



# 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

<b>Abstract</b>	<b>i</b>
<b>Acknowledgements</b>	<b>ii</b>
<b>Declarations</b>	<b>iii</b>
<b>Submitted Abstracts</b>	<b>iv</b>
<b>Associated Publications</b>	<b>v</b>
<b>Contents</b>	<b>vi</b>
<b>List of Figures</b>	<b>viii</b>
<b>List of Tables</b>	<b>ix</b>
<b>Abbreviations</b>	<b>x</b>
<b>1 Material and Methods</b>	<b>1</b>
1.1 Ethical approval and recruitment of study participants . . . . .	1
1.1.1 Psoriasis patient recruitment . . . . .	1
1.1.2 PsA patient recruitment . . . . .	2
1.1.3 Healthy volunteer recruitment . . . . .	3
1.2 Sample processing . . . . .	3
1.2.1 PBMC and synovial fluid cells isolation . . . . .	3
1.2.2 Primary cell isolation using magnetic-activated cell sorting	4
1.2.3 Primary cell isolation using fluorescence-activated cell sorting . . . . .	4
1.2.4 Skin biopsies processing and adherent assay . . . . .	5
1.3 Experimental protocols . . . . .	6
1.3.1 Cryopreservation and cell culture . . . . .	6
1.3.2 ATAC-seq, Fast-ATAC and Omni-ATAC . . . . .	7
1.3.3 Chromatin Immunoprecipitation with sequencing library preparation by Tn5 transposase . . . . .	10
1.3.4 RNA extraction and RNA-seq . . . . .	12
Bulk RNA-seq . . . . .	12
PsA memory CD4 <sup>+</sup> and CD8 <sup>+</sup> single-cell RNA-seq . . . . .	15
Small-bulk RNA-seq . . . . .	15
1.3.5 Sequencing of the T cell receptor V(D)J region . . . . .	15
1.3.6 DNA extraction and genotyping of the SNP rs4672505 . . .	16
1.3.7 CyTOF . . . . .	16
1.4 Computational and statistical analysis . . . . .	16
1.4.1 ATAC-seq data analysis . . . . .	16

## CONTENTS

---

	Next generation sequencing data analysis . . . . .	17
	Peak calling and filtering . . . . .	17
	Combined peak master list and differential analysis . . . .	18
1.4.2	ChIPm data analysis . . . . .	19
	Next generation sequencing data analysis . . . . .	19
	Peak calling and filtering . . . . .	19
	Combined peak master list and differential analysis . . . .	19
1.4.3	RNA-seq data analysis . . . . .	20
	Bulk RNA-seq analysis . . . . .	20
	Single-cell RNA-seq analysis . . . . .	20
1.4.4	Pathway enrichment analysis . . . . .	20
1.4.5	Genomic region annotation and pathway visualisation . . .	20
1.4.6	Statistical fine-mapping . . . . .	20
	<b>Bibliography</b>	<b>21</b>



# List of Figures

# List of Tables

1.1	Antibody panel used for FACS separation of primary cell populations in controls and PsA samples . . . . .	5
1.2	Modified Illumina Nextera indexing primers . . . . .	9
1.3	ChIPm buffers modified from Schmidl <i>et. al</i> , 2015 . . . . .	13
1.4	Antibody panel used for immunoprecipitation of histone marks in ChIPm . . . . .	14

# Abbreviations

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

# Chapter 1

## Material and Methods

### 1.1 Ethical approval and recruitment of study participants

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

#### 1.1.1 Psoriasis patient recruitment

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

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

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

## Material and Methods

---

- fulfillment of the clinically accepted Psoriasis Area and Severity Index (PASI) classification for psoriasis diagnosis (**Fredriksson1978**)
- moderate to severe disease (PASI<sub>≥</sub>5)
- less than 2 weeks without antibiotics unless used for prophylaxis
- available clinical information and written consent

Detailed clinical information of the psoriasis cohort is included in

### 1.1.2 PsA patient recruitment

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

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

- over 18 years old
- previously or newly diagnose, with concomitant psoriasis, in a flare and going into biologic therapy for the first time
- fulfillment of the clinically accepted PsA Response Criteria (PSARC) including a physician global assessment questionnaire (**Philipp2011; Clegg1996**)
- oligoarticular phenotype and naïve for any treatment
- less than 2 weeks without antibiotics unless used for prophylaxis
- available clinical information and written consent

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

### 1.1.3 Healthy volunteer recruitment

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

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

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

## 1.2 Sample processing

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

### 1.2.1 PBMC and synovial fluid cells isolation

PBMC were isolated from blood samples through density gradient separation using Ficoll-Paque. Total synovial fluid (SF) cells (SFC) were isolated by centrifugation at 500g for 5 min. Both were washed twice in Hanks balanced salt solution without calcium or magnesium (Thermo Fisher Scientific) and

resuspended in phosphate saline buffer (PBS, Gibco) supplemented with 0.5% fetal calf serum (FCS, Invitrogen) and 2mM ethylenediaminetetraacetic acid (EDTA, Sigma) prior to cell types separation. Cell numbers and viability were determined by manual count using a haemocytometer and trypan blue (Sigma).

### **1.2.2 Primary cell isolation using magnetic-activated cell sorting**

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

### **1.2.3 Primary cell isolation using fluorescence-activated cell sorting**

Primary cell subpopulations from controls to study the effect of cryopreservation in chromatin states (Chapter 3) or PsA blood and SF samples were isolated by FACS. PBMC and SFC were resuspended in PBS 1mM EDTA (FACS buffer) at  $10 \times 10^6$  cells/mL, stained with the appropriate antibody cocktail (Table 1.1) for 30 min at 4°C, washed with FACS buffer and centrifuged at 500g for 5 min at 4°C. For the samples used in Chapter 3, a modified FACS buffer supplemented with 3 mM EDTA , 2% FCS and 25mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, Invitrogen) was used to avoid cell clumping after cryopreservation and short recovery. After removing the supernatant, cells were resuspended in FACS buffer prior to separation.

## Material and Methods

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

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

**Table 1.1:** Details regarding target molecule, fluorochrome, clone, supplier and dilution used for PBMC and SFC staining are provided for each of the antibodies in the panel. In controls only CD3, CD4 and CD4 markers were used.

### 1.2.4 Skin biopsies processing and adherent assay

KC enrichment from skin biopsies was performed as described in Gutowska-Owsiak and colleagues (GutowskaOwsiak and Schaupp 2012). Skin biopsies (approximately 4mm) were washed with PBS, cut in 1mm width strips and incubated in 2U/mL of dispase II (Sigma) overnight at 4°C. The epidermis was separated from the dermis and either snap-frozen in liquid nitrogen (for RNA extraction) or further digested in trypsin (Invitrogen) at 37°C for 5 min,



when used for chromatin accessibility assay. After digestion the resulting cell suspension was filtered through a 70µm nylon strainer (BD) and washed with PBS. In some instances cells were manually counted and aliquoted for ATAC-seq processing. In others, cell from each of the biopsies were resuspended in KGM-2 BulletKit (Lonza) supplemented with 0.06mM Ca<sup>2+</sup> and cultured in a collagen IV coated 96-well plate at 4°C for 10 min or 3 hours, upon experimental requirements (see Chapter X). After culturing, cells were washed twice with 200µL of PBS and kept at 37°C for downstream processing.

## 1.3 Experimental protocols

### 1.3.1 Cryopreservation and cell culture

For the controls samples in Chapter 3, 40-50x10<sup>6</sup> of PBMC were freeze-thawing using a modified version of the (Kent 2009) protocol, where cells were pre-conditioned in RPMI 1640 (brand) complete medium supplemented with 2 mM L-glutamine, 100U penicillin and strep 100µg/mL and 50% FCS for 30 minutes and afterwards diluted 1 in 2 in complete RPMI 1640 (supplemented as previously described) with 20% dimethyl sulfoxide (DMSO, Sigma). PBMC underwent slow cryopreservation at -80°C in isopropanol at -1°C per minute and stored for a minimum of two weeks in liquid nitrogen. PBMC were thaw, resuspended in supplemented complete RPMI 1640 with 10% FCS at a density of 10<sup>6</sup> cells/mL and rested for 30 min at 37°C, 5% CO<sub>2</sub> in 25mL non-adherent polypropylene cell culture flasks followed by filtering through a 40µm to obtain an homogenous cell suspension undergoing FACS separation. Frozen Normal human epidermal keratinocytes (NHEK) in passage 3 were recovered and cultured at a cell density of 5x10<sup>6</sup> cells/mL in a 75 mL adherent cell culture flask (brand) in EpiLife basal medium (Gibco) following manufacturer's instructions. After recovery NHEK were trypsinised at room temperature for 8

## Material and Methods

---

minutes followed by trypsin inactivation with EpiLife 10% FCS, centrifugation at 180g for 10 min at room temperature and manual counting with trypan blue. NHEK were seeded in a 96-well plate in 100uL of medium at a cell density of 160 cells/ $\mu$ L. NHEK were cultured for 2 days to a 90-100% confluence before being used downstream.

### 1.3.2 ATAC-seq, Fast-ATAC and Omni-ATAC

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

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

After appropriate determination of the amount of DNA amplification using qPCR, samples were amplified and singled indexed for 11 PCR cycles using modified Nextera primers from Buenrostro *et al.*, 2013 (Table1.2). The resulting DNA libraries were purified using the MinElute PCR purification kit (Qiagen) and additional Agencourt AMPure XP Magentic Beads (Beckman Coulter),

## Material and Methods

---

according to the manual specifications, to remove the remaining adaptors and primer dimers.

Primer name	Full sequence
Ad1.noMX	AATGATACGGCGACCCAGAGATCTACACTCGTCGGCAGCGTCAGATGTG
Ad2.1	CAAGCAGAAAGACGGCATAACGAGATTCGCCCTAGTCTCGTGGCTCGGAGATGT
Ad2.2	CAAGCAGAAAGACGGCATAACGAGATCTAGTACGGTCTCGTGGCTCGGAGATGT
Ad2.3	CAAGCAGAAAGACGGCATAACGAGATTCTGCCCTGTCTCGTGGCTCGGAGATGT
Ad2.4	CAAGCAGAAAGACGGCATAACGAGATGCTCAGGAGTCTCGTGGCTCGGAGATGT
Ad2.5	CAAGCAGAAAGACGGCATAACGAGATAGGAGTCCGTCTCGTGGCTCGGAGATGT
Ad2.6	CAAGCAGAAAGACGGCATAACGAGATCATGCCCTAGTCTCGTGGCTCGGAGATGT
Ad2.7	CAAGCAGAAAGACGGCATAACGAGATGTAGAGAGGTCTCGTGGCTCGGAGATGT
Ad2.8	CAAGCAGAAAGACGGCATAACGAGATCCTCTGTGTCGTCTCGTGGCTCGGAGATGT
Ad2.9	CAAGCAGAAAGACGGCATAACGAGATAGCGTAGCGTCTCGTGGCTCGGAGATGT
Ad2.10	CAAGCAGAAAGACGGCATAACGAGATCAGCCTCGGTCTCGTGGCTCGGAGATGT
Ad2.11	CAAGCAGAAAGACGGCATAACGAGATGCCCCTCTTGTCTCGTGGCTCGGAGATGT
Ad2.12	CAAGCAGAAAGACGGCATAACGAGATTCCTCTACGTCTCGTGGCTCGGAGATGT
Ad2.13	CAAGCAGAAAGACGGCATAACGAGATATCACGACGTCTCGTGGCTCGGAGATGT
Ad2.14	CAAGCAGAAAGACGGCATAACGAGATACAGTGGTGTCTCGTGGCTCGGAGATGT
Ad2.15	CAAGCAGAAAGACGGCATAACGAGATCAGATCCAGTCTCGTGGCTCGGAGATGT
Ad2.16	CAAGCAGAAAGACGGCATAACGAGATACAAACGGGTCTCGTGGCTCGGAGATGT
Ad2.22	CAAGCAGAAAGACGGCATAACGAGATTGTGACCA GTCTCGTGGCTCGGAGATGT
Ad2.23	CAAGCAGAAAGACGGCATAACGAGATAGGGTCAA GTCTCGTGGCTCGGAGATGT

Table 1.2: Name and full sequence of the PCR primers used for amplification, indexing and pooling of the ATAC-seq and ChIPm samples in this thesis. These primers were designed by Buenostro *et al.*, 2013 and they are an modified version of the Nextera Illumina primers optimised for larger molecular weight DNA fragments from low input samples. All samples were indexed with the universal primer Ad1.noMx and one of the additional 18 primers. The indexing sequence of each of the primers is in blue text.

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

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

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

### **1.3.3 Chromatin Immunoprecipitation with sequencing library preparation by Tn5 transposase**

Assessment of histone marks modification in the chromatin of psoriasis patients from Cohort 1B and four age matched healthy volunteers was performed

## Material and Methods

---

using a low cell input Chromatin Immunoprecipitation (ChIP) method known as ChIPmentation (ChIPm). For each individual and cell type three histone marks, including H3K27ac, H3K4me1 and X were tested in chromatin isolated from 100,000 cells and compared to an input chromatin sample processed in parallel. Samples were processed following the protocol published by Schmidl and colleagues (**Schmidl2015**) with some modifications. Aliquots of 600,000 cells of MACS sorted cell types, as described in 1.2, were fixed with 1% formaldehyde (Sigma) and snap frozen in dry ice and ethanol prior to storage at -80°C. Chromatin sonications of the different individuals and cell types were performed in one batch using Covaris M220(Covaris). Each of the aliquots was resuspended in 130µL of SDS lysis buffer (Table1.3), sonicated for 8 min using a duty factor of 5% and aliquoted for single ChIPm reactions prior to long term storage at -80°C.

Sonicated chromatin aliquots were thawed and 1.5 volumes of ChIP equilibration buffer (Table1.3) was added in order to reduce the SDS concentration to 0.1% and to achieve the appropriate concentration of NaCl and Triton-X100. For the immunoprecipitation step, samples were incubated with the appropriate amount of antibody (Table 1.4) overnight in rotation at 4°C. Protein-A Dynabeads (Invitrogen) were also washed in beads wash buffer (Table1.3), blocked with yeast tRNA (Ambion) and BSA (NEB) overnight in rotation at 4°C and added to the sample-antibody mix for incubation. After appropriate washes, tagmentation was performed on the bead-antibody bound crosslinked complexes, which prevents overtagmentation of the DNA. This was followed by protein K digest, reverse crosslinking, elution of the DNA and from the beads and purification using PCR MinElute kit. The chromatin input control samples were quantified with Qbit after reverse crosslinking and 1ng of chromatin was used for tagmentation.

## Material and Methods

---

Amplification by qPCR was done in each of the samples and control inputs to identify the number of full cycles required to reach one-third of the final fluorescence. Libraries were amplified for the number of cycles minus one determined with this strategy to minimise the total number of PCR replicates. The primers used for amplification and indexing were the ones optimised by Buenrostro and colleagues (Table 1.2). The number of amplification cycles for each of the samples is recorded in

### 1.3.4 RNA extraction and RNA-seq

#### Bulk RNA-seq

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

## Material and Methods

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

**Table 1.3: Composition of the three modified buffers in house for the ChIPm protocol: SDS lysis buffer, ChIP equilibration buffer, beads washing buffer and ChIP buffer. For each of the buffers the reagents, composition and supplier are indicated. The final volume prepared for each buffer was adjusted depending on the number of samples processed at the time. Sodium dodecyl sulfate (SDS), PI (proteinase inhibitor) Supplier for each of the reagents as follows: SDS (Sigma), EDTA (xxx), Tris-HCl pH8 (xx), Triton-X100 (xxx), NP-40 (Sigma) NaCl (xx), PI (Roche) and water (Ambion)**



## Material and Methods

---

Histone mark	Feature	$\mu\text{g}$ per sample	Manufact
H3K27ac	Active enhancer, promoter	2	Diagenode (C1
H3K4me1	Enhancer	1	Diagenode (C1
H3K4me3	Active promoter, enhancer	1	Diagenode (C1

**Table 1.4: Details regarding the histone marks, the the most likely chromatin state delineated, the amount of antibody required per reaction and the supplier and catalog num of the antibodies.**

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

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

### PsA memory CD4<sup>+</sup> and CD8<sup>+</sup> single-cell RNA-seq

Single CD4<sup>+</sup> or CD8<sup>+</sup> memory T cells were sorted in 2µL of cell lysis buffer into 96-well plates. Four and three plates were prepared for CD4<sup>+</sup> and CD8<sup>+</sup>, respectively, including wells containing 50 cells in each of the plates for control purposes. Libraries were prepared, indexed and sequenced at the Sanger Institute (Cambridge) using the Smart-seq2 methodology as described by Picelli and colleagues (Picelli2014). Samples were sequenced using 50-bp single-end Illumina HiSeq2500 at a depth of approximately 4 million reads per cell.

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

### Small-bulk RNA-seq

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

### 1.3.5 Sequencing of the T cell receptor V(D)J region

Single-cell sequencing of the V(D)J segments from TCR transcripts was performed simultaneously with the 10X Genomics technology Chromium 5' expression library kit (PN-1000014) by the Oxford Genomics Centre at the Wellcome Trust Centre for Human Genetics. In short, full-length cDNA was amplified by PCR with primers against the 5 and 3 ends of the barcode sequences

inherent to the 10X Genomics bead technology. The amplified material was divided for use in 5' total sc-RNAseq (as specified previously in 1.3.4) and also for enrichment of the TCR by PCR amplification with specific primers. TCR enriched cDNA was followed by enzymatic fragmentation and size selection in order to obtain variable length fragments spanning the V(D)J segments. Library prep and indexing was followed by Illumina HiSeq4000 xxxx bp single-end sequencing at a depth of XXX reads per cell.

### 1.3.6 DNA extraction and genotyping of the SNP rs4672505

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

### 1.3.7 CyTOF

## 1.4 Computational and statistical analysis

### 1.4.1 ATAC-seq data analysis

ATAC-seq, Fast-ATAC and Omni-ATAC data were analysed using an in house pipeline towards which development I made an important contribution.

## Material and Methods

---

The pipeline performs single sample data processing and it also builds a combined master list for each of the comparisons of interest to later perform differential analysis.

### Next generation sequencing data analysis

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

### Peak calling and filtering

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

## Material and Methods

---

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

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

### Combined peak master list and differential analysis

To perform differential open chromatin analysis a non-overlapping 500bp peak master list was built for each cell type separately by union of all the peaks that were present in at least 30% of the samples, regardless patient and control group. Reads overlapping each of those peaks were retrieved for each sample using HTSeq-count algorithm () to build a combined count matrix. The count matrix was quantile normalised using the function `normalize.quantiles` from R. An empirical 95% confidence cut off for the number of counts in absent peaks was calculated on the normalised count matrix (**Xinmin2005**; **Jonker2014**) and used to replace the counts in each of the peaks in the master list not called (absent) in the corresponding sample. Differential analysis was performed using linear model analysis with `limma voom` R package and correcting for covariates xxxxx A combined master list for all cell types and tissues (when applicable) was also built following the same strategy for principal component analysis (PCA) purposes.

### 1.4.2 ChIPm data analysis

#### Next generation sequencing data analysis

ChIPm NGS data from samples and inputs were processed similarly to ATAC-seq (*ATAC<sub>a</sub>analysis*) for trimming, mapping and filtering. For visualisation of each sample subtract and bedGraphToBigWig tools.

#### Peak calling and filtering

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

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

#### Combined peak master list and differential analysis

To perform differential histone modification analysis a peak master list was built and counts at each of the locations were performed as described for ATAC-seq (*ATAC<sub>a</sub>analysis*). PCA analysis was performed for samples from all cell types and individuals.

#### Differential analysis

### 1.4.3 RNA-seq data analysis

#### Bulk RNA-seq analysis

#### Single-cell RNA-seq analysis

### 1.4.4 Pathway enrichment analysis

### 1.4.5 Genomic region annotation and pathway visualisation

Annotation of genomic regions and signalling-pathways visualisation were performed with two functionalities of the R package and web-app Atlas and Analysis of systems-biology-led pathways, developed by Dr. Hai Fang and towards which I have made some contribution (submitted manuscript)(). ATAC-seq and ChIPm peaks were annotated with gene entities taking into account not only the closest gene but also publicly available data regarding chromatin conformation capture (), eQTL () and enhancer RNA () in several cell types. This approach was not biased by the physical vicinity to a gene when annotating genomic intervals located at intergenic regions or gene deserts and it integrates the knowledge in the field about regulatory regions affecting the expression of distal genes. For visualisation, manually curated KEGG pathways including all the genes for each gene family were colored based on the fold-change from the corresponding differential analysis and highlighted in bold when passing the FDR threshold for significance .

### 1.4.6 Statistical fine-mapping

# Bibliography

- Buenrostro, Jason D., Wu, Beijing, Chang, Howard Y., and Greenleaf, William J. (2015). "ATAC seq: A Method for Assaying Chromatin Accessibility Genome Wide". *Current Protocols in Molecular Biology* 109: pp. 1–9. doi: 10.1002/0471142727.mb2129s109. URL: <http://dx.doi.org/10.1002/0471142727.mb2129s109>.
- GutowskaOwsiak, D. and Schaupp, A. L. (2012). "IL17 downregulates filaggrin and affects keratinocyte expression of genes associated with cellular adhesion". *Experimental dermatology* 21: pp. 104–110. doi: 10.1111/j.1600-0625.2011.01412.x. URL: <http://dx.doi.org/10.1111/j.1600-0625.2011.01412.x>.
- Kent, L. (2009). "Freezing and thawing human embryonic stem cells". *Journal of visualized experiments: JoVE*. URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3152242/>.