\section{Discussion}

\subsection{Characterising chromatin accessibility in PsA samples}

Technological advances in the field of epigenetics has enabled characterization of the epigenetic landscape in immune cells isolated from PB and SF in PsA patients. The cell types included in this study (CD14$^+$ monocytes, mCD4$^+$, mCD8$^+$ and NK) represent key players in the innate and adaptive immune response dysregulated in PsA pathophysiology. In particular, expansion of mCD4$^+$ and mCD8$^+$ cells in the synovium of PsA patients has been described, and GWAS have highlighted significant association with MHC-I and other pivotal genes involved in T cell immune response. \parencite{Taams2018}.

I have used ATAC to characterise the chromatin landscape of four relevant immune cell types isolated from PB and SF of PsA patients, revealing genome-wide differences between the two tissues. CD14$^+$ monocytes demonstrated the largest number of DARs between SF and PB (23.3\% of the investigated regions). Across all four cell types, the majority of the DARs were located at intergenic and intronic regions annotated as enhancers (weak or strong), and also enriched for eRNAs identified in each of the four relevant cell types by the FANTOM project \parencite{ FANTOM2014}. Altogether these findings suggested a putative role of the differential regions in long-range regulation of gene expression \parencite{Qu2015}.

Pathway enrichment analysis of the genes placed proximal to the DARs revealed both commonalities and differences between regions, with increased accessibility in SF or PB for each of the cell types. In SF, enrichment for open DARs proximal to genes in the NF-$\kappa$B pathway was found in CD14$^+$ monocytes, consistent with the TNF-$\alpha$ signature of PsA and the efficacy of anti-TNF therapies already reported for PsA \parencite{ Ahil2016}. Interestingly, in mCD8$^+$ T cells enrichment of DARs open in PB was found proximal to genes in the Wnt signaling pathway, such as \textit{SMAD3}. Wnt signalling is involved in many biological processes and its dysregulation is associated with a number of autoimmune diseases, such as RA \parencite{Miao2013}. In this data, the significantly increased chromatin accessibility nearby Wnt signalling genes in PB mCD8$^+$ cells may be related to an increased recall proliferation capacity of the circulating cells compared to the tissue-resident, a phenomenon previously described \parencite{Boudousquie2014}. In NK cells, increased accessibility proximal to gene members of the NK-mediated cytotoxicity pathway was found in PB compared to SF cells. NK CD56$^{bright}$ cells resident at sites of inflammation are more specialised for cytokine production than cytotoxicity \parencite{Michel2016}. Notably, matched mass cytometry data in these patients have shown a reduced proportion of CD56$^{bright}$ cells in PB compared to SF which would be consistent with the enrichment of DARs open in PB for proximal to genes in the NK-mediated cytotoxicity pathway.

Overall, this approach has identified robust differences in chromatin accessibility between SF and PB in relevant immune cells isolated from PsA patients. Although these differences may suggest meaningful functional differences, a limitation of pathway analysis in ATAC is linking putative regulatory regions to genes. Annotation of ATAC peaks with genes in proximity has been widely used in the literature \parencite{Scharer2016,Ackermann2016, Corces2016,Wang2018}. However this approach fails to evaluate long range interactions, assuming accessible chromatin to regulate expression of the neighbouring regions. Moreover, accessible chromatin is not a definitive marker for regulatory regions and mapping of histone marks such as H3K4me1 and H3K27ac together with eRNA quantification will help to refine the functional relevance of the identified DARs.

\subsection{Biological insights in the integration of chromatin accessibility, gene expression and proteomics data in PsA}

\subsubsection{Chromatin accessibility landscape and bulk gene expression profiling in paired samples}

In contrast to chromatin accessibility profiling, recent research has been conducted to deepen the understanding ofdifferences in the transcriptional profile between PBMCs and synovial tissue from PsA patients; with some studies also incorporating synovial membrane biopsies from controls samples \parencite{Dolcino2015}. However, comparative analysis in matched discrete cell subpopulations from SF and PB from the same patient have yet to be reported.

In this chapter, I have presented a pilot study characterising the transcriptomic profile for relevant immune genes in CD14$^+$ monocytes, mCD4$^+$ and mCD8$^+$ cells isolated from SF and PB and for the first time, integrated this with paired chromatin accessibility data from the same samples. CD14$^+$ monocytes and mCD8$^+$ presented the largest number of consistently modulated genes between SF and PB in this pilot analysis. The top most significantly dysregulated genes between SF and PB in the four cell types were \texit{SPP1} and \textit{FN1}, the same as the reported by Dolcino \textit{et al.}. Other highly DEGs in Dolcino \textit{et al.} study, including \textit{TNFA}, \textit{CXCL13} or \textit{CCL18}, were found also to be modulated between SF and PB in at least one of the cell types in this pilot data. Consistently with their role in Th17 cell biology, \textit{CXCL13} and \textit{IL26} only appeared significantly up-regulated in SF mCD4$^+$ and/or mCD8$^+$ cells but not in CD14$^+$ monocytes \parencite{Takagi2008}.

The subsequent integration of transcriptome profiling with paired-ATAC data in CD14$^+$ monocytes, mCD4$^+$ and mCD8$^+$ revealed that genes with differentially modulated expression in SF and PB corresponded with nearby DARs showing changes in the same direction. Some of those DARs had also been identified as eRNAs in CD14$^+$ monocytes (e.g \textit{NFKB1}) and mCD4$^+$ cells (e.g \textit{CCR6}) by the FANTOM project, which may suggest a downstream role of those regions in active transcriptional regulation. Furthermore, exploratory integration of differences in gene expression in PB between PsA patients and controls with the cross-tissue comparison within matched PsA samples from SF and PB led to identification of systemic, tissue-specific and putative-disease specific modulated genes according to this pilot data.

CD14$^+$ monocytes and mCD8$^+$ presented a larger number of genes modulated between SF and PB in PsA patients only (tissue-specific genes) than genes changing expression only between PB of PsA versus control samples (systemic genes). Interestingly, \textit{CCR10} gene was one of the systemic genes for mCD8$^+$ cells. \textit{CCR10} is a chemokine receptor co-expressed by subset of memory cell which preferentially migrate to skin, and has also been identified as an up-regulated gene in tCD8$^+$ cells in the psoriasis cohort (Chapter \ref{ch:Results1} and in patients with atopic dermatitis \parencite{Hijnen2005}. Interestingly, \textit{SPP1} and \textit{FN1}, the top two modulated genes between SF and PB in all cell types, appeared in the tissue-specific category in the three cell types. These two genes were reported by Dolcino and colleagues as the two most dysregulated when comparing bulk synovial membrane transcriptome from PsA patients and healthy individuals \parencite{Dolcino2015}. My data suggests that CD14$^+$ monocytes, mCD4$^+$ and mCD8$^+$ cells may all be contributing to the up-regulation of \textit{SPP1} and \textit{FN1} in the PsA synovium membranes found by Dolcino and colleagues. The \textit{SPP1} protein product, osteopontin, is a cytokine and chemokine expressed by many cell types, including monocytes/macrophages and T cells. It is involved in cell migration, adhesion and cell-mediated immune response through regulation of T cells, importantly in Th-17 \parencite{Morimoto2010}. \textit{SPP1} has also been described to play a role in other chronic inflammatory and autoimmune diseases \parencite{Rittling2015}. On the other hand, \textit{FN1} is the gene coding for fibronectin-1, a main component of the cartilage matrix, involved in cell adhesion, migration, growth and differentiation and found to be highly expressed in RA inflamed synovium \parencite{Chang2005}. Moreover, \textit{FN1} has been shown to induce bone resorption mediated by pro-inflammatory mediators, such as nitric oxyde and IL-1$\beta$ \parencite{Gramoun2010}. In Dolcino’s data \textit{FN1} up-regulation was only found when comparing synovial membranes of PsA versus controls, but no changes were reported between PsA and healthy controls PB samples, altogether suggesting the tissue-specificity of this dysregulation in PsA. Furthermore, the identification of DARs open in SF at the promoter and 3’ downstream of \textit{FN1} in CD14$^+$ monocytes suggest a link between changes in modulation of gene expression and chromatin accessibility in this particular cell type. In contrast to the observation in my pilot study, Dolcino also showed moderate up-regulation of \textit{SPP1} expression in PB from PsA patients compared to controls. This could be explained by the fact that Dolcino's study was performed in bulk PBMCs and that dysregulation of \textit{SPP1} in CD14$^+$ monocytes, mCD4$^+$ and mCD8$^+$ cells may be tissue-specific. However, this could also be the result of the small number of qPCR replicates and \textit{SPP1} dysregulation in PB between PsA and healthy samples failing to reach significance in my study. Overall, up-regulation of these two genes in SF cells reflect the activation of chemotaxis and infiltration of circulating monocytes and T cells, activation of the Th-17 immune response and dysregulation of osteoclast bone remodeling , all of which are pivotal in PsA pathophysiology, particularly at the inflamed tissue \parencite{Durham2015, Mensah2017}.

An example of an interesting putative-disease specific gene identified by this study was \textit{GPR68}, shown to be up-regulated in mCD4$^+$ PsA PB compared to controls, with expression further increased in SF. \textit{GPR68} is a G protein-coupled receptor (GPCR), expressed in T cells, amongst other cells, that undergoes activation through pH acidification, characteristic of synovial tissues under inflammation \parencite{Biniecka2016, Saxena2011}. Interestingly, \textit{GPR65}, another member of the acid-sensing GPCR family, has been associated with a number of immune mediated diseases, including AS, CD and MS \parencite{Cortes2013,Lassen2016,Wirasinha2018}. GPR65 was found to be a marker of pathogenic Th17 cells in the murine and human systems \parencite{Gaublomme2015, Al-Mossawi 2017}. Unfortunately, GPR65 was not included in the gene array used in this study. Indeed, the use of a gene array rather than a transcriptomics approach using RNA sequencing is a limitation of this study.

Pathway analysis using consistently modulated genes between SF and PB in this data revealed enrichment for relevant pathways in CD14$^+$ monocytes and mCD4$^+$ cells. In CD14$^+$ monocytes, TLR and NOD-like signalling pathways were enriched, consistently with the relevance of TLR and NOD-like receptors for rheumatic diseases \parencite{McCormack2009}. Up-regulated expression of \textit{TLR1} and \textit{TLR2} presented significant differential modulation in CD14$^+$ monocytes in SF; however a similar trend in \textit{TLR4} expression failed to reach significance in this pilot study. This is consistent with some studies that have identified increased \textit{TLR-2} and \textit{TLR-4} expression in SFMCs compared to PBMCs in patients with juvenile idiopathic arthritis \parencite{Myles2011}. Moreover, the transcriptomic analysis in lesional and uninvolved psoriatic skin presented in Chapter \ref{ch:Results2} also highlighted the role of NOD-like signalling in the inflammatory response at the site of inflammation. The cross-talk between the TLR and NOD-like signalling pathways, a well-known phenomenon in inflammation, was further evidenced by network-based analysis in this data. TLR and NOD-like signalling pathways lead to activation of the NF$\kappa$B TF, which appeared transcriptionally up-regulated in SF CD14$^+$ monocytes in this data, and transcription of pro-inflammatory cytokines and chemokines \parencite{Mitchell2010}. In my data, enrichment of DARs open in SF in the proximity of genes within the NF-$\kappa$B pathway was found and further supported at the transcriptional level by the consistent up-regulation of \textit{TNFA}, textit{CCL2} and \textit{CCL5} expression in SF CD14$^+$ monocytes. Moreover, analysis for conserved TFBS motifs in the ATAC data revealed enrichment for NF-$\kappa$B binding motifs within DARs open in SF CD14$^+$ monocytes (data not shown).

Interestingly, pathway analysis between SF and PB mCD4$^+$ demonstrated an enrichment for significant gene modulation in IL-10 signalling. Flow cytometry data in matched PBMCs and SFMCs from the same three patients revealed an expansion of Tregs in SF compared to PB (data not shown), characterised by the expression of anti-inflammatory cytokines, including IL-10 \parencite{O’Garra2004}. qPCR analysis showed a significant increase in expression of the IL-10 receptor subunit $\alpha$ \textit{IL10RA} in SF mCD4$^+$, and a trend (not statistically significant) in increased \textit{IL10} expression in SF versus PB for the three cell types. Altogether, this could suggest that inflammation in PsA is refractory to the immunomodulatory effects for IL-10 signalling in SF or is counterbalanced by the immunostimulatory properties of this cytokine, hypothesised as one of the reasons for failure of IL-10 agonist therapy in CD \parencite{Marlow2013}.

Altogether, this study sought to characterise gene expression and chromatin accessibility differences between SF and PB, and has revealed interesting observations across all cell types. In this exploratory data, CD14$^+$ monocytes showed the larger number of DARs and confidently modulated genes between SF and PB as well as functional relevant pathway enrichment for those changes. The work presented is part of a collaborative multi-omics PsA pilot study. In light of this thesis, and for the cohort size available at the time of analysis, CD14$^+$ monocytes were chosen as the cell type to further explore by scRNA-seq and further integrate with mass cytometry data. mCD4$^+$ and mCD8$^+$ also relevant changes in chromatin accessibility and gene expression, and have been identified as the two cell types undergoing greater expansion in the synovium of PsA patients. In fact, colleagues involved in this project have focused on performing differential TCR clonality analysis between T cells in SF and PB some of the PsA samples included in this thesis. Interestingly, a significant number of mCD8$^+$ TCR clones with potentially shared antigen recognition across patients were found to be enriched in SF compared to PB (manuscript in preparation).

\subsubsection{Further characterisation of CD14$^+$ monocytes by scRNA-seq and mass cytometry}

Monocytes are very plastic cell populations that undergo cell differentiation at the site of inflammation, and these differences may be better captured at the single-cell level. In this pilot experiment, cluster identification in scRNA-seq from combined SF and PB CD14$^+$ monocytes using a conservative approach (see \ref{Discussion\_scRNAseq) revealed two main subpopulations (CC-mixed and CC-IL7R). CC-mixed appeared as a large heterogenous cluster in contrast to the CC-IL7R subpopulation, characterised by cells consistently expressing \textit{IL7R}, \textit{IL32} and textit{CCL5}, amongst other markers. The equal contribution from SF and PB monocytes to the CC-IL7R scRNA-seq cluster (3 and 2.7\%, respectively) contrasted with flow cytometry from the same three PsA patients, which showed approximately 35\% of the total SF CD14$^+$ monocyte population to be IL-7R$^+$ (in revision \parencite{Al-Mossawi2018}). This could be due to discrepancies between gene expression and protein translation acknowledged in the literature but may also be a consequence of lower sensitivity of scRNA-seq in quantifying gene expression, as only 20\% of the transcriptome of a cell is reverse-transcribed \parencite{Liu2016,Islam2014}. Thus the number of cells expressing \textit{IL7R} may be underestimated by scRNA-seq and the size of this cluster may be larger. These differences in sensitivity between qPCR and scRNA-seq may also partly explain the limited overlap of differentially expressed genes between SF and PB when using qPCR and scRNA-seq. Although the same top dysregulated genes (including \textit{SPP1}, \textit{FN1} or \textit{OLR1}) were identified by both techniques, a discrete number of significantly modulated genes found by qPCR were reproduced by the scRNA-seq analysis in the CC-mixed cluster. In terms of chromatin accessibility, the comparison of FCs from contrasting SF and PB chromatin accessibility in CD14$^+$ monocytes and scRNA-seq expression in the CC-mixed cluster only showed moderate correlation. This discrete correspondence between chromatin accessibility and gene expression has also been reported by other studies and may be also the result of aforementioned limitations in annotating accessible chromatin with a putative target gene \parencite{Wang2018}.

The identification of sub-populations of CD14$^+$ monocytes expressing IL-7R$^+$ is of biological interest as \textit{IL7R} polymorphisms are associated with a number of chronic inflammatory diseases, including AS and MS \parencite{Gregory2007, Cortes2007}. Although the role of IL-7 and IL-7R in mediating the immune response was only characterised in T cells, the relevance of IL-7R in CD14$^+$ monocytes under LPS stimulation has been demonstrated in eQTL studies and also at the protein level in a manuscript under review, to which I have contributed towards \parencite{Fairfax2014, Al-Mossawi2018}. In this work, a distinct transcriptional profile of the IL7R cluster in PsA SF CD14$^+$ monocytes was found to be very similar to the gene expression profile from IL7R$^+$ \textit{in vitro} stimulated monocytes. Interestingly the IL7R locus showed differentially accessible chromatin in PsA synovial fluid and was one of the top DEGs in the monocyte qPCR array. Taken together, this data supports a role for IL-7 signaling in the innate immune response and chronic inflammation.

As for the qPCR modulated genes, pathway enrichment analysis using scRNA-seq DEGs between SF and PB in the CC-mixed cluster also identified a number of significantly enriched and biologically relevant pathways, including MHC-II Ag processing, IFN signaling and extracellular-matrix components, amongst others. Interestingly, up-regulation of \textit{IFI6} and \textit{IFITM3}, two of the genes contributing to the enrichment of this pathway, have also been identified as markers of a subpopulation of IFN-$\gamma$ activated monocytes in from RA synovial tissue using scRNA-seq \parencite{Zhang2018}. Genes enriched for the extracellular-matrix pathway included members of the S100 protein family, previously reported to be dysregulated in lesional skin from psoriasis patients (Chapter \ref{ch:Results2}), which are also involved in joint erosion and development of arthritis \parencite{Raghunatha2012}. Two genes of this family, \textit{S100A8} and \textit{S100A9}, were up-regulated in lesional skin compared to uninvolved, whereas they appeared as down-regulated in SF CC-mixed monocytes in this data. The lack of overlap between the pathways identified for the DEGs in the CC-mixed and those found by qPCR in bulk CD14$^+$ monocytes could be due the result of the qPCR array being biased to a small number of genes versus the unbiased approach of the scRNA-seq.

%Follow to overlap with ATAC and qPCR example CCL2

Single-cell mass cytometry characterisation in matched PBMCs and SFMCs from the same patients was conducted in order identify monocytes subpopulations at the proteomic level. The data from the three PsA patients revealed clear separation between classical, intermediate and non-classical monocytes. However, no further subpopulations within the classical and intermediate monocytes (CD14$^+$ monocytes investigated at by scRNA-seq) have been distinguished. Despite study limitations, such as small sample size, mass cytometry data has confirmed increased chemokine and cytokine production in SF CD14$^+$ monocytes, previously suggested by chromatin accessibility and transcriptomic data. For example, up-regulation of \textit{SPP1} in SF CD14$^+$ monocytes was validated at the protein level though increased levels of osteopontin in this cell type. Moreover, NF-$\kappa$B activation in SF compared to PB was confirmed by increased number of SF CD14$^+$ monocytes producing TNF-$\alpha$ compared to the PB counterpart.

\subsection{Challenges and future perspectives in multi-omic approaches}

\label{Discussion\_scRNAseq}

Advances in epigenetic and single-cell techniques has enabled investigation of the chromatin accessibility, transcriptomic and proteomic profiles in clinical samples. This multi-omics approach represents a very powerful strategy to dissect disease pathophysiology in a cell type specific manner. The work presented in this chapter is an exploratory study and a proof of principle for the implementation of a multi-omics approach. Nevertheless, a number of limitations and challenges have been encountered and need to be taken into account to contextualise these results. One of the limitations is the small sample size (n=3) and the lack of paired data across all the techniques here presented. This results from difficulties of recruiting samples from PsA patients na\'{i}ve for any treatment, the logistical difficulties to coordinate all of the techniques from the same sample, and the high cost this approach. A further limitation in this study is the lack of PB from healthy controls or SF from another autoimmune or non-inflammatory joint disease, as included by other studies \parencite{Fumitaka2018, Dolcino2015,Zhang2018}. The definition and categorisation of the qPCR significantly modulated genes into systemic, tissue-specific and putative disease-specific involves limitations due to lack of control samples to compare to the SF cells and the use of a biased transcriptomic analysis using a qPCR array.

Another challenge in this study relates to the analysis and integration of scRNA-seq and mass cytometry data. Both techniques still represent an emerging field were no consensus has been reached on the best strategy to combine samples across patients and experiments, accounting for batch effect. This is particularly relevant when aiming to identify further subpopulations within cell types, where the defined clusters may result from unsuccessful correction of this batch effect. In this exploratory study, CCA was chosen as the best strategy to combine CD14$^+$ monocytes scRNA-seq data first individually subseted in each sample. However standard resolution for cluster identification yielded potentially spurious subpopulations, leading to the adoption of a more conservative approach to define clusters in this particular analysis. This may be the consequence of remaining batch effect and alternative methods of combining samples from the different experiments should be investigated. In this respect, the identification of robust and stable subpopulations through cluster analysis will benefit from the implementation of algorithm design for cluster validation such as Silhouettes, which has been successfully used in the field of single-cell \parencite{Rousseeuw1987, Zhang2018}. In addition to this, incorporation of bulk RNA-seq data to from CD14$^+$ monocytes will help to interpret and validate the scRNA-seq results. In mass cytometry data, to reduce batch effect, subsequent recruited patient samples are undergoing ex-vivo fixation and cryopreservation followed by simultaneous staining and barcoding. Moreover, different methodologies for cluster identification and annotation are also being explored.

In terms of data integration, this pilot study shows a simplistic multi-omics data integration limited the sample size, technical aspects and time scale, and provides a platform for future validation studies. A more systematic integrative approach should be implemented for the expanded cohort to establish appropriate correlation across datasets. Currently Zhang and colleagues have presented the most exhaustive methodology to integrate bulk-RNA-seq, scRNA-seq, mass cytometry and flow cytometry into multi-modal transcriptomic and proteomics profiles in RA \parencite{Zhang2018}. This strategy has revealed disease-specific functional expanded subpopulations amongst the most relevant cell types in RA pathophysiology. Additionally, the correlation between bulk ATAC and scRNA-seq is clearly limited by the different scales of the two approaches. Therefore, generation of scATAC-seq data, identification of clusters based on chromatin profiling and appropriate methods for the overlap with scRNA-seq populations should be conducted to have a better understating of the correlation between chromatin accessibility and gene expression at the single-cell level \parencite{Duren2018}.

\subsection{The use of PsA functional data to inform fine-mapping GWAS loci}

The integration of epigenetic data with fine-mapped SNPs from GWAS studies have been widely demonstrated to be a powerful tool to further narrow down the candidate putative causal variants, particularly for intergenic or intronic signals not driven by LD with coding SNPs \parencite{Bunt2015,Farh2014}. Although DARs between SF and PB in four cell types did not show any overlap with SNPs from the credible set, significant enrichment of fine-mapped SNPs for accessible chromatin in all four cell types was demonstrated. A number of SNPs from the 5q31 credible set in my analysis have revealed SNPs overlapping accessible chromatin and eQTL signals in the same cell type. Integration of tCD4$^+$ and tCD8$^+$ eQTL from Kasela \textit{et al.} also confirmed the association between the SNPs from the 5q31 GWAS signal and T cells \textit{SLC22A5} expression, previously reported in a smaller study by Bowes \textit{et al.}. Moreover, this integration also highlighted a potential role for the 5q31 PsA-specific GWAS association in regulating expression of \textit{P4HA2} and \textit{SLC22A5} in unstimulated and stimulated monocytes. Interestingly \textit{SLC22A5}, a cell membrane transporter of carnitine involved in fatty acids metabolism, has been prioritised by our in-house pipeline as the third most promising druggable candidate for the treatment of psoriasis. Supporting evidence highlights the implications of this gene in other inflammatory conditions such as CD and the potential for a good therapeutic candidate in PsA \parencite{Leung2006}.

Contrary to the initial hypothesis, the integration of fine-mapping data and DARs between SF and PB in PsA relevant cell types failed to show any overlap in any of the four studied cell types. This may result from the small sample size used for the identification of DARs or the limited power of the fine-mapping analysis using only a subset of the PsA GWAS cohort. For associations driven by exonic non-synonymous missense mutations, such as \textit{TRAIF3IP2}, changes in chromatin accessibility are not necessarily expected. In other instances, for non-coding signals hypothesised to have a regulatory role, these results may suggest that fine-mapped SNPs from PsA GWAS loci do not have a tissue specific effect in chromatin accessibility changes for any of these cell types. Thus, an approach focused on variation of chromatin accessibility upon genotyping data, similarly to the example of chr2p15 presented in Chapter \ref{ch:Results2}, may be more informative when integrating epigenetics at a GWAS associated locus.