



Functional genomics of psoriasis

Alicia Lledo Lara
Hertford College
University of Oxford

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

Abstract

Functional genomics of psoriasis

Alicia Lledo Lara, Hertford College, Trinity Term 2018

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

This is my abstract...

Acknowledgements

Thank you, thank you, thank you.

Declarations

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

Submitted Abstracts

Title	Year
Authors	

Associated Publications

Title

Journal

Authors

Other Publications

Title

Journal

Authors

Contents

Abstract	i
Acknowledgements	ii
Declarations	iii
Submitted Abstracts	iv
Associated Publications	v
Contents	vi
List of Figures	vii
List of Tables	viii
Abbreviations	ix
1 Cross-tissue comparison of chromatin accessibility, gene expression signature and immunophenotypes in PsA	1
1.1 Introduction	1
1.2 Results	1
1.2.1 PsA patients cohort description and datasets	1
1.2.2 Immune cellular composition of blood and synovial fluid in the PsA cohort	4
1.2.3 Chromatin accessibility differential analysis in immune cells reveals differences between SF and PB	5
1.2.4 Pathway and TFBS enrichment analysis highlight tissue-specific functional differences in chromatin accessibility . .	16
1.2.5 Differential gene expression analysis in paired circulating and synovial immune cells	19
1.2.6 Mass cytometry reveals CD14 ⁺ monocytes as the most active cytokine producing cells	32
1.3 Discussion	35
Appendices	36
A Establishment of methods to assess genome-wide chromatin accessibility	37
B Cross-tissue comparison analysis in PsA	40

List of Figures

1.1	Comparative percentages of PB and SF immune cellular composition from the PsA cohort.	5
1.2	QC of FAST-ATAC PsA samples in four cell types	7
1.3	Combined PCA analysis of all four cell types isolated from blood and SF.	8
1.4	Enrichment of eQTLs publicly available data in the combined cell type and tissue chromatin accessibility master list for the PsA cohort.	10
1.5	Annotation with genomic regions and chromatin states of the PsA DOCs from the four cell types differential analysis.	13
1.6	Enrichment of PsA DORs for the FANTOM5 eRNA dataset.	14
1.7	Differentially accessible regions located within gene bodies in CD14 ⁺ monocytes and NK cells from PsA patients.	15
1.8	Distinct enriched pathways across SF and PB in CD14 ⁺ ,CD4m ⁺ ,CD8m ⁺ and NK.	18
1.9	Enrichment of eRNA cell type-specific PsA DORs for conserved TFBS.	19
1.10	Representation of the fold-change in gene expression between SF and PB for the significant genes (pval<0.05) in at least one of the cell types.	21
1.11	Expression changes in immune-relevant genes between SF and PB in CD14 ⁺ monocytes, mCD4 ⁺ and mCD4 ⁺ cells.	23
1.12	Chromatin accessibility at the <i>FN1</i> gene in CD14 ⁺ monocytes isolated from SF and PB.	24
1.13	Protein network analysis based on the immune PCR array expression data.	29
1.14	Expression changes in immune-relevant genes between SF and PB in CD14 ⁺ monocytes, mCD4 ⁺ and mCD4 ⁺ cells.	33
1.15	Mean expression of IL-2, IL-8, TNF- α and IFN- γ in SF and PB from PsA patients.	34
1.16	Mass cytometry analysis of TNF- α production in SF and PB by CD14 ⁺ monocytes after protein transport blockade with BFA.	35
A.1	FAST-ATAC and Omni-ATAC NHEK tapestation profiles.	38
A.2	Assessment of TSS enrichment from ATAC-seq and FAST-ATAC in healthy and psoriasis skin biopsies samples.	39
B.1	Permutation analysis SF vs PB in CD14 ⁺ ,CD4m ⁺ ,CD8m ⁺ and NK.	40

List of Tables

1.1	Description of PsA patients cohort recruitment and metadata.	3
1.2	Datasets generated for the PsA cohort samples.	4
1.3	Summary results of the chromatin accessibility analysis between SF and PB in PsA samples	11
1.4	Annotation of gene body DORs in four cell types from PsA samples.	14
1.5	Immune genes with significant modulated expression in SF proximal to a DOR in ATAC-seq.	22
1.6	Pathway enrichment analysis for the modulated genes between SF and PB in CD14 ⁺ and mCD4 ⁺	27

Abbreviations

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

Chapter 1

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

1.1 Introduction

The techniques to generate the scRNA-seq data have also evolved and SmartSeq2 and 10X Chromium are the two main I haven't talked at all about IL-2 and IL-8 in the relevant cytokines in disease. The reason is because they are not the most relevant ones but due to cytotoxic constraints they are together with TNF α and IFN γ among the ones yielding the best quality results.

1.2 Results

1.2.1 PsA patients cohort description and datasets

In this study peripheral blood (PB) and SF were collected from a cohort of six PsA patients, with equal numbers of males and females (Table 1.1). All the patients presented oligoarticular joint affection and had been first diagnosed with psoriasis. The cohort presented a mean of 1.5 tender or swollen affected joints (TJC66 and SJC66), which is characteristic of the oligoarticular form of disease (involving four or fewer joints). Regarding global assessment, the mean scores

Cross-tissue comparison analysis in PsA

for the patient and physician evaluation were X and 3, respectively, in a scale of 1 to 5. These four measurements including joints and global assessment compose the PsARC disease activity scores, used by clinicians as the main indicator of response to treatment by recommendation of the National Institute for Health and Care Excellence (NICE) (Chapter ??).

The mean age of the cohort at the time of diagnosis was 44.3 years old and the mean disease duration 8.8 years. Interestingly, PsA1728 was diagnosed at a later age compared to the other patients in the cohort (late PsA onset clinical significance??). Moreover, C-reactive protein (CRP) levels, other marker of inflammation, presented an average of 17.45 mg/L and was particularly higher in PsA1719 and PsA1728 compared to the other patients. At the time of sample recruitment all the PsA patients were naive for treatment and only PsA1505 had been on MTX therapy in the past for xxx months/years (how many years ago?). Post-visit, most of the patients qualified for TNFi biologic therapy xxxx.

Cross-tissue comparison analysis in PsA

Table 1.1: Description of PsA patients cohort recruitment and metadata.. PsARC disease activity score is composed of tender joint count (TJC66) and swollen joint count 66 (SJC66), joint pain (4 point score) and self-patient and physician global assessment (5 point score). Joint pain and global assessment use a likert scale based on questionnaire answers that measure the level of agreement with each of statements included. C-reactive protein (CRP).

Sample ID	Sex	Age at diagnosis	Disease duration (months)	Type	TJC66/SJC66 assessment	Physician assessment	Patient assessment	CRP (mg/L)
PsA1718	Female	17	180	Oligo	2/2	3	3	6
PsA1719	Male	33	24	Oligo	1/1	3	4	36.6
PsA1607	Male	42	108	Oligo	1/1	4	3	8
PsA1728	Female	72	48	Oligo	2/2	3	4	43.2
PsA1801	Female	53	168	Oligo	2/2	3	3	9.9
PsA1505	Male	35	108	Oligo	1/1	2	2	1
Total	-	44.3	106	-	1.5/1.5	3	3.2	17.45

For each of the patients, paired data in PB and SF was generated from bulk mononuclear cells or specific isolated cell types of interest (detailed in Table 1.2 and Chapter ??). Due to project contrains, ATAC-seq, PCR gene expression array, scRNA-seq and mass cytometry were not generated for all six individuals in the cohort.

Sample ID	% FAST-ATAC	RNA PCR array	scRNA-seq	Mass cytometry
PsA1718	Yes	No	No	Yes
PsA1719	Yes	Yes	No	Yes
PsA1607	Yes	Yes	Yes	Yes
PsA1728	No	Yes	No	No
PsA1801	No	No	Yes	No
PsA1605	No	No	Yes	No

Table 1.2: Datasets generated for the PsA cohort samples. Four types of data were generated in a paired way between PB and SF from the same individual. The available data sets vary between individuals due to project constrains. FAST-ATAC data was generated for CD14⁺, mCD4⁺,mCD8⁺ and NK cells. RNA expression by PCR array was performed for CD14⁺, mCD4⁺ and mCD8⁺. scRNA-seq data was generated using 10X technology in bulk PBMCs, bulk SFMCs and sorted mCD4⁺ and mCD8⁺ isolated from both tissues.

1.2.2 Immune cellular composition of blood and synovial fluid in the PsA cohort

The immune cellular composition of three PsA samples (PsA1718, PsA1719 and PsA1607) was characterised in SF and PB using the ICS mass cytometry panel in Chapter ?. For both tissues, mCD4⁺ (between 32.1 and 55.6%) constituted the most abundant cell type followed by mCD8⁺ (between 16.9 and 24.9%) and CD14⁺ "non-classical" monocytes (between 6.9 and 21.7%). Consistently with previous studies, a trend of increased percentage of mCD8⁺, NK and pDC in SF compared to PB were observed (**Ross2000; Jongbloed2006; Spadaro2004**). Conversely, a tendency towards reduced proportions of B cells in SF compared to PB reinforced the lack of contribution of the humoral immune response in

Cross-tissue comparison analysis in PsA

PsA pathophysiology (**Spadaro2004**). The observed differences cell composition between SF and PB were not statistically significant for any of the twelve analysed populations likely due to the small samples size (only three individuals) available for the analysis. Further increase in the sample size will probably proof statistical significance for the observed differences in immune cell composition between the two tissues in lines with the results published by other studies.

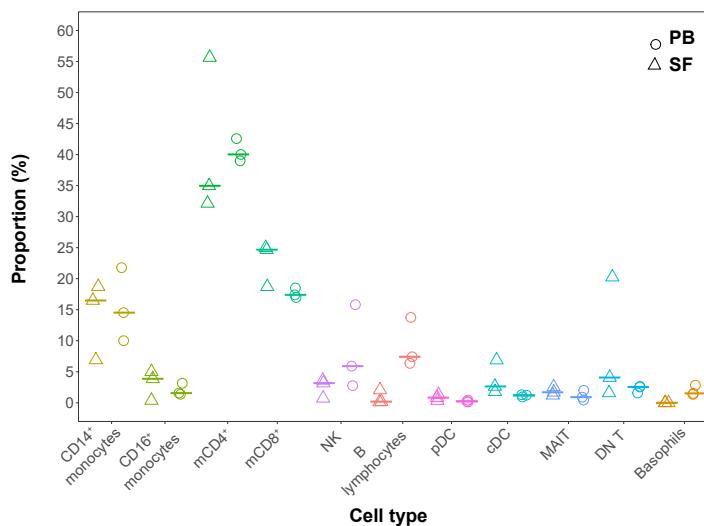


Figure 1.1: Comparative percentages of PB and SF immune cellular composition from the PsA cohort. Percentages of each of the twelve cell types identified by mass cytometry are shown by individual and tissue for PsA1718, PsA1719 and PsA1607. Horizontal line represents the median percentage for a particular cell type in the appropriate tissue (SF or PB). Each of the cell types is displayed in a different colour. Central DC=cDC, mucosal-associated invariant T=MAIT, DN=double negative.)

1.2.3 Chromatin accessibility differential analysis in immune cells reveals differences between SF and PB

Quality control of open chromatin regions

The twenty four FAST-ATAC PsA samples from four different cell types and two tissues (PB and SF) were sequenced to a median of 158M reads (79M paired-end) per sample. After filtering for low quality mapping, duplicates and MT reads, the median total number of reads ranged between 46.6 and 70.2 M reads, upon cell type (Figure 1.2 a). Overall, MT and duplicated reads accounted for

Cross-tissue comparison analysis in PsA

a median between 40 to 62.2% of the total number of reads depending on cell type (Figure 1.2 b) and importantly contributing to the loss of reads FAST-ATAC. The final number of reads remaining after filtering was inversely related to the percentage of MT and duplicated reads identified. For example, mCD14⁺ and mCD8⁺ presented the lowest median of total number of reads after filtering concomitantly with the greatest percentage of combined MT and duplicated reads.

Regarding sample quality, TSS enrichment analysis showed differences in the levels of background noise across cell types and highlighted the variability of FAST-ATAC performance (Figure 1.2 c). A trend towards greater TSS enrichment in PB samples compared to SF was observed in all four cell types. mCD4⁺ and mCD8⁺ presented the best signal-to-noise ratios, with median of 19.1 and 23.1 fold enrichment, respectively. In contrast, NK was the cell type with the lowest TSS enrichment values. Particularly, the fold enrichment for PsA1719 and PsA1607 were 7.3 and 6.2, respectively, both just above the 6 to 10 acceptable range from ENCODE. If a larger cohort size was available these two samples would be removed from the differential analysis, since high background levels can reduce the power of this approach.

Cross-tissue comparison analysis in PsA

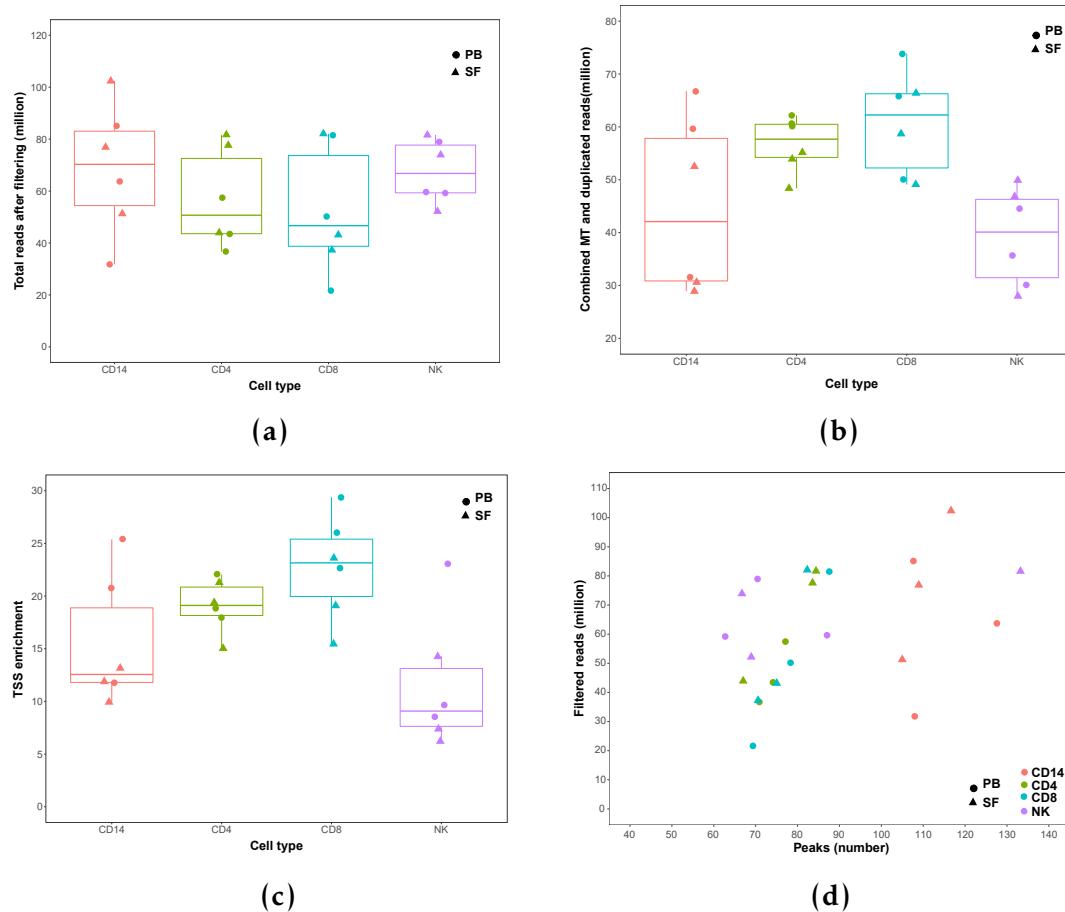


Figure 1.2: QC of FAST-ATAC PsA samples in four cell types.

When identifying open chromatin regions by peak calling using standard filtering (FDR<0.01), the number of accessible regions per sample ranged between $\sim 62 \times 10^3$ to $\sim 133 \times 10^3$ (Figure 1.2 d). A positive correlation between the number of called peaks and number of reads after filtering could be observed in the data. For example, CD14⁺ was the cell type with greatest number of called peaks (108.4×10^3) as well as the greater median of reads remaining after filtering when compared to the other three cell types (Figure 1.2 a). Overall, appropriate number of peaks were called in all samples and no concerning outliers were identified.

Accessible chromatin reflects cell type specificity and functional relevance

To determine the ability of in house pipeline open chromatin to recapitulate cell type specificity in the PsA sample cohort, a combined master list including all four cell types and the two tissues was built. Following Chapter ?? and Chapter ??, the combined master list contained open chromatin regions identified in at least 30% of the samples (7 out of 24 samples), regardless cell type and tissue to avoid any bias. PCA analysis based on chromatin accessibility showed that most of the samples variability (PC1 65.6% of the variability) correlated with cell type, leading to sample separation in four cluster (Figure 1.3). As expected, the myeloid ($CD14^+$ monocytes) and lymphoid (m $CD4^+$ and m $CD8^+$) clusters were the ones presenting greatest differences in chromatin accessibility pattern whereas the m $CD4^+$ and m $CD8^+$ clusters were the most similar between them. Modest clear separation between SF and PB samples was found in the m $CD4^+$, m $CD8^+$ and NK clusters.

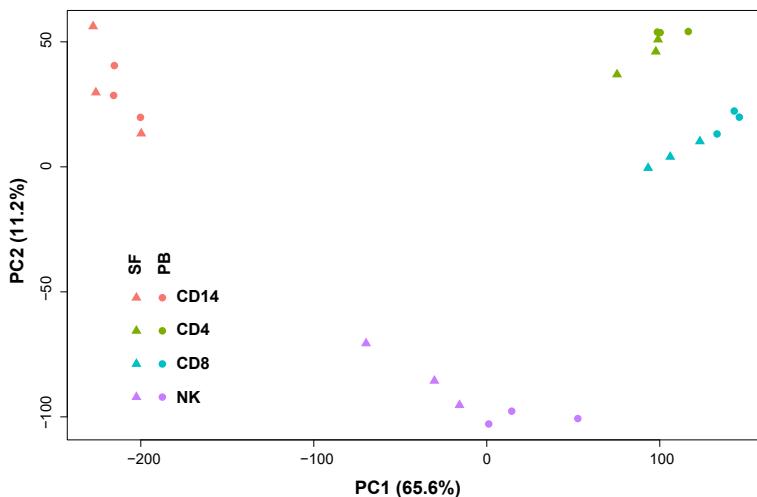


Figure 1.3: Combined PCA analysis of all four cell types isolated from blood and SF.

The ability to capture putative regulatory regions within the identified accessible chromatin in the four cell types was also explored. Enrichment analysis of different eQTL publicly available data sets for the combined chromatin accessibility master list was performed. Amongst the GTEx eQTL

Cross-tissue comparison analysis in PsA

data, the greatest (z-score) and most significant (-log₁₀FDR) enrichment was found for the venous blood data set (red dot), consistently with the cell types included in the study (Figure 1.4 a). Moreover, the greatest enrichments of immune-related cells in the accessible chromatin regions were found for CD14⁺ monocytes (importantly unstimulated, LPS 2h and IFN- γ 24h) followed by mCD8⁺ T cells (Figure 1.4 b). Interestingly, the B cells eQTLs were the least enriched in the chromatin accessibility data, consistently with the limited relevance of this cell type in psoriasis and PsA pathophysiology as well as the absence of this cell type amongst the ones included in the study.

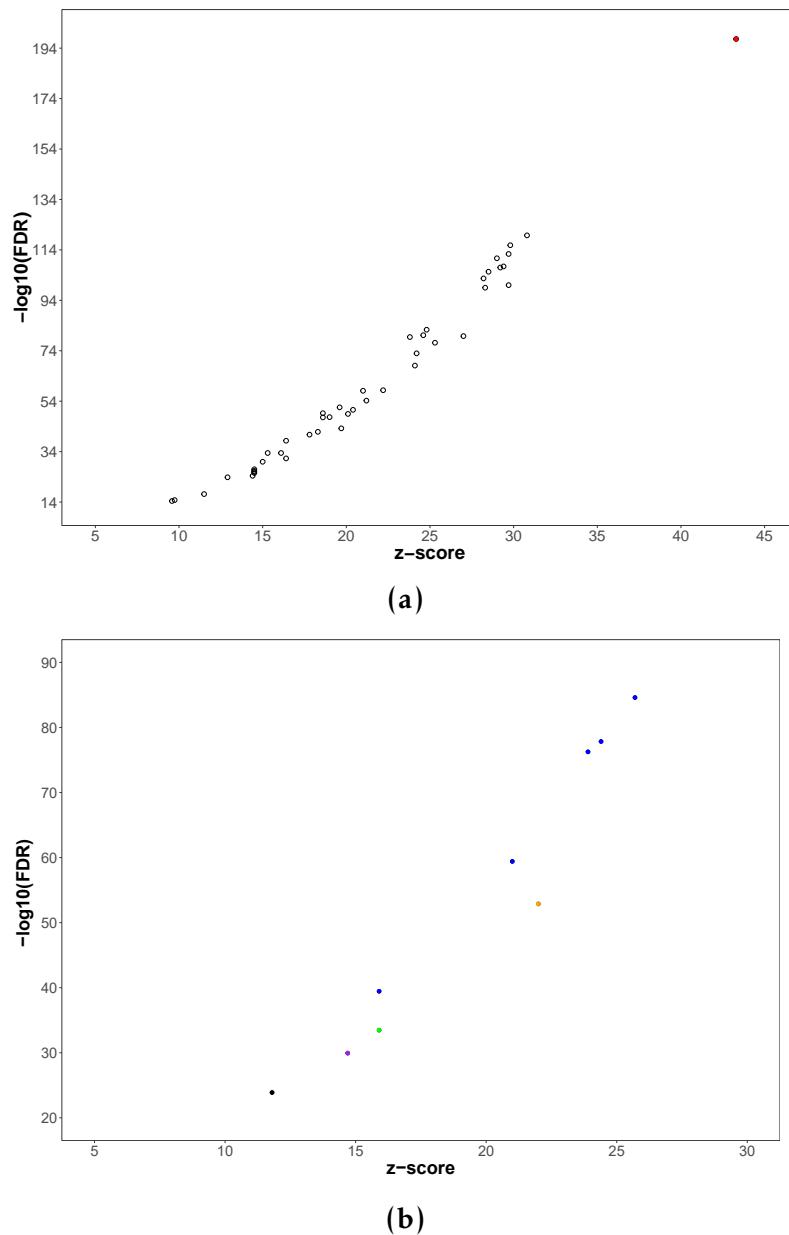


Figure 1.4: Enrichment of eQTLs publicly available data in the combined cell type and tissue chromatin accessibility master list for the PsA cohort. The dot plots showed the z-score values of the enrichment analysis in the x-axis and the significance ($-\log_{10}\text{FDR}$) in the y-axis for a) GTEx eQTL datasets and b) non-GTEx immune-related cell types including CD14⁺ monocytes (unstimulated, 2 or 24h LPS stimulated and 24h IFN γ stimulated) in blue, B-cells in black, total CD4⁺ in green, total CD8⁺ in orange and neutrophils in purple.

Characterisation of the differential accessible chromatin regions

Differential chromatin accessibility analysis was performed per cell type using a paired design between SF and PB for each of the four cell types (Table 1.3).

Cross-tissue comparison analysis in PsA

For each of the cell types a master list containing chromatin accessible regions in at least 30% of the samples (~2 samples), regardless the tissue. In all for analysis an 80% cut-off for background noise was applied to the count matrix, as previously explained in Chapter 3. Only differentially open regions (DORs) identified with DESeq2 and also shared with quantile normalisation limma voom analysis were considered downstream. The CD14⁺ monocytes and NK showed a greater proportion of DORs (23.3 and 8.9%, respectively) compared to mCD4⁺ and mCD8⁺ T cells. In CD14⁺ monocytes a significantly greater number of DORs more accessible in SF (3,779 DORs) were found compared to PB (1,506 DORs). Conversely, the number of open DORs in Sf and PB were evenly distributed between in the other three cell types.

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

Table 1.3: Summary results of the chromatin accessibility analysis between SF and PB in PsA samples.

Permutation analysis using the ten unique possible combinations demonstrated that greater number of DORs than expected by chance were obtained in the differential analysis (Figure B.1). This reinforced the fact that the identified changes in chromatin accessibility are driven by true differences between SF and PB in all the analysed cell types and therefore are specific.

When performing genomic annotation of the DORs in each of the cell types, intronic and intergenic regions consistently represented 80% or more of all regions with differential accessibility (Figure 1.5 a). Universal promoter regions represented the third genomic feature with approximately between 5 to 15% of the DORs annotation. In addition to this, the fifteen cell type-specific chromatin states from the Epigenome Roadmap segmentation maps were also

Cross-tissue comparison analysis in PsA

used for annotation (Figure 1.5 b). For all four cell types DORs, between 44.96 and 72.11% were annotated as weak enhancers, which represented the most prominent category and the most significantly enriched (data not shown). This was consistent with the predominance of introns and intergenic regions, as those are the preferred location for enhancer elements. Although modest percentages of heterochromatin and repetitive regions were observed, enrichment was not significant for these two chromatin states in any of the four cell types (data not shown).

Cross-tissue comparison analysis in PsA

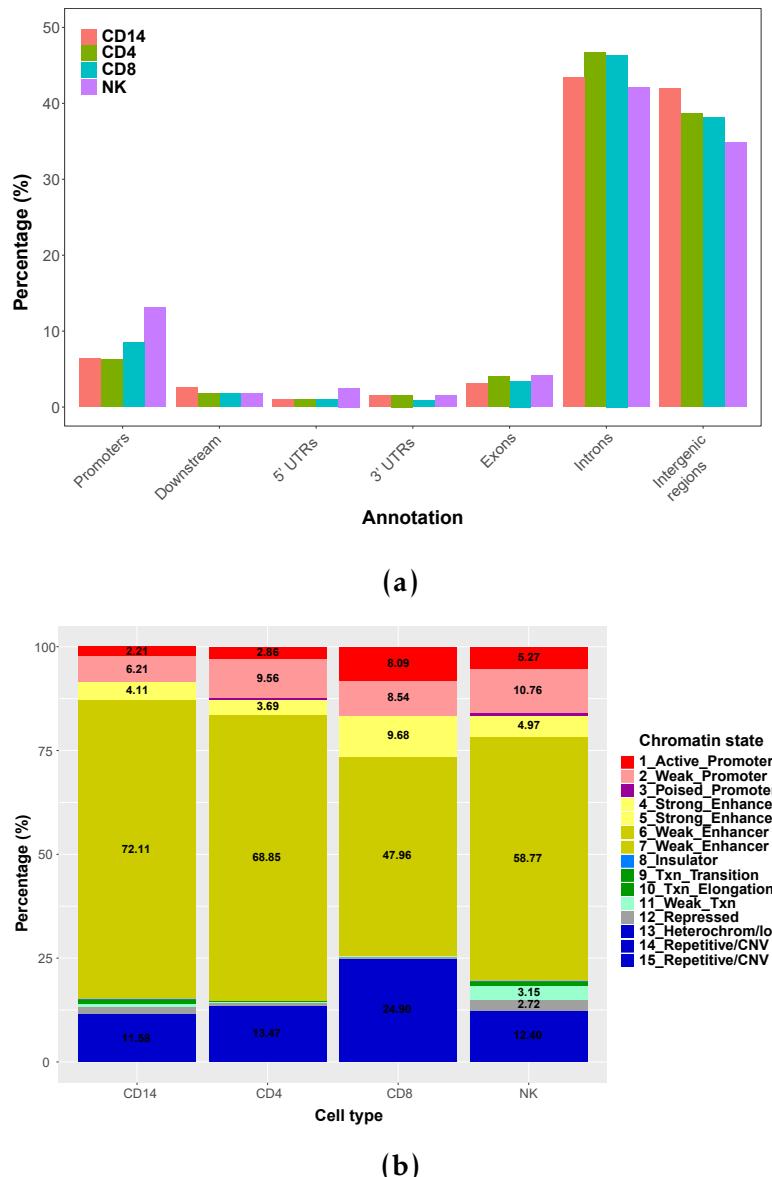


Figure 1.5: Annotation with genomic regions and chromatin states of the PsA DOCs from the four cell types differential analysis. xxxx

The functional relevance of the differential chromatin accessibility in terms of regulation of gene expression was further investigated by integration of the eRNA data from the FANTOM5 project. Statistically significant enrichment for robust and permissive enhancers was found for the DORs identified in the four cell types (Figure ??). Robust enhancers are those for which transcription was significantly detected at the genome-wide level in at least one primary cell type or tissue, whereas the permissive set included also those not passing the

Cross-tissue comparison analysis in PsA

filtering criteria (Andersson2014). Moreover, DORs from all four cell types also presented significant enrichment for the corresponding cell type eRNA set amongst the top two most enriched. The proportion of DORs overlapping the appropriate cell type set of expressed eRNAs ranged between 19.8% (83 open in SF and 160 open in PB) in NK and 31.8% (83 open in SF and 160 open in PB) in CD4⁺.

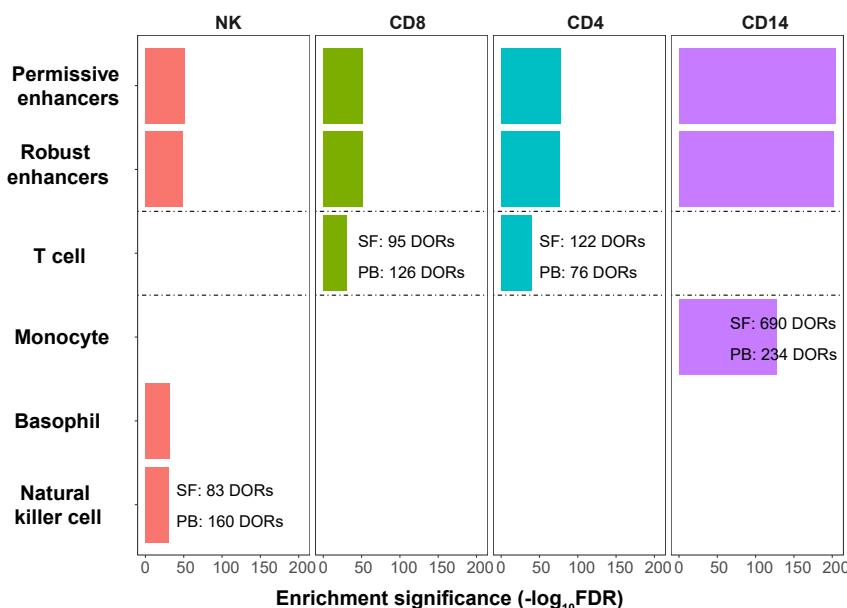


Figure 1.6: Enrichment of PsA DORs for the FANTOM5 eRNA dataset. xxx

From the differential analysis between SF and PB in all four cell types, a number of DORs were overlapping a gene body (Table 1.4). Interestingly, the majority were located within introns instead of untranslated regions (UTRs) and have also been annotated as weak or strong enhancers according to the cell type specific chromatin segmentation map.

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

Table 1.4: Annotation of gene body DORs in four cell types from PsA samples. xxxx

Cross-tissue comparison analysis in PsA

As an example, a more accessible PB DOR located in an intron of the VAV3 gene and significantly expressed as eRNA was identified in NK (Figure 1.7 a). Additionally, for all four cell types, a number of gene entities contained more than one DOR, showing the same direction of chromatin accessibility between SF and PB. For example, in CD14⁺ two DORs located at the 5' and 3' UTRs for IL7R gene were found to be more accessible in SF compared to PB (Figure 1.7 b).

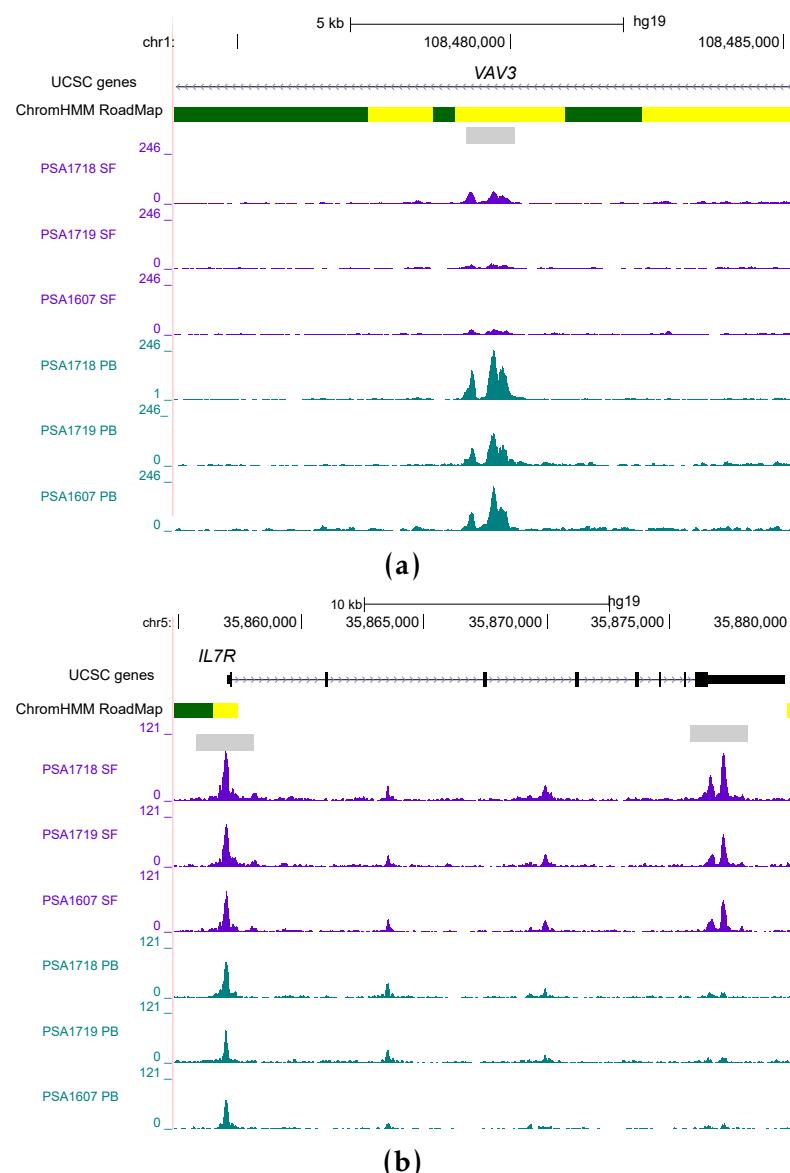


Figure 1.7: Differentially accessible regions located within gene bodies in CD14⁺ monocytes and NK cells from PsA patients. xxxx

The relevance of the differences in chromatin accessibility in the context of psoriasis and PsA GWAS hits was also addressed. Enrichment analysis of psoriasis and PsA GWAS hits for the DORs in each of the four cell types was performed using XGR co-localisation and permutation analysis. At the SNP level, no significant enrichment was reported between DORs and GWAS lead SNPs and those in LD $r^2 \geq 8$. However, significant enrichment (2-fold enrichment and empirical p-val 0.043) of the psoriasis and PsA GWAS LD blocks was observed only for the CD14⁺ DORs.

1.2.4 Pathway and TFBS enrichment analysis highlight tissue-specific functional differences in chromatin accessibility

For each of the four sets of DORs, pathway enrichment analysis was conducted separately for those regions more accessible in SF and PB. Gene annotation of the DORs was performed by physical proximity, as detailed in Chapter ???. Despite commonalities, differences in significant enriched pathways (FDR<0.01 or 0.05) were identified within the same cell type between SF and PB (Figure 1.8). In CD14⁺ monocytes open chromatin in SF presented enrichment for pathways involved in regulation of immunity, inflammation and cell survival such as the NF-B pathway or cytokine production, including IL-2 signalling and IL-3, 5 and granulocyte-macrophage colonystimulating factor (GM-CSF) pathways (Figure 1.8 a). mCD4⁺ DORs more open in SF compared to PB showed enrichment for T cell receptor signaling as well as chemokine signaling, which included DORs in proximity to IFN- γ , IL-2 receptor alpha (*IL2RA*) and IL-5 receptor alpha (*IL5RA*), amongst others (Figure 1.8 b). This is consistent with the IL-2, IL-3 and IL-5pathway enrichment in open CD14⁺ SF DORs. Although T cell signaling pathway appears only enriched for open SF DORS in mCD4⁺, PB open regions in this cell type were enriched for focal adhesion members, also involved in the TCR signaling (XY). Enriched pathways for open mCD8⁺ were

Cross-tissue comparison analysis in PsA

only significant when using an FDR< 0.5 threshold. (Figure 1.8 c). Interestingly, G protein coupled receptor (GPCR) signalling was enriched for mCD8⁺ open PB DORs, consistently with the role of this pathway in chemotaxis. mCD8⁺ DORs open in SF also showed enrichment for the Wnt signaling pathway involved in the production of memory cells with enhanced proliferative potential and stronger protective capacity.

NK DORs open in SF presented enrichment for Fc-gamma receptor(FC γ R)-mediated phagocytosis that could be triggered by occurrence of monoclonal gammopathy of undetermined significance (MGUS) in PsA patients and consequently induce NK activation (Figure 1.8 d). Moreover, members of the HIF-1 pathway were also enriched in NK open SF DORs, as hypoxic environment is induced in joint inflammation. Interestingly, enrichment of PB open DORs in the proximity of NK cell mediated toxicity genes was unveiled. According to FACS analysis, the proportion of NK CD56^{bright} was greater in PB compared to SF in this sample cohort. This is consistent with the observed enrichment for NK cytotoxicity, since previous studies have demonstrated that CD56^{bright} NK cells are preferentially cytokine producers than the cytotoxic active than the tissue resident ones.

Cross-tissue comparison analysis in PsA

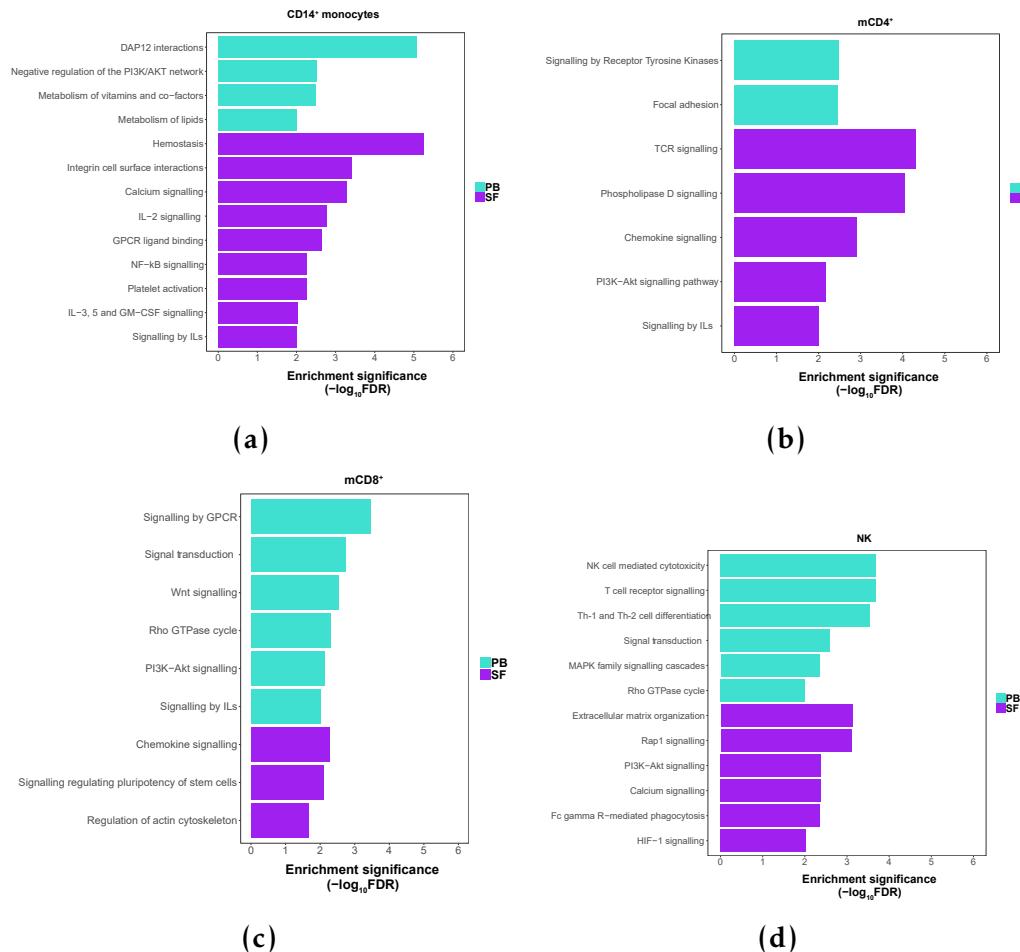


Figure 1.8: Distinct enriched pathways across SF and PB in CD14⁺, CD4m⁺, CD8m⁺ and NK. All pathways shown have an FDR <0.01.

In addition to pathway level analysis, individual enrichment analysis for conserved TFBS within SF and PB open DORs in each cell type was conducted. To further enhance the reliability of the analysis, only DORs overlapping FANTOM5 eRNA expressed in that particular cell type were considered.

-TFBS

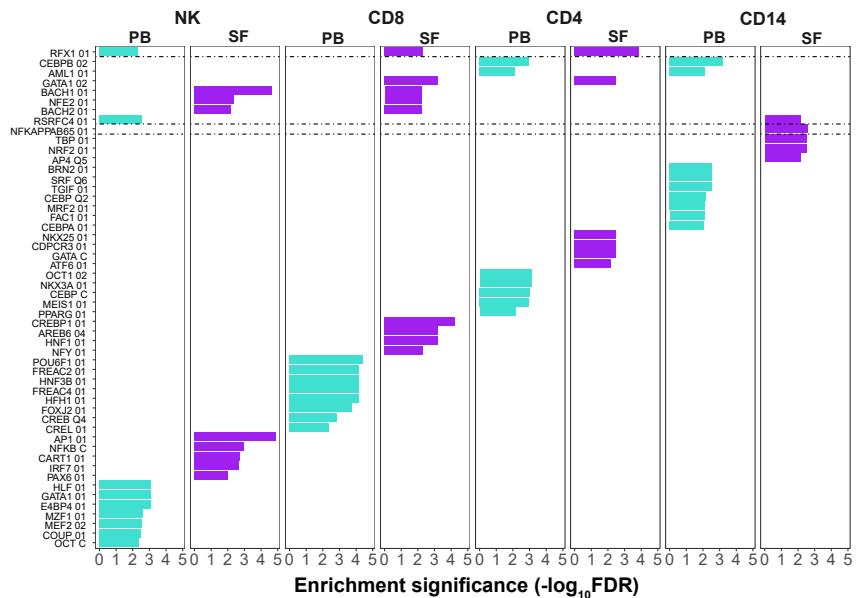


Figure 1.9: Enrichment of eRNA cell type-specific PsA DORs for conserved TFBS. xxxx

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

Immune-relevant gene expression and correlation with chromatin accessibility

Mapping chromatin accessibility represents an informative tool to identify regulatory elements undergoing histone modifications, DNA methylation and TF binding, as previously explained. All those elements are involved in the regulation of gene expression, making the study of chromatin accessibility a good proxy for the inference of genes expression. Nevertheless, the characterisation of the chromatin landscape also presents some limitations, including the discordance between open chromatin and functionality of the regulatory element, shown by CAGE studies, as well as the identification of the target gene regulated by a particular element.

In order to contextualise the ATAC-seq data, PCR gene expression analysis for 370 key genes in the inflammatory and autoimmune response was conducted in CD14⁺ monocytes, mCD4⁺ and mCD8⁺ cells isolated from SF and PB of three

Cross-tissue comparison analysis in PsA

PsA patients (Table 1.2). Those appeared as the most abundant cell types in PB and SF from patients and particularly for mCD8⁺ cells have been shown to expand in PsA inflamed synovium, as previously mentioned. The PCR array represented a cost-effective approach to study gene expression between PB and SF focusing in a relevant subset of genes of notably importance, given the pathophysiological characteristics of PsA. In each cell types, fold-change (FC) in expression was calculated pair-wise for SF in respect to PB within each sample for each of the tested genes. Likely due to the small sample size, most of the modulated genes between SF and PB lacked of significance after multiple testing correction (FDR<0.05). Therefore, to explore biological relevance of this data, pval<0.05 was used as the filtering threshold.

When exploring for each individual and each cell type, differences in magnitude and reproducibility in FCs were observed across samples and cell types (Figure ??). Some of the modulated genes showed upregulation ($\log_2\text{FC}>0$) in SF compared to PB across the three cell types, for example *IFN1*, *SPP1* or *CCL2*, amongst others (Figure ?? orange). On the other hand, a number of genes presented reduced expression in SF ($\log_2\text{FC}<0$) in at least one of the three cell types, including *FOS*, *IL16*, *PPBP* and *TPST1* (Figure ?? purple). Also, a number of only seemed consistently modulated in CD14⁺ monocytes but not in T cells (Figure ?? dark blue). For example, *CCR7* and *IL7R* were upregulated in SF CD14⁺ monocytes compared to PB; however the FCs between SF and PB were largely variable across the three patients in mCD4⁺ and mCD8⁺. Moreover, differences in the magnitude of FCs across the three cell types were observed for some of the genes modulated in the same direction across the three cell types, for instance *VEGFB* and *CXCR6* (Figure ?? black and green).

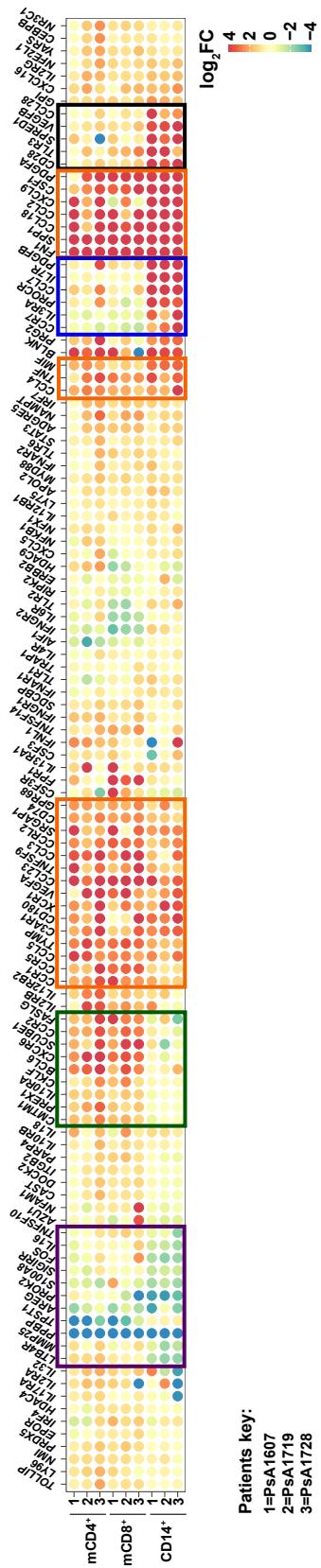


Figure 1.10: Representation of the fold-change in gene expression between SF and PB for the significant genes ($pval < 0.05$) in at least one of the cell types. xxxx

Cross-tissue comparison analysis in PsA

Further filtering based on statistical significance ($pval<0.05$) and FCs ($0.6<FC<1.5$), revealed that $CD14^+$ monocytes and $mCD8^+$ presented greater number of significantly modulated (70 and 73 genes, respectively) compared to $mCD4^+$ cells (46 genes) (Figure ?? a, b and c). For the three analysed cell types, the majority of modulated immune genes presented upregulation in the SF (Figure ?? a, b and c). For example, 56 out of the 70 significantly modulated genes in $CD14^+$ monocytes showed a \log_2 mean $FC>0.586$ versus the 14 genes with \log_2 mean $FC<-0.586$ (Figure ?? a).

Amongst the significantly modulated genes (based on $pval$ and FC cut-offs) between SF and PB, overlap with DORs were observed for the three cell types (Table 1.5). In $CD14^+$ monocytes, 13 out of the 56 significantly upregulated genes in SF overlapped with DORs more accessible in this tissue.

Cell type	Genes upregulated and overlapping open chromatin in SF	Genes downregulated and overlapping closed chromatin in SF
$CD14^+$	13 (<i>BLNK, CCL2*, CCR1*, CD180, CXCL10, FN1, IL18, IL31RA*, IL7R*, NFKB1*, PRG2, SRGAP1, STAT3</i>)	2 (<i>FOS, PROK2*</i>)
$CD4^+$	3 (<i>CXCL13, CXCR6*</i>), <i>IL2RA</i>	0
$CD8^+$	6 (<i>CCL3, CCR2, CCR5 ,IRF4 TNFSF10, YARS</i>)	1 (<i>EPHX2</i>)

Table 1.5: Immune genes with significant modulated expression in SF proximal to a DOR in ATAC-seq. An overlap is defined by significant change in expression ($pval<0.05$) of a particular gene where there is also a proximal DOR showing changes in chromatin accessibility in the same direction. (*) indicates that the proximal DOR overlapping an eRNA identified by FANTOM5 project in that particular cell type (see subsection Characterisation of the differential accessible chromatin regions).

For example, the greater chromatin accessibility at the *IL7R* 5' and 3' UTR, previously shown (Figure 1.7 b) correlated with greater mRNA expression in SF $CD14^+$ monocytes. Another relevant example is the *FN1*, involved in cell adhesion, migration and osteoblast biology. Upregulated expression in synovial

Cross-tissue comparison analysis in PsA

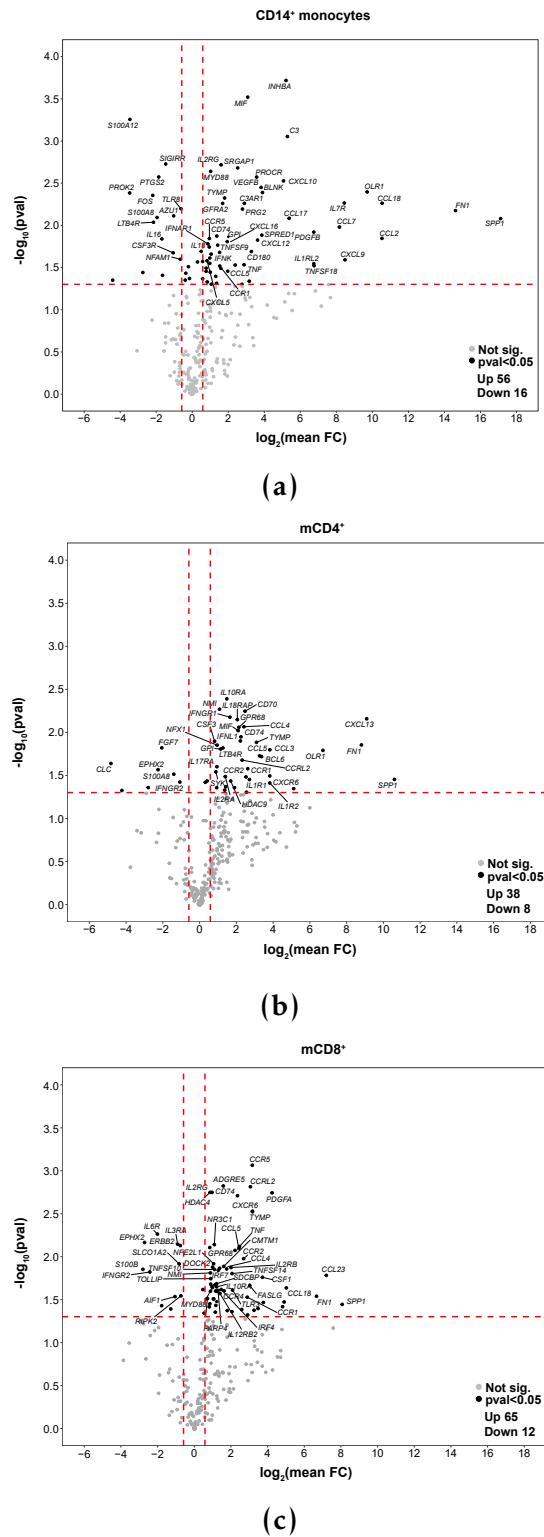


Figure 1.11: Expression changes in immune-relevant genes between SF and PB in CD14⁺ monocytes, mCD4⁺ and mCD8⁺ cells.

Cross-tissue comparison analysis in PsA

biopsies compared to PB has already been reported (**Dolcino2015**). In this cohort, *FN1* expression was upregulated in SF for all three cell types, with the greater FC found in CD14⁺ monocytes (Figure ?? a), concomitantly with more accessible chromatin at the promoter and 3' UTR of the gene (Figure ??). Lesser overlap between gene expression and chromatin accessibility was observed in mCD4⁺ and mCD8⁺ (6 and 3 targets, respectively) between genes with greater expression and accessible chromatin in SF compared to PB. Only significantly downregulated genes in CD14⁺ monocytes and mCD8⁺ isolated from SF presented overlap with proximal less accessible chromatin regions (2 and 1, respectively) in the same tissue.

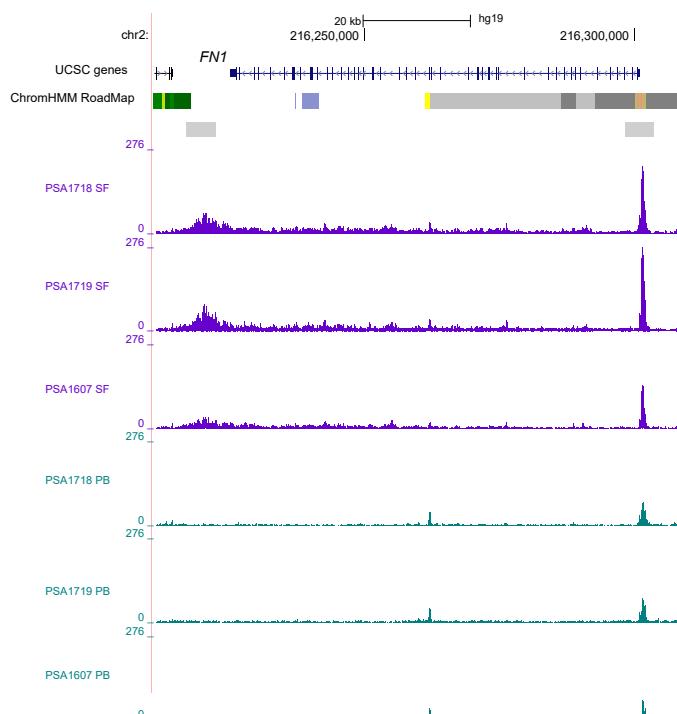


Figure 1.12: Chromatin accessibility at the *FN1* gene in CD14⁺ monocytes isolated from SF and PB. xxxx

Pathway enrichment and network analysis highlights the role of synovial CD14⁺ monocytes in cytokine and chemokine production

Modulated gene expression in CD14⁺ monocytes shows enrichment for cytokine and chemokine production

To identify relevant pathways amongst the modulated genes between SF and PB, enrichment analysis was performed for each of the three cell types. Upregulated and downregulated genes showing $0.6 < \text{mean FC} > 1.5$ and $\text{pval} < 0.05$ were combined for the analysis. Interestingly, the modulated genes between SF and PB in CD14⁺ monocytes were enriched for chemokine, NOD-like signalling and TLR signalling pathways (Table 1.6). All three pathways are involved in the activation of cytokines and chemokines gene expression, leading to T cell recruitment and inflammatory response.

The TLR signalling pathways enrichment involved *FN1* (previously mentioned) and *SPP1*, two of the top three most differentially expressed genes reported by Dolcino and colleagues in a study comparing synovial biopsies between healthy and PsA individuals (**Dolcino2015**) (Table 1.6). Together with *FN1*, *SPP1* was also highly upregulated (mean $\log_2 \text{FC} \geq 4$) in the three cell types (Figure ?? orange), showing the greatest FC in monocytes (Figure ?? a). Moreover, some of the genes driving enrichment, such as *CCL5* and *NFKB*, were also shared across the three pathways of interest. Others genes, including *TNF*, *IRF7* and *MyD88*, highlighted the cross-link between the NOD-like and the TLR signalling pathways. The enrichment of SF open chromatin regions in CD14⁺ monocytes for the NFB pathway is closely related to the enrichment for TLR and NOD-like signalling pathways at the transcriptomic level (Figure 1.8 a). Both, TLR and NOD-like pathways lead to the activation of the NFB TF, which induces transcriptional activation of pro-inflammatory cytokines, also supported by the enrichment of SF more accessible DORs for IL-2, IL-3, IL-5 and GM-CSF pathways (Figure 1.8 a). Moreover, the pivotal role of NFκB in the immune

Cross-tissue comparison analysis in PsA

transcriptional profile of SF CD14⁺ monocytes is additionally sustained at the chromatin accessibility level by the enrichment of SF accessible chromatin sites for this TF (Figure ??).

The enrichment for the chemokine pathway in CD14⁺ monocytes (Table 1.6) included modulated genes highly upregulated (mean log₂FC≥4) in SF compared to PB (e.g *CCL18* and *CCL2*) for all three cell types (Figure ?? orange) as well as genes only consistently modulated between SF and PB in CD14⁺ monocytes (e.g *CCL28*, Figure ?? black). This pathway involves production of chemottractant molecules involved in the recruitment of leukocytes to the site of inflammation and the production of reactive oxygen species (ROS) through Ca²⁺ mobilisation, a pathway presenting enrichment for more accessible chromatin regions in SF CD14⁺ monocytes (Figure 1.8 a).

At the transcriptional level, significantly modulated genes between SF and PB in mCD4⁺ T cells were enriched for the IL-10 signalling pathway (Table 1.6), in lines with the enrichment for IL signalling of open chromatin in SF cells (Figure 1.8 b).

Cross-tissue comparison analysis in PsA

Table 1.6: Pathway enrichment analysis for the modulated genes between SF and PB in CD14⁺ and mCD4⁺. The analysis was performed using only those genes showing $0.5 \leq \text{mean fold-changes} \leq 1.5$ across the three patients and $\text{pval} < 0.05$. Enriched pathways based on FDR < 0.05 .

Cell type	Pathway	Genes
CD14 ⁺	Chemokine signalling	CCL17, CCL18, CCL2, CCL28, CCL5, CCR1, CCR5, CXCL10 CXCL12, CXCL16, CXCL5, CXCL9, NFKB1, PPBP, PF4V1, STAT3, XCR1
	NOD-like receptor signalling	CCL2, CCL5, IFNAR1, IL18, IRF7, MEFV, MYD88, NFKB1, NAMPT, TNF
mCD4 ⁺	TLR signalling	CCL5, CXCL10, CXCL9, IFNAR1, IRF7, MYD88, NFKB1, SPP1, FOS, TLR1, TLR2, TLR8, TNF
	IL-10 signaling	CCL3, CCL4, CCL5, CCR1, CCR2, CSF1, CSF3, IL10RA, IL1R1, IL1R2

Cross-tissue comparison analysis in PsA

In addition to pathway enrichment, network analysis was performed to understand the interaction and relationship between the analysed genes. A gene subnetwork was identified from the STRING functional interaction database using as input the filtered expression PCR array information and ranking the genes based on the best pval across the three cell types. The identified subnetwork predominantly included significant modulated genes between SF and PB. Amongst the most interesting node, the single Ig and Toll-interleukine domain containing gene (*SIGIRR*) which is a negative regulator of the TLR signalling pathway (Figure ??). *SIGIRR* is significantly downregulated in SF CD14⁺ monocytes only and it is consistent with the significant upregulation (pval<0.05) of *TLR1*, *TLR2*, *MyD88* and the enrichment for the TLR pathway in this cell type (Figure ?? and 1.6). Moreover, the significant upregulation of *NFKB* and *TNF* in the SF CD14⁺ monocytes could be a downstream result of the functional connection with TLR pathway members such as *MyD88* and as a downstream result of the TLR signalling pathway. Conversely, in mCD4⁺ and mCD8⁺ the modulation of these members did not appear to be significant between SF and PB. However, in mCD8⁺ *TNF* is also significantly upregulated in the synovium.

Another interesting part of the network is the connection of the TLR pathway and the chemokine production through *NFKB*, *TNF* and *CCL2* (Figure ??). *CCL2* is connected to *CXCL10* and subsequently with *CCL18* and *CCR5*, all chemokines regulating migration and infiltration of monocytes and memory T cells. This network analysis also highlighted relationship between *IL7R* and *IL2RG* coding for the two chains of the IL-7 receptor (IL-7R). Interestingly, these two nodes are only significantly upregulated in SF CD14⁺ monocytes, supporting a novel the cell and context specific role of IL-7R and polymorphism in this gene in this cell type under inflammatory conditions

Cross-tissue comparison analysis in PsA

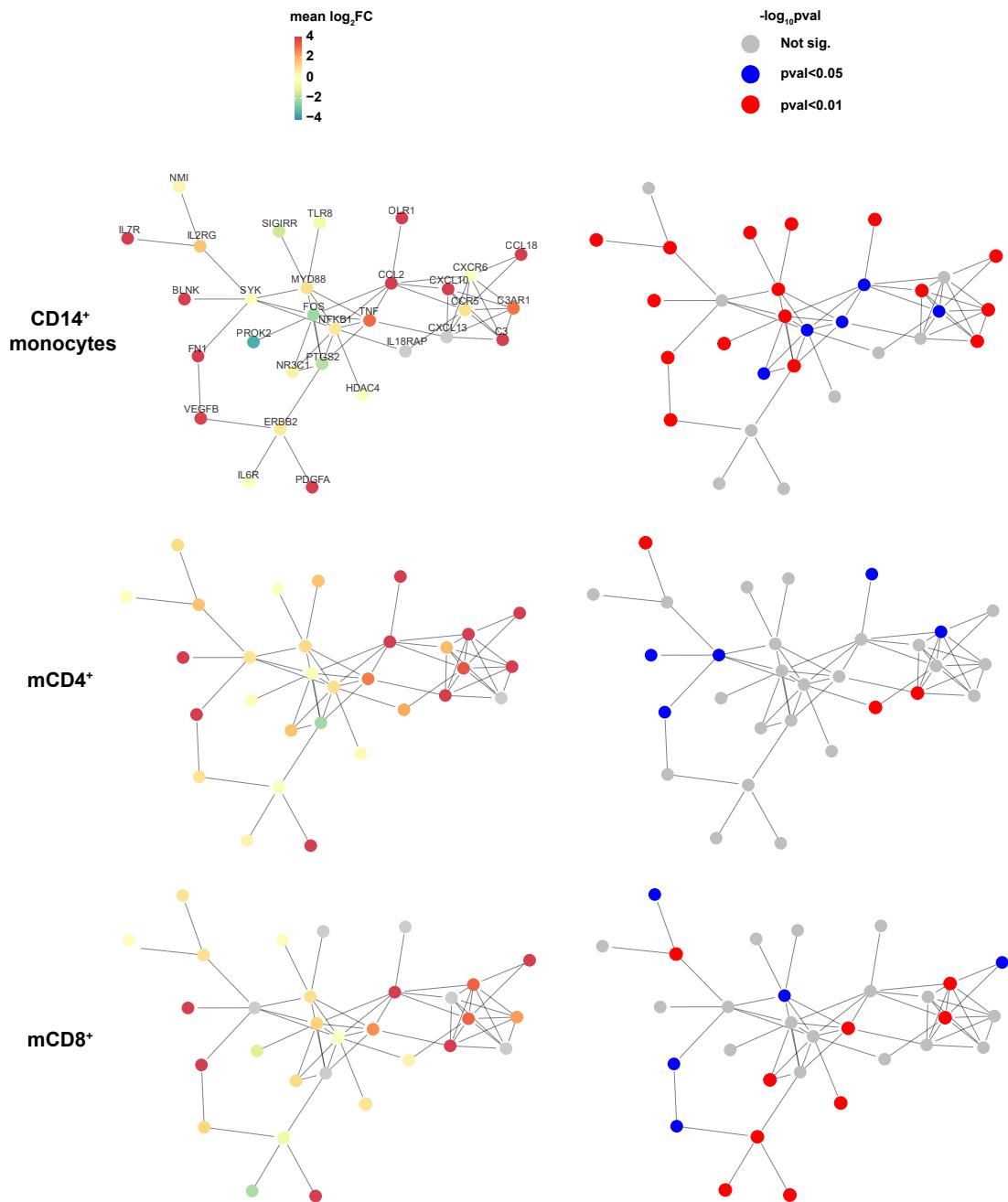


Figure 1.13: Protein network analysis based on the immune PCR array expression data. xxxx

Cross-tissue comparison analysis in PsA

Overall, the integration of the chromatin accessibility and immune transcriptional data reinforced a relevant role of synovial CD14⁺ monocytes in the production of cytokines and chemokines, leading to activation of the innate immune response and the recruitment of T cells to this site of inflammation.

Tissue and disease specificity in gene expression modulation and relevant biological pathways

In order to better understand the disease and tissue specificity of the prior transcriptomic results, gene expression was analysed in CD14⁺ monocytes, mCD4⁺ and mCD8⁺ isolated from PB in three healthy controls (HC) using the same PCR array. In each of the cell types, the FC in was calculated for the PsA PB expression in respect to the healthy controls. Similarly to the previous analysis, pvals were calculated for the FC in each of the particular genes. When comparing to the modulated gene expression between SF and PB in PsA, three group of genes could be identified (Figure ??). The genes only significantly modulated (based on pval and FC threshold criteria) in PB between HC and PsA were designated as systemic genes (Figure ?? green). Those genes were not significantly modulated in this data when comparing SF versus PB within PsA patients and could then be considered as the circulating disease "footprint". In this respect, CD14⁺ monocytes was the cell type with lower number of systemic modulated genes (14), compared to mCD8⁺ (23) and mCD4⁺ (42) (Figure ??a, b and c).

Another group of genes were designated as tissue specific, since they were significantly modulated between SF and PB in PsA patients but did not show significant changes between HC and PsA at the circulating level (Figure ?? red). Interestingly, in CD14⁺ monocytes the tissue specific modulated genes considerably outnumber the systemic ones (62 versus 14), showing a more pronounced change in the expression profile of immune genes across patients tissues than between healthy and diseased PB (Figure ?? a in red). For example,

Cross-tissue comparison analysis in PsA

the aforementioned *NFKB* and *MyD88*, *TLR2* genes were only upregulated in PsA SF CD14⁺ monocytes and their expression was not significantly modulated between HC and PsA circulating CD14⁺ monocytes. Similarly to CD14⁺ monocytes, mCD8⁺ cells also presented greater disease tissue-specific modulation than genes modulated compared to HC in PB (Figure ?? c in red).

The third category of genes were those significantly modulated for each cell type between HC and PsA patients in PB as well as between SF and PB from PsA individuals. These were defined as putative disease-specific genes and numbers were similar across CD14⁺ monocytes, mCD4⁺ and mCD8⁺ (10, 9 and 8, respectively) (Figure ?? blue in a, b and c). In CD14⁺ monocytes two of those genes, *GPI* and *PRG2*, were upregulated in both comparisons with further exacerbation in SF when comparing the log₂FCs (Figure ?? a). Evidence of the glucose-6-phosphate isomerase *GPI* upregulation has been found in RA synovial fibroblast and linked to increased levels of TNF- α and IL-1 β in the synovium (Zhong2015). Another example of exacerbated upregulation in SF was the expression of *GPR68* in mCD4⁺. This gene was upregulated in PsA PB mCD4⁺ when compared to the HC counterparts and further upregulated in SF when compared to PB in PsA individuals (Figure ?? b). *GPR68* is a G protein-coupled receptor, expressed in T cells , amongst others, which gets activated through pH acidification, particularly relevant in synovial tissues, increasing Ca²⁺ levels and leading to activation of immune pathways. *GPR68* was also upregulated in SF compared to PB in mCD8⁺ cells, reinforcing the relevance of this gene in the synovial pathophysiological aspect of PsA. An opposite example is the epidermal growth factor-like amphiregulin (*AREG*) gene which is significantly upregulated in PsA PB mCD8⁺ cells compared to the HV but is downregulated in SF compared to PB in mCD8⁺ from PsA individuals (Figure ?? c). Lack of *AREG* in mouse models have been shown an impaired immunosuppressive response by Treg cells (Zaiss2013).

Cross-tissue comparison analysis in PsA

Despite the interesting findings, the identification of disease-specific and disease tissue-specific genes is clearly limited by the lack of HC SF in the experimental design, partly due to invasive sampling procedure, and further commented in the Discussion.

When performing pathway enrichment analysis using the significantly modulated genes between healthy and PsA patients PB, none specific pathway enrichment beyond Reactome immune system were identified only for CD14⁺ monocytes and mCD4⁺. This result reinforced the tissue-specificity of the pathways enriched for the modulated genes between SF and PB in CD14⁺ monocytes PsA patients and clearly suggest a more pronounced inflammatory phenotype of the pathological CD14⁺ monocytes in SF compared to PB.

1.2.6 Mass cytometry reveals CD14⁺ monocytes as the most active cytokine producing cells

Measurement of cytokine production using mass cytometry was conducted in collaboration with Dr Hussein Al-Mossawi and Dr Nicole Yager. Cytokine production by CD14⁺ monocytes, mCD4⁺ T cells, mCD8⁺ T cells and NK was quantified in SF and PB after incubation with BFA for 6 h. BFA blocks cytokine release and enables to measure the cell cytokine production rate in absence of any inflammatory stimuli, as previously explained in Chapter ??(Figure ??). Following data QC, IL-2, IL-8, IFN- γ and TNF- α were the cytokines consistently released with the most reliable quantification results within each of the manual gated cell type clusters. CD14⁺ monocytes isolated from both, SF and PB, was the only active cell type in producing IFN- γ , IL-2 and IL-8. Nevertheless, no significant differences in the production (mean) of those three cytokines was observed in CD14⁺ monocytes between SF and PB. On the other hand, only CD14⁺ monocytes isolated from SF showed positive detectable mean levels of TNF- α , which was undetectable in the PB counterpart cells.

Cross-tissue comparison analysis in PsA

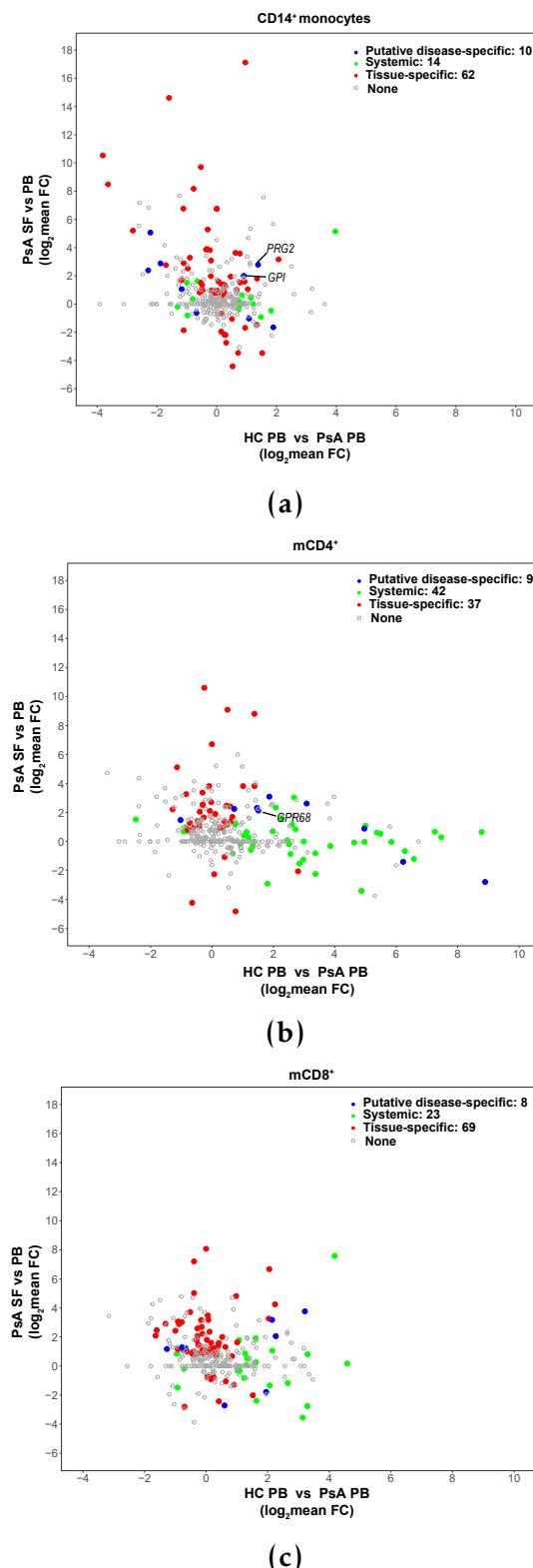


Figure 1.14: Expression changes in immune-relevant genes between SF and PB in CD14⁺ monocytes, mCD4⁺ and mCD8⁺ cells.

Cross-tissue comparison analysis in PsA

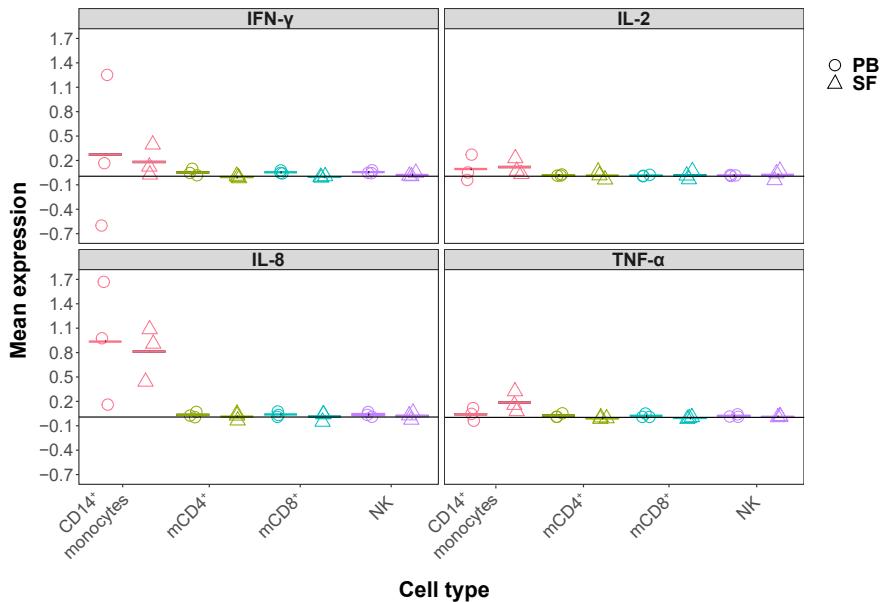


Figure 1.15: Mean expression of IL-2, IL-8, TNF- α and IFN- γ in SF and PB from PsA patients. Expression levels quantified by mass cytometry. xxxx

The increased rate of TNF- α production by SF CD14 $^{+}$ monocytes was reinforced when comparing the proportion of TNF- α positive cells after 6h of blocked cytokine secretion in each of the two tissues. The mass cytometry analysis of CD14 $^{+}$ versus TNF- α signal showed a greater number of SF CD14 $^{+}$ monocytes producing TNF α after 6 h of BFA treatment compared to PB for each of the three analysed patients (Figure 1.16 a). The basal (0h) percentage of CD14 $^{+}$ monocytes positive for TNF- α was negligible for all three patients in both, SF and PB (Figure 1.16 b). Conversely, upon inhibition of protein transport (6h), the percentage of CD14 $^{+}$ monocytes producing TNF- α experienced an increase, ranging between 2 and 11.8%, in SF whereas the increment in PB did not reach 1%. This trend of increased percentage of SF CD14 $^{+}$ monocytes TNF- α did not reach significance likely due to the small sample size.

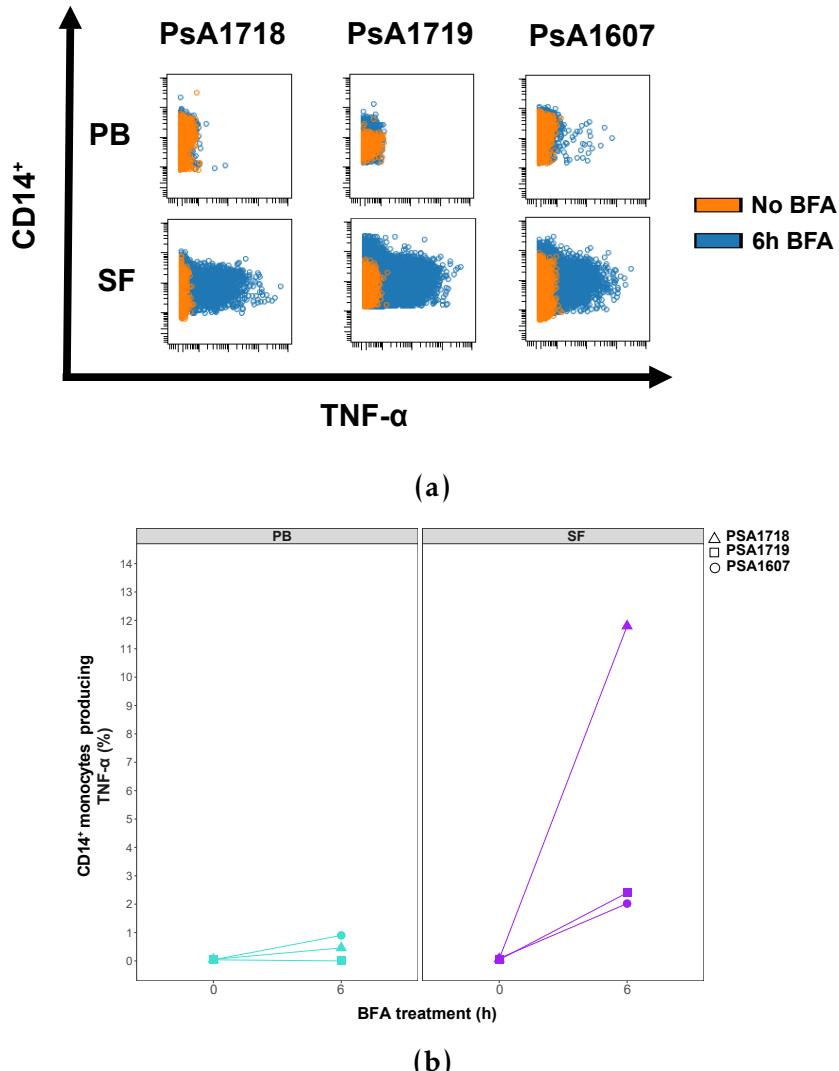


Figure 1.16: Differentially accessible regions located within gene bodies in CD14⁺ monocytes and NK cells from PsA patients. xxxx

1.3 Discussion

In Dolcino PSA vs HV SF comparison also detected more up genes than down in the array analysis Say which cell types drive more the top changed genes fGWAS analysis as Matthias did would be of interest but needs appropriate GWAS data I am going to try using XGR to do some of this

Appendices

A Establishment of methods to assess genome-wide chromatin accessibility	37
B Cross-tissue comparison analysis in PsA	40

List of Figures

List of Tables

Appendix A

Establishment of methods to assess genome-wide chromatin accessibility

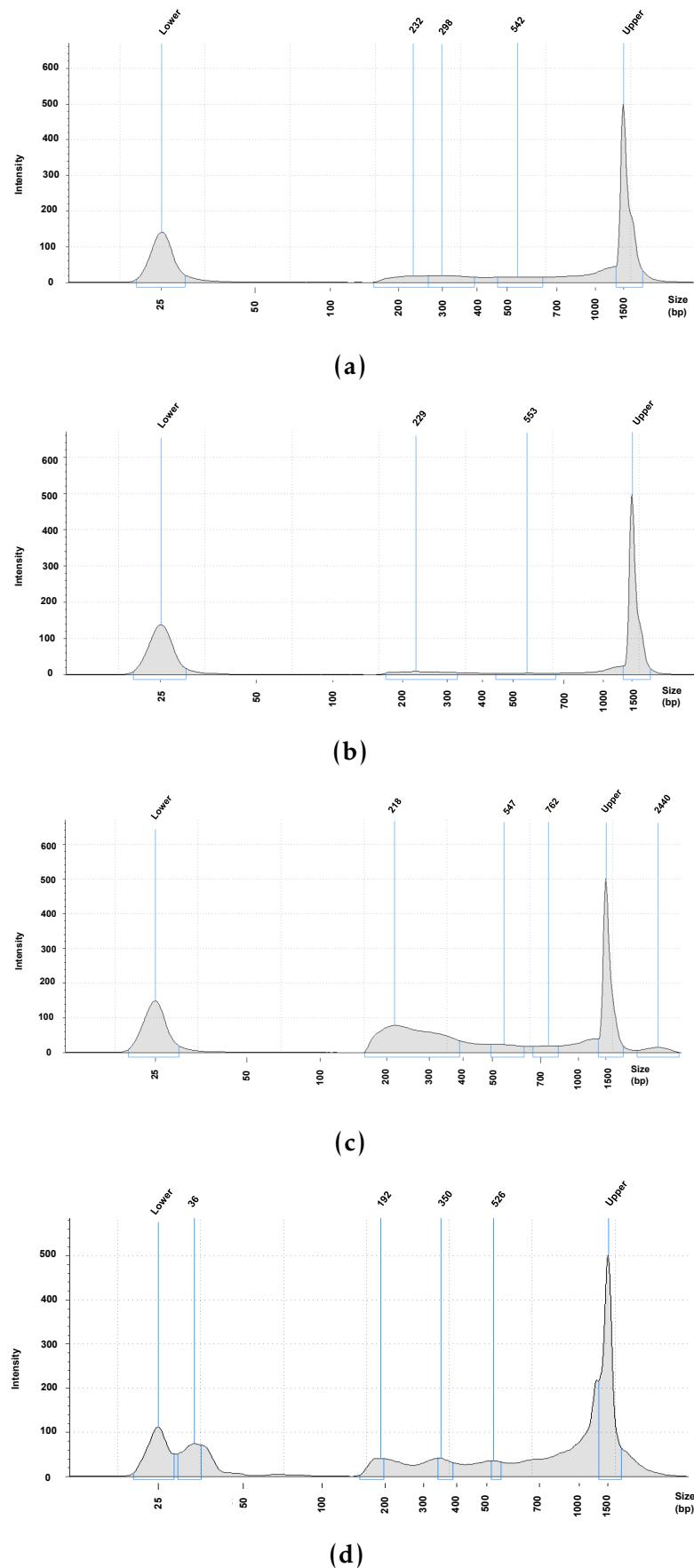


Figure A.1: FAST-ATAC and Omni-ATAC NHEK tapestation profiles.

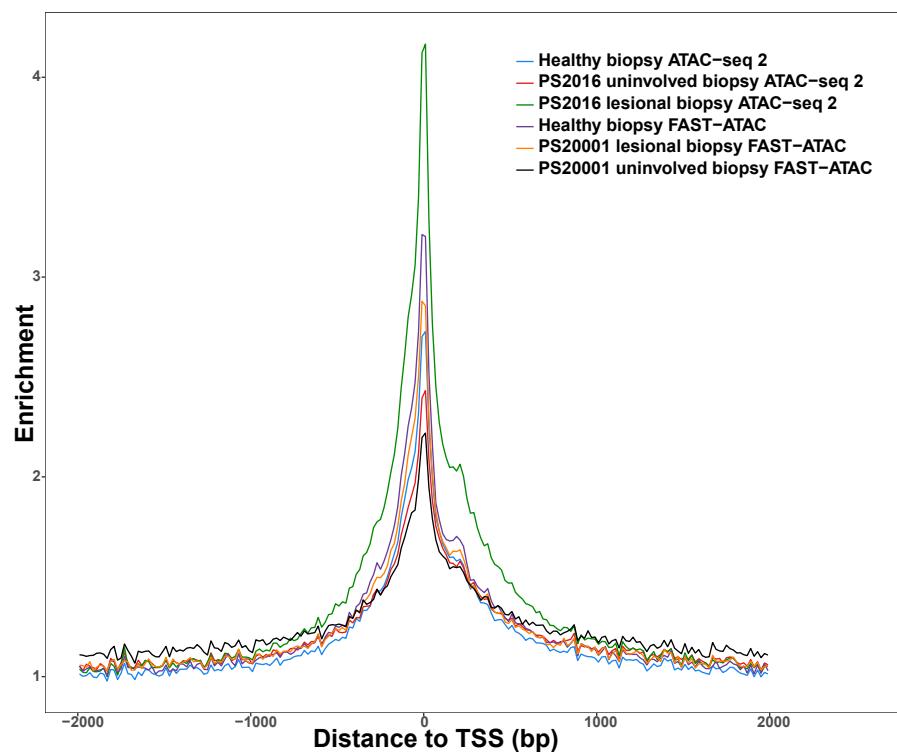


Figure A.2: Assessment of TSS enrichment from ATAC-seq and FAST-ATAC in healthy and psoriasis skin biopsies samples.

Appendix B

Cross-tissue comparison analysis in PsA

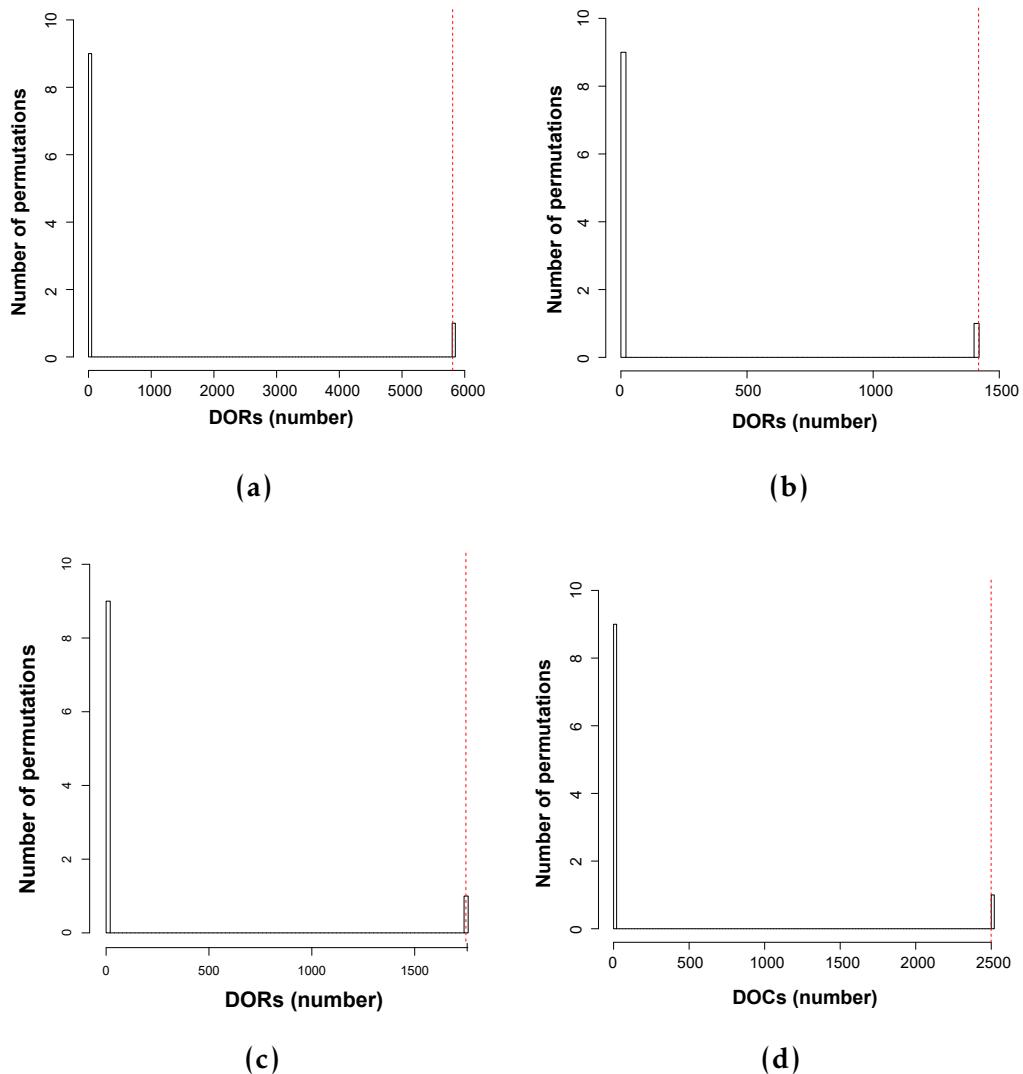


Figure B.1: Permutation analysis SF vs PB in CD14⁺,CD4m⁺,CD8m⁺ and NK.