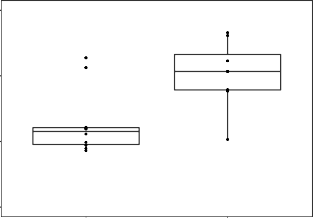
0 \_

69 \_

0 \_

69 \_

**RNA-seq expression**

11

**TIAM1 (log2 counts)**

10

9

8

Controls Psoriasis

**Condition**

**CTL3**

**CTL4**

**CTL6**

**CTL7**

**CTL8**

**CTL9**

**CTL10**

**PS2014**

**PS2015**

**PS2016**

**PS2000**

**PS2001**

**PS2314**

**PS2319**

**PS1011**

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

69 \_

0 \_

Figure 4.19: Differential gene expression and chromatin accessibility landscape for *TIAM1* gene in CD8+ cells. Boxplot on the left represents RNA-seq log2 normalised counts for the *TIAM1* gene in psoriasis and healthy controls CD8+ cells. The right panel is the UCSC Genome Browser view illustrating the normalised ATAC read density (y- axis) at an intron of the *TIAM1* gene (x-axis) in CD8+ cells. This region was identified as more accessible in psoriasis patients compared to healthy controls. Tracks are colour- coded by condition and assay: control(CTL)=dark turquoise and psoriasis (PS)=red. The Epigenome Roadmap chromatin segmentation map for CD8+ cells is included.

* + 1. Allele-specific differences in chromatin accessibility at a GWAS locus 2p15

The chr2p15 psoriasis risk locus (Tsoi *et al.*,2012) is shared with other chronic inflammatory diseases including AS and Crohn’s disease (Ellinghaus et al. 2016). Fine-mapping using summary statistics from the psoriasis GAPC Immunochip cohort failed to successfully fine-map this locus (log10BF¡3). The lead SNP reported in both the AS GWAS association analysis and multi-disease meta-analysis from Ellinghaus and colleagues was rs4672505 (risk allele rs4672505\_A) (Ellinghaus et al. 2016).

rs4672505 overlaps a CD8+ T cell specific ATAC region (Figure 4.20 a) that was not differentially accessible between psoriasis patients and controls but did show marked variability between individuals (Figure 4.20 b), with some individuals (PS2314 and CTL1) demonstrating no ATAC signal at this location. Integration with publicly available ENCODE and Epigenome Roadmap DHS data confirmed accessible chromatin at this site in Th-1, Th-2 and Th-17 cells and CD8+ T cells respectively (Figure 4.20 a). together with evidence of RUNX3 transcription factor binding (a psoriasis and AS GWAS associated gene involved in CD8+ cell differentiation (Wong et al. 2011)). In addition *in silico* TFBS prediction using PROMO (Messeguer et al. 2002) and ENCODE genomic DNase-I footprint in GM128778 predicted STAT1 binding at rs4672505. Altogether, integration of ATAC and publicly available epigenetic data indicated that rs4672505 was the most likely variant, amongst the three fine-mapped SNPs included in the 95% credible set, to have a functional role explaining the association of chr2p15 with psoriasis risk.

Scale



2 kb hg19

chr2: 62,560,000 62,565,000

**Fine-mapped SNPs** rs6759298 rs4672505

**ChromHMM RoadMap**

**Txn Factor ChIP Roadmap CD8** 136 \_

**DNase-I**

**footprint** 1 \_

**Peaks ATAC CD8** 56 \_

Scale chr2:

**Fine-mapped SNPs**

**Peaks Th-1**

14 \_

**ENCODE DHS Th-1**

1 \_

**Peaks Th-2**

12 \_

**ENCODE DHS Th-2**

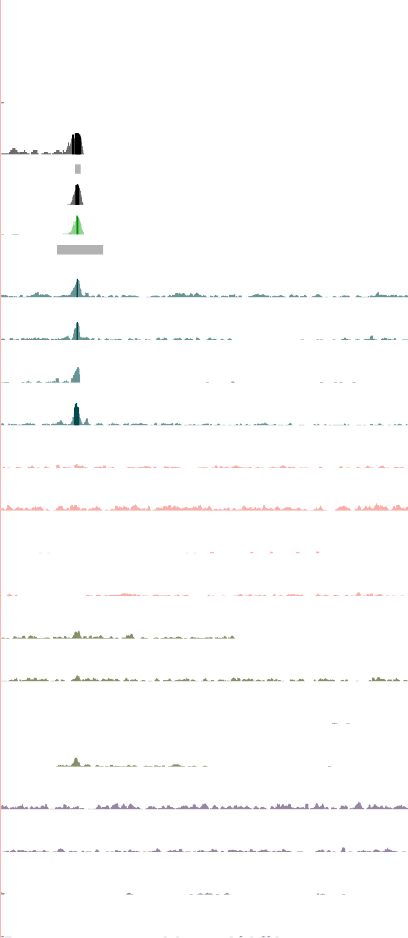
5 kb hg19

62,560,000 62,565,000

rs6759298 rs4672505 rs6759003

**CTL1**

**CTL2**



**CTL3**

**CTL4**

**(GA)**

**(GG)**

**(GA)**

**(GG)**

0 \_

56 \_

0 \_

56 \_

0 \_

56 \_

0 \_

56 \_

**Peaks Th-17**

1 \_

80 \_

**CTL6**

**(GG)**

0 \_

56 \_

**ENCODE DHS Th-17**

1 \_

**Roadmap CD56 DNase-I**

**footprint Peaks ATAC CD8**

56 \_

**CTL2 CD8**

0 \_

56 \_

**CTL7**

**CTL8**

**CTL9**

**(GA)**

**(GG)**

**(GA)**

0 \_

56 \_

0 \_

56 \_

**CD8+**

**CTL3 CD8**

**CTL6 CD8**

**CTL7 CD8**

0 \_

56 \_

0 \_

56 \_

**CTL10 (GA)**

0 \_

56 \_

0 \_

56 \_

0 \_

56 \_

**CTL2 CD14**

0 \_

56 \_

**CTL3 CD14**

**PS1011 (GA)**

**PS2014 (GA)**

0 \_

56 \_

0 \_

**CD14+**

**monocytes**

**CTL6 CD14**

**CTL7 CD14**

0 \_

56 \_

0 \_

56 \_

**PS2015 (GA)**

**PS2016 (GA)**

56 \_

0 \_

56 \_

0 \_

56 \_

**CTL2 CD4**

0 \_

56 \_

**CTL3 CD4**

**PS2000 (GG)**

0 \_

56 \_

0 \_

56 \_

**CD4+**

0 \_

56 \_

**CTL6 CD4**

0 \_

56 \_

**PS2001 (GA)**

0 \_

56 \_

**CTL7 CD4**

0 \_ 56 \_

**CTL2 CD19**

0 \_

56 \_

**CTL3 CD19**

0 \_

**PS2314 (AA)**

**PS2319 (GA)**

0 \_

56 \_

0 \_

19 \_

**CD19+**

**CTL6 CD19**

**CTL7 CD19**

56 \_

0 \_

56 \_

0 \_

(a)

**H3K27ac CTL7**

**H3K27ac PS2314**

0 \_

19 \_

0 \_

(b)

Figure 4.20: Epigenetic landscape at the location of the SNPs in the chr2p15 GWAS psoriasis locus. a) UCSC Genome Browser view illustrating normalised read density for in-house ATAC and a number of other publicly available epigenetic data (DHS, DNAse-I footprint, chromatin segmentation map) (y-axis) spanning the three SNPs (rs6759298, rs4672505 and rs6759003) (x-axis) from the 95% credible set obtained in the fine-mapping analysis of the chr2p15 GWAS association in AS. Representative ATAC data from the same four controls in the cohort and the four cell types included in this study are shown. b) UCSC Genome Browser view illustrating the normalised read density for CD8+ ATAC (x-axis) generated in psoriasis patients and healthy controls, in-house H3K27ac ChIPm, ENCODE TF ChIP-seq and DNase-I fooprint (y-axis) at the location of the SNP rs4672505 (y-axis). For each of the patients and controls of the cohort the Sanger sequencing genotype of rs4672505 is included.

The genotype of each individual for rs4672505 was determined using Sanger sequencing. Amongst the eighteen samples (ten controls and eight psoriasis patients), one (PS2314) was homozygous for the risk allele (A, MAF=0.43), eleven were heterozygous and six were homozygous for the protective allele (G) (Figure 4.20 b). Interestingly, PS2314, the only homozygous individual for the risk allele, showed complete absence of the peak at rs4672505.

To further investigate the role of rs4672505 genotype in the variability of chromatin accessibility across individuals, the normalised read counts at the ATAC peak (chr2:62,559,749-62,561,442) were used as a dependent variable in linear model analysis based on rs4672505 genotype, using batch as a covariate. Significant negative correlation (pval=0.035) was found, suggesting allele-dependent chromatin accessibility (Figure 4.21 a). Furthermore, allelic imbalance for the ATAC reads at rs4672505 position was investigated on those individuals identified as heterozygous by Sanger sequencing and for which 50% of the ATAC reads were expected to map to each of the alleles. This analysis demonstrated a larger percentage of ATAC reads (greater than the expected 50%) preferentially tagging the protective allele G (Figure 4.21 b). This finding was not driven by mapping bias, since A was the reference allele in the hg19 build

100

100

75

**Normalised ATAC counts at rs4672505**

**rs4672505 (% of allele G)**

80

50

60

25

0 (GG)

1 (GA)

**rs4672505 genotype**

(a)

40

2 (AA)

ATAC

reads

**Method**

(b)

Sanger

Figure 4.21: rs4672505 genotype and chromatin accessibility at chr2:62,559,749-62,561,442 in CD8+ cells. a) Boxplot illustrating the effect of the rs4672505 genotype on chromatin accessibility at the chr2:62,559,749-62,561,442 ATAC peak. Log2 normalised ATAC counts adjusted for batch effect, also included as a covariate for the linear model, are plotted for each sample against the number of copies of the minor allele (G=0, AG=1, AA=2). b) Representation of the percentage of ATAC reads overlapping rs4672505 and mapping to the major allele (G) in comparison to the Sanger genotype results for the eleven heterozygous individuals at this SNP.

used to map the ATAC data. Overall, these results showed evidence of greater chromatin accessibility in presence of the fine-mapped protective allele rs4672505(G) at the chr2p15 locus.

A major challenge with intergenic GWAs signals is the difficulty in determining the specific gene they may be modulating, for example through differential enhancer activity affecting gene expression. rs4672505 is located 140Kb downstream of *B3GNT2* and 150Kb upstream of *TMEM1*. Publicly available promoter capture data from Javierre *et al.* 2016 in CD8+ revealed a genome- wide significant interaction (CHiCAGO score=7.67) between a region containing rs4672505 and the promoter of the *B3GNT2* gene (Figure 4.22 a). This interaction was not found in any of the additional sixteen human primary hematopoietic cell included in the study. No upstream interaction with *TMEM1* promoter was identified. Investigation of the publicly available T cell eQTL dataset from Kasela *et al.* and Raj *et al.* did not show a

**CTL2 CD8**

62410000 62420000 62430000 62440000 62450000 62460000 62470000 62480000 62490000 62500000 62510000 62520000 62530000 62540000 62550000 62560000 62570000

**CTL3 CD8**

**CTL6 CD8**

**CTL7 CD8**

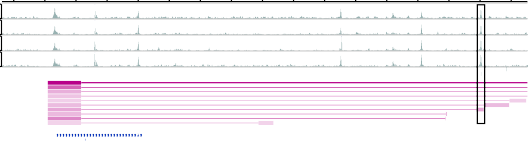
**Fine Mapping SNPs**

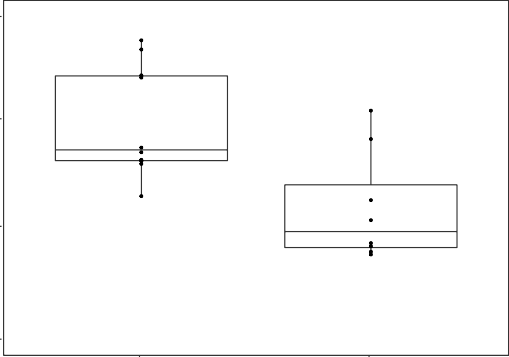
**Promoter capture HiC**

**rs6759003 rs6759298 rs4672505**

**Genes**

*B3GNT2 MIR5192*

(a)

11.0

10.5

***B3GNT2* (log2 normalised counts)**

10.0

9.5

Controls Psoriasis

**Condition**

(b)

Figure 4.22: Potential role of rs4672505 in regulating *B3GNT2* gene expression. a) WASHU Genome Browser track showing in CD8+ cells normalised ATAC read density in four of the healthy controls, the location of the three SNPs of the 95% credible set and the promoter capture HiC data depicting the regions interacting with the bait at the *B3GNT2* promoter. b) Boxplot illustrating the *B3GNT2* log2 normalised RNA-seq counts adjusted for batch effect in the psoriasis and healthy control groups.

significant eQTL for this SNPs or SNPs in high LD (r2*>*0.8) either in CD8+ or CD4+ (Raj et al. 2014; Kasela et al. 2017). Similarly, no eQTL effect of rs4672505 was found in unstimulated or stimulated CD14+ monocytes (Fairfax et al. 2014). However, a whole blood eQTL study from Jansen and colleagues revealed a significant *cis*-eQTL (FDR=1.34x10−5) with moderate effect size (*β*=-0.16) for the minor allele (Jansen et al. 2017). In terms of gene expression based on disease state, significant (FDR*<*0.05) down-regulation of *B3GNT2* was observed in psoriasis patients when compared to controls in CD8+ cells (-log2FC=-0.317) (Figure 4.22 b). These data provide evidence of allele-specific chromatin accessibility at chr2p15 in CD8+ T cells involving a GWAS SNP, with chromatin conformation capture and gene expression data suggesting that this modulates at a distance expression of *B3GNT2*, providing insights into the possible mechanistic basis of the disease association seen in psoriasis and other inflammatory diseases including AS.

# Discussion

## Chromatin accessibility and H3K27ac landscape in psoriasis immune cells

Comparison of chromatin accessibility and H3K27ac histone modifications has revealed a small number of differential regions between patients and controls in the four cells types under study. For both epigenetic features, CD14+ monocytes and CD8+ cells had the largest number of discrete changes. In ATAC, greater accessibility in CD8+ cells from patients compared to controls was found at two regions proximal to *IL7R* and *TNFSF11*, respectively, that also overlap FANTOM eRNA in the same cell type. Both genes are well known for having a pro-inflammatory effect and be involved in chronic inflammatory diseases. For example, *TNFSF11* is downstream of the lead SNP for a CD risk locus, and its protein product RANKL was found to be overexpressed in epidermis from

psoriasis patients, highlighting the role of this gene in the pathophysiology of psoriasis (Toberer et al. 2011).

Integration of the ATAC and H3K27ac ChIPm differential analysis only found one overlapping region at an intron of the *DTD1* gene, which participates in initiation of DNA replication and is associated with aspirin-intolerance in asthmatics (Pasaje et al. 2011). However, evidence of *DTD1* involvement in chronic inflammation has yet not been reported. The lack of overlap between DARs and differentially H3K27ac modified regions might be expected given that chromatin accessibility is driven by the interaction between a number of histone modifications, TFs, and structural proteins, such as CTCF.

The results in this chapter suggest that disease status does not involve global differences in chromatin accessibility and H3K27ac between patients and controls in the studied circulating immune cells. Recent similar studies performing ATAC in B cells from SLE patients have revealed larger differences in the chromatin accessibility landscape between patients and controls (Scharer et al. 2016) while H3K27ac mapping in mCD4+ cells isolated from

juvenile idiopathic arthritis synovial fluid found approximately one thousand differential enhancers when compared to healthy control circulating cells (Peeters et al. 2015). Conversely only small differences were found when comparing mCD4+ from peripheral blood of patients and controls, highlighting the specificity of the disease signature at the site of inflammation and importance of studying the most disease relevant cell and tissue types. Moreover it may be necessary to study changes in chromatin accessibility in the context of

genotype and under exogenous inflammatory stimuli that may manifest those differences (Calderon2018; Alasoo et al. 2018).

## Dysregulation of gene expression in psoriasis circulating immune cells

Comparison of gene expression between psoriasis and healthy controls in a cell type specific manner identified larger numbers of DEGs compared to DARs or differential H3K27ac modifications. As for ATAC and ChIPm, CD14+ monocytes and CD8+ cells showed the largest number of transcriptomic changes in disease. This may suggest greater relevance of these two cell types in the systemic footprint of psoriasis. The more dysregulated gene expression in CD8+ compared to CD4+ may suggest that, as in skin, CD8+ are the main effector cells upon induced-activation by CD4+ cells (Nickoloff and Wrone-Smith 1999). The importance of monocytes/macrophages in psoriasis has also been demonstrated by their presence in psoriatic skin where TNF-*α* production contributes towards maintenance of inflammation (Nickoloff2000; Wang et al. 2006).

The cell type specific analysis conducted in my thesis identified significant enrichment of relevant biological processes, including MAPK and IL-12 signalling, in CD14+ monocytes and CD8+ T cells. Interestingly, some of the well-known pro-inflammatory genes contributing to the enrichment of these pathways were down-regulated in psoriasis compared to controls.

For example, *MAP3K4* down-regulation in LPS stimulated PBMCs has been identified as an immune-suppressive feature in CD leading to reduced expression of the cytokine IL-1*α*. In the IL-12 signalling pathway, leading to T cell proliferation and IFN-*γ* production through activation of TFs from the STAT family, CD14+ cells showed down-regulation of *STAT4* and *STAT5A* in psoriasis versus controls. Other member of the STAT family, such as *STAT2*, was found to be down-regulated in psoriasis PBMCs and in AS monocyte- derived macrophages when compared to controls (Coda et al. 2012; Smith et al. 2008). Monocytes do not express *STAT4* in basal conditions but up-regulation follows IL-12, IL-18 and IFN-*α* stimulation (Frucht et al. 2000; Schindler et al. 2001). STAT5 phosphorylation in monocytes is mainly induced by granulocyte macrophage-colony stimulating factor (GM-CSF) and promotes differentiation into macrophages. Interestingly, STAT5 downstream gene targets such as prostaglandin synthase 2 (*COX2*) and *IL-10* were not dysregulated in CD14+ monocytes in my data. In other chronic inflammatory diseases such as T2D, persistent STAT5 phosphorylation has been found in circulating monocytes isolated from T2D upon GM-CSF (Litherland et al. 2005). Further investigation to determine phosphorylated STAT4 and STAT5 protein abundance will be require to determine if the down-regulation at the transcript level observed in psoriasis CD14+ monocytes is biologically relevant. In CD8+, expression of *IFNG* a gene activated by the IL-12 signalling pathway, was down-regulated when compared to healthy controls. Down-regulation of *IFNG* has previously been reported in unstimulated and stimulated macrophages derived from AS patients, in SF from SpA patients compared to RA and in a SpA rat model (Smith et al. 2008; Fert et al. 2014). This down-regulation was accompanied by an overall inverse transcriptional response of IFN-regulated genes, which was not seen in my data. Moreover, the reduced expression of *IFNG* in knock-out mice has been shown to increase activation of the IL-23/IL-17 axis, which is pivotal in psoriasis

pathogenesis (Can˜ete et al. 2000; Chu et al. 2007). Therefore, down-regulation of *IFNG* may actually result in a pro-inflammatory effect.

In CD8+ specifically, DEGs showed significant enrichment for three very

relevant pathophysiological pathways in psoriasis: NF-*κ*B, TNF and chemokine signaling. Important cross-talk between the NF-*κ*B and TNF signaling pathway was observed, with a number of dysregulated genes contributing to both. Interestingly, the enrichment of these pathways involved up-regulation of pro-inflammatory genes (e.g *ATF2*, *ATF4*, *RELA*, *RELB*) but also increased expression of well-characterised immunoregulatory genes. These included *NFKBIA* and *TNFAIP3*, also up-regulated in CD14+ monocytes and CD4+ cells respectively, and both associated with psoriasis GWAS signals. Polymorphisms within *TNFAIP3* or in its vicinity have also been associated with a number of chronic inflammatory diseases including MS, RA, SLE and T1D (Vereecke et al. 2011). *NFKBIA* codes for I*κ*B*α* which inhibits NF-*κ*B by binding it and preventing translocation to the nucleus. *TNFAIP3* codes for the zinc finger protein and ubiqitin-editing enzyme A20, and is up-regulated in the presence of inflammation and NF-*κ*B activation in order both to inhibit the NF-*κ*B TNF- mediated response and promote return to homeostasis. Either *NFKBIA* or *TNFAIP3* were found to be dysregulated in psoriasis PBMCs by Coda *et al.*, Lee *et al.* and Mesko *et al.* or in PBMCs from PsA patients versus controls (Dolcino et al. 2015). Interestingly, qPCR analysis in PBMCs from mild (PASI*<*4.84) and severe (PASI*>*4.84) psoriasis vulgaris revealed a significant negative correlation between *TNFAIP3* expression and disease severity (Jiang et al. 2012). Furthermore, this study also demonstrated that in the mild group of patients but not in the severe *TNFAIP3* expression was down-regulated when compared to healthy control PBMCs. This is in line with my findings with the caveat that all patients from my cohort would be classified as severe by Jian *et al.*. Altogether, the up-regulated expression of *TNFAIP3* and *NFKBIA* compared to healthy controls may not be

unexpected as it reflects a persistent inflammatory stimuli in psoriasisperipheral bloodand a mechanism that limits the systemic inflammatory response to some extent (Idel et al. 2003).

Of interest was also the up-regulation of the chemokine receptor *CCR10* in CD8+ cells from psoriasis patients. In circulation, expression of *CCR10* is restricted to a subset of circulating mCD4+ and mCD8+ T cells expressing the cutaneous lymphocyte-associated antigen (CLA), and preferentially recruited to cutaneous sites of inflammation (Hudak et al. 2002). Indeed, an increase of CCR10+ infiltrated T lymphocytes in psoriatic skin, where keratinocytes express *CCR10* ligand *CCL27*, has been demonstrated (Homey et al. 2002). The up-regulation of *CCR10* in my data could potentially suggest an increase of mCD8+ CCR10+ cells ready to migrate into the skin lesions. Moreover, correlation between the frequency of CTLA+ CD8+ cells and disease severity measured by PASI score has been found (Sigmundsdttir and 2001). Overall, these data have revealed dysregulation between psoriasis patients and controls for relevant immune genes showing pro- and anti-inflammatory effects in circulating immune cells. Although down-regulation of pro-inflammatory genes and up-regulation of anti- inflammatory genes has been detected, understanding the overall effect of those interactions in the inflammatory response requires further investigation.

## Correlation between changes in chromatin accessibility and gene expression

In this chapter, greater changes in gene expression have been identified compared to chromatin accessibility. Strikingly, in CD8+ cells, 687 transcripts were differentially expressed between psoriasis and healthy controls but only

55 regions showed differential chromatin accessibility when performing the same contrast and only six of the 687 were proximal to a DAR. This may relate to the complexity of gene regulation with a multi-component regulatory process manifesting in differential gene expression fo which chromatin accessibility is only one part. Correlation between chromatin accessibility measured by ATAC and gene expression has been reported to some extent in a number of studies, with limitations in establishing relationships between enhancer regions and the regulated target (Ackermann et al. 2016; Wang et al. 2018). An example of a relevant DEG nearby a DAR was *TIAM1*, with increased chromatin accessibility and gene expression in psoriasis CD8+ cells compared to healthy controls. *TIAM1* is involved in IL-17 expression and cell migration into the inflamed tissue (Kurdi et al. 2016; Grard et al. 2009). However, no eQTL or chromatin conformation data in this cell type has been found to formally establish a link between the region harbouring this DAR and *TIAM1* expression.

## Trancriptomic profiles in lesional and uninvolved psoriatic epidermis

Investigation of differences in the transcriptomic profile between paired lesional and uninvolved skin was conducted for three psoriasis patients. Most previous transcriptional studies in psoriasis have used full thickness skin biopsies, comprising a mix of cell types including fibroblast, adipocytes, keratinocytes from the epidermis and dermis and infiltrated immune cells. A study from Ahn and colleagues demonstrated large differences in gene

expression between whole biopsies and FACS-isolated keratinocytes, and the former may be masking keratinocyte-specific pathophysiological differences in many previous studies using psoriasis skin biopsies (Ahn et al. 2016). In this chapter, RNA-seq was conducted on epidermal sheets isolated from whole biopsies and a total of 1,227 DEGs were identified. Comparison with the Tervaniemi *et al.* study contrasting gene expression between lesional and uninvolved epidermis split biopsies, mainly formed by epidermis, revealed an overlap of only 359 out of the 1,227 DEGs detected in my data (12.1% of Tervaniemi *et al.* DEGs). Interestingly, the overlap with the Tsoi *et al.* study using whole biopsies was similar (505, 13.1% of Tsoi *et al.* DEGs) and only 5 genes had opposite direction of change, in contrast with the 75 showing discrepancies with Tervaniemis study. The similar percentage overlap with the Tsoi study despite the different source material could simply be the result of greater power in that study.

Genes consistently up-regulated across the three studies included genes from the *S100A* family. The *S100* family are located in the chr1p21 locus, which harbours genes involved in keratinocytes differentiation, act as calcium sensors and may also have a chemotactic effect (Eckert et al. 2004). In particular, *S100A9* and *S100A12* undergo up-regulation in psoriasis (Broome et al. 2003), with the latter involved in the T cell proliferative response and IFN-*γ* and IL-2 production (Moser et al. 2007). *LCE3B*, also at the chr1p21 locus, was also upregulated in lesional skin compared to uninvolved in all three studies. *LCE3B/C del*, a psoriasis GWAS association, is found in approximately 60 to 70% of European psoriasis patients (Cid et al. 2009). As explained in Chapter 1, *LCE* gene expression is induced upon disruption of the skin barrier, and expression of *LCE3B* and *LCE3D* has been only detected in lesional but not uninvolved psoriatic skin of heterozygous individuals (Cid et al. 2009; Bergboer et al. 2011).

Pathway enrichment analysis for the DEGs between lesional and uninvolved skin revealed a number of relevant biological processes for psoriasis pathophysiology. These highlighted alterations in cell cycle and metabolic processes, including amino acid metabolism, glycolysis, and hypoxia (HIF-I signalling), which had been identified in other studies performing DGE analysis between lesional and uninvolved skin or genome-wide pathway analysis (Coda et al. 2012; Gudjonsson et al. 2010; Aterido et al. 2016; Tervaniemi et al. 2016). Up-regulation of the hypoxia-inducible TFs HIF-1*α* and HIF-2*α* has been reported in lesional skin, correlating with an increase in *VEGF* transcript levels, a gene regulated by HIFs that mediates pathological angiogenesis also characteristic of psoriasis (Rosenberg2007). No correlation was observed between *HIFA* and *VEGF* in the data presented here, likely due to the small sample size. Moreover, HIF-I signalling is also involved in regulating Th-17/Treg ratios and therefore in perpetuation and termination of the immune response (Dang2013).

Immune-related pathway enrichment were also found including Th-17, IL-12, cytokine-cytokine and NOD-like signalling. Interestingly, NOD-like signalling was reported to be enriched in DEGs between lesional and uninvolved skin in a contemporary study by Tervaniemi and colleagues (Tervaniemi et al. 2016). Tervaniemi mainly attributed this novel pathway to the greater sensitivity of RNA-seq compared to microarrays to detect changes in gene expression for genes involved in this pathway. The fact that Tsoi *et al.* also used RNA-seq and did not show enrichment for NOD-like signalling is likely due to the type of biopsy, highlighting the value of studying epidermis instead of full thickness skin to uncover dysregulation of functional pathways in keratinocytes. NOD-like signalling involves signal transduction by NOD-like receptors, a type of pattern-recognition receptors, which can recruit and activate caspases

into the inflammasomes or trigger inflammation through NF-*κ*B and MAPK. Amongst the genes contributing to this pathway *CARD6*, *IFI16*, *NOD2* and *NLRX1* overlapped with Tervaniemis data and showed up-regulation in both. Notably, polymorphisms in *NOD2* have been linked to inflammatory diseases such as CD, atopic eczema and arthritis and potentially with psoriasis and PsA (Zhong et al. 2013; Zhu et al. 2012).

## LncRNAs in psoriasis

## The largest number of differentially expressed lncRNAs between psoriasis patients and controls were found in CD14+ monocytes and CD8+ T cells (28 and 31, respectively for FDR<0.05) with 46 lncRNAs differentially expressed between lesional and uninvolved skin (FDR<0.05). S but no study has been conducted to identify differentially lncRNAs in a cell type-specific manner in peripheral blood from psoriasis patients. The role of lncRNAs has been studied in RA, SLE, AS, and PsA (Muller2014; Shi et al. 2014; Zhang et al. 2017; Dolcino et al. 2018).

Characterisation of lncRNA biological function is a developing field, which represents a limitation when interpreting these results (B et al. 2018). Some of the well-characterised dysregulated lncRNAs have a role in the immune response. For example, in psoriasis CD14+ monocytes the up-regulation of *HOTAIRM1* was found to be associated with down-regulation of the predicted target gene *USP1*. *UPF1* is involved in nonsense-mediated decay

and in partnership with the monocyte chemotactic protein-1-induced protein- 1 (*MCPIP1*) gene drives degradation of inflammation-related mRNAs to ensure maintenance of homeostasis (Mino et al. 2015). Down-regulation of *UPF1* in psoriasis CD14+ monocytes may suggest impairment of this homeostatic mechanism, contributing to disease pathophysiology in this cell type. Another example of a relevant lncRNA up-regulated in CD14+ monocytes was *NEAT1*, which has also shown up-regulation in SLE CD14+ monocytes. Knock- down demonstrated impairment of TLR4 signalling and down-regulation of inflammatory genes including IL-6 and CXCL10 (Zhang et al. 2016).

*MIR146A* was found to be differentially expressed between lesional and uninvolved skin and also when comparing psoriasis CD8+ T cells to healthy controls. Molecular studies have suggested a role for miR-146a as a negative regulator of the TLR4 pathway through inhibition of TNF associated factor 6 (*TRAF6*) and IL-1 receptor-associated kinase 1 (*IRAK1*) expression (Taganov et al. 2006). Down-regulation of miR-146a levels in CD8+ cells would support failure of one of the check-points controlling sustained inflammation and the subsequent pathophysiological implications while differential expression has been previously reported in psoriasis lesions (Lerman2014; Li2014; Tsoi et al. 2015) .

Other dysregulated non-coding RNAs in lesional epidermis relevant to psoriasis pathophysiology were *HG19* and *MIR31HG*, found to be down- and up-regulated respectively. Silencing miR-31hg in the keratinocyte immortal cell line HaCaT induced cell cycle arrest and inhibited cell proliferation, consistent with two characteristic aspects dysregulated in psoriatic keratinocytes (Gao et al. 2018). Overall, the lncRNA differential analysis conducted in this chapter gives an overview of dysregulation in blood and skin of psoriasis patients. A more comprehensive analysis could be performed to identify putative targets for all the identified lncRNAs. Those interactions could then be used to identify relevant biological processes through

network and pathway analysis using only those dysregulated lncRNAs matching dysregulated target genes, similarly to the strategy used by Dolcino *et al.,* 2018. However, such an analysis would likely require increased sample size to be appropriately powered.

## Differences in transcriptional dysregulation in blood and skin

Comparison of the dysregulated genes in circulating immune cells and in psoriatic skin revealed very limited overlap. CD14+ monocytes and CD8+ cells showed the greatest overlap of DEGs however almost half of genes showed opposite directions in differential expression consistent with a report by Coda and colleagues comparing DEGs in psoriasis patients and controls PBMCs to genes dysregulated between lesional and uninvolved skin biopsies (reference). Genes showing opposite changes in circulation and in skin included the GWAS genes *TNFAIP3*, *EGR2*, and *EGR3*. As previously mentioned, *TNFAIP3* down-regulation in lesional skin may reflect complete loss of an NF-*kappa*B pathway check-point to control and terminate the inflammatory response at the site of inflammation. Similarly, *EGR2* and *EGR3* are pivotal for control of inflammation and antigen-induced proliferation. In addition to its role in

regulating inflammation, down-regulation of *EGR2* in skin may also increase keratinocyte proliferation as has been shown in certain types of cancer (Wu et al. 2010).

Differences are also observed in the distinct enriched pathways. For example, DEGs in skin not only showed enrichment for immune-related functions but also highlighted metabolic dysregulation. Moreover, immune related pathway such as NOD-like signalling were only seen in skin. Likewise, up-regulation of genes from the *S100* family in lesional skin, such as *S100A7*, *S100A8*, *S100A9*, contributed to enrichment of IL-17 signalling and appeared to be a feature of dysregulated inflammation only in skin. Notably, these genes had also been reported as a specific hallmark of skin inflammation when compared to inflamed synovium from matched PsA patients, supporting the better outcomes for IL-17 antagonists in skin lesions compared to the inflammation of the synovium (Belasco et al. 2015).

## Allelic differences in chromatin accessibility

One of the particularly interesting psoriasis GWAS associations is the chr2p15 locus, where the lead SNP is located in an intergenic region 140Kb and 150Kb away from *B3GNT2* and *TMEM1* respectively. Although fine- mapping at chr2p15 failed in psoriasis (probably due to lack of power), fine- mapping analysis in AS implicated rs4672505 which was found to overlap a CD8+-specific

ATAC peak. Chromatin accessibility was observed to vary between individuals, with complete loss of the ATAC peak in some individuals. Integration of genotyping data with ATAC revealed allele-dependent chromatin accessibility, with loss of chromatin accessibility correlating with the risk allele. Promoter capture-C data linked rs4672505 to the *B3GNT2* promoter only in CD8+

cells, suggesting the accessible chromatin at rs4672505 may be highlighting an enhancer element interacting with *B3GNT2* promoter as a priming event (Javiere2016). *B3GNT2* is a major polylactosamine synthase involved in the post-translational modifications of carbohydrate chains, which are essential for cellcell, receptorligand and carbohydratecarbohydrate interactions. Interestingly, *B3GNT2* knock-out mice demonstrated more sensitive and strongly proliferating T cell and B cell responses to stimulation compared to wild-type (Togayachi et al. 2010). In T cells, this effect was linked to a reduction of polylactosamine chains in co-stimulatory accessory molecules such as CD28, overall leading to enhanced initiation of the immune response *in vitro*.

Up-regulation of *B3GNT2* in this context could be contributing to attenuation and modulation of CD8+ activation. Under this scenario, the presence of the risk allele (A) at this stimulus-specific enhancer could increase risk of disease

by reducing chromatin accessibility, in both homozygous and heterozygous individuals. Establishing a more comprehensive and accurate model to further explain the functional role of rs4672505 in psoriasis susceptibility will require additional work, such as increasing the sample size to acquire more homozygous individuals for the risk allele, studying chromatin accessibility and *B3GNT2* expression in relation to rs4672505 genotype in stimulated CD8+ and performing genome editing to establish a causal relationship between the risk allele and altered gene regulation.

## Limitations in the approach and future work

Although the work in this chapter has shed light on the chromatin landscape and gene expression in psoriasis in a cell type and tissue specific manner, a number of limitations are noted. Due to difficulties in optimising ATAC protocols to yield good quality data, mapping chromatin accessibility in lesional and uninvolved keratinocytes was not achieved. This may have revealed larger differences in chromatin accessibility between psoriasis patients and controls compared to circulating immune cells, as in other studies performed in affected tissues (Scharer et al. 2016; Wang et al. 2018). Additionally, chromatin and transcriptomic profiles from skin infiltrated cells could be generated using FACS or single-cell technologies to better understand the changes in chromatin accessibility and gene expression driven by the inflammatory stimuli at the site

of inflammation. Moreover, generating this data would also allow comparison to the profiles obtained in blood to better understand disease pathophysiology.

Other limitations in this work include its relatively small sample size, lack of genome-wide genotyping data and skin biopsies only being available for three patients in the cohort. These limitations are intrinsic to time and project budget constraints and will be addressed as the study continues. Recruitment of additional patients would allow validation of the findings described in this chapter. Genotyping data would permit the study of chromatin accessibility in a genotype-specific manner, using the current samples with prospective integration of chromatin conformation data (Kumasaka et al. 2018). Importantly, this will enable exploration of changes in chromatin accessibility at GWAS loci in combination with fine-mapping, similarly to the chr2p15 locus analysed in this chapter. Furthermore, new sample recruitment could be used to study chromatin accessibility and gene expression in additional cell populations sorted by FACS and also to include *in vitro* stimulations. Overall, this strategy would allow better characterisation of the differences and similarities between patients and controls in context-specific regulatory elements *in vivo* and *in vitro* (Peeters et al. 2015).

Finally, improvements in analytical methods will also be required to ascribe chromatin accessibility changes in enhancers to target genes potentially regulated by these regions. These could involve a more systematic integration of available chromatin conformation data, eRNA FANTOM data and also use of analytical models and tools currently available or that may be further developed in the future to specifically address this challenge (Wang et al. 2016; Cao et al. 2018).

## Conclusions

In this chapter, use of contemporary methods for epigenetic profiling (as established in the previous chapter) together with gene expression quantification has allowed

characterisation of the regulatory landscape in relevant cell types isolated from psoriasis patients and healthy individuals. Minor differences in chromatin accessibility and H3K27ac modifications between psoriasis and healthy controls have been identified in circulating immune cells. Conversely, a number of relevant biological processes dysregulated in the context of psoriasis have been shown at the transcriptional level both, in circulating cells and in psoriatic epidermis. Moreover, this chapter illustrates how GWAS signals may be interpreted through integration of multiple data types. Overall, the protocols established and data generated in this chapter provide a valuable resource that may be built upon in future work.

%SNPs tagging all psoriasis GWAS loci showed enrichment for H3K27ac in twenty-six relevant cell types and/or tissues, with approximately 50\% of variants in LD with those lead SNPs overlapping annotated enhancers in total CD4$^+$ T cells \parencite{Lin2018}.

*MIR22HG*

*CACNA2D3*

*ZNF550*

*SDC2*

*RASGEF1B*

*MFSD2A*

*ACTA2*

*ZNF2 HSPH1 SPN*

*SERTAD3*

*MGST1*

*CDKN2D*

*IFI30*

*CD83*

*MB21D1 HSPA8*

*ACCS*

*HLA−A DIEXF*

*YPEL5*

*BCL6 DSC2*

*ZNF786 DDX31 ANXA6 ZNF574*

*PRR5*

*TRIM65*

*PIM3*

*DIRC2*

*SLC2A3*

*KBTBD6 ITPRIPL1*

*URB2*

*PPP3R1*

*NFIL3*

*ZNF250*

*CBX8*

*RPP14 PAFAH2*

*STIP1*

*HLX*

*ARID5B FBXL8 CD69 MAL RPL17 UNC45B*

*ZNF597 HS3ST3B1*

*GOLGA5P1*

*DLGAP1−AS1*

*DPP4 EIF1*

*HIST1H1D*

*TNFSF8 CTD−3184A7.4*

*HLA−A*

*SERTAD1 TAF1A KBTBD7 INTS5*

*MCOLN3*

*CCDC117*

*GPR18*

**Not sig. FDR<0.05**

*EIF1 LMNA*

*CYB5D2 NCF4*

*APC*

*ZNF436*