Dissecting the role of ISWI chromatin remodellers in inflammatory gene expression

JÚLIA MELIÀ ALOMÀ

Milano, 2020
This thesis, submitted to the Open University for the degree of Doctor of Philosophy, has been carried out and completed by me, Júlia Melià Alomà, under the supervision of Dr. Gioacchino Natoli. I hereby declare that I designed and performed all the experimental work described in this study, including preparation of libraries for next-generation sequencing (NGS). NGS samples were sequenced by Dr. Sara Polletti on an Illumina NextSeq 500 platform, while Dr. Chiara Balestrieri, Dr. Viviana Piccolo, Dr. Francesco Gualdrini and Pierluigi Di Chiaro were responsible for the bioinformatic analyses of NGS data.
Dedico aquesta tesi als que m’han donat sempre el seu suport incondicional: la meva mare i el meu pare

“Caminante, son tus huellas el camino y nada más; Caminante, no hay camino, se hace camino al andar. Al andar se hace el camino, y al volver la vista atrás se ve la senda que nunca se ha de volver a pisar.”

Antonio Machado
Proverbios y cantares (XXIX)
# TABLE OF CONTENTS

LIST OF ABBREVIATIONS ........................................................................... 5  
FIGURES INDEX ...................................................................................... 8  
TABLES INDEX ........................................................................................ 10  
ABSTRACT ................................................................................................. 11  
INTRODUCTION .......................................................................................... 12  
   1. Transcriptional regulation of the inflammatory response .............. 12  
      1.1 Overview of the transcription process .................................. 12  
      1.2 Gene regulation in murine macrophages ............................... 17  
   2. The role of chromatin during transcription .................................. 21  
      2.1 Nucleosomes: the fundamental unit of chromatin ............... 21  
      2.2 Nucleosome positioning at regulatory elements .................. 23  
      2.3 Families of ATP-dependent chromatin remodellers .......... 24  
   3. ISWI family of ATP-dependent chromatin remodellers ............. 27  
      3.1 Structure and regulation of the ISWI ATPase .................... 28  
      3.2 Biological functions of the ISWI complexes ..................... 29  
AIMS .......................................................................................................... 33  
MATERIALS AND METHODS ................................................................. 34  
   1. Mouse biology ................................................................................ 34  
      1.1 Strains .................................................................................. 34  
      1.2 Genotyping .......................................................................... 34  
   2. Cell culture ..................................................................................... 34  
      2.1 Isolation, growth conditions and treatment of bone marrow-derived macrophages (BMDMs) ............................................. 34  
      2.2 Cell lines: growth conditions ............................................. 35  
      2.3 Plasmids .............................................................................. 35  
      2.4 Lentiviral infection ............................................................... 36  
      2.5 Retroviral infection ............................................................... 37  
   3. Molecular biology .......................................................................... 37  
      3.1 Protein extraction and western blotting ............................... 37  
      3.2 Co-immunoprecipitation analysis ....................................... 38  
      3.3 Chromatin immunoprecipitation ....................................... 39
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADP</td>
<td>Adenosine DiPhosphate</td>
</tr>
<tr>
<td>AP-1</td>
<td>Activator Protein-1</td>
</tr>
<tr>
<td>ATAC</td>
<td>Assay for Transposase-Accessible Chromatin</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine TriPhosphate</td>
</tr>
<tr>
<td>BAZ1</td>
<td>Bromodomain Adjacent to Zinc finger domain 1</td>
</tr>
<tr>
<td>BMDMs</td>
<td>Bone Marrow Derived Macrophages</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>BPTF</td>
<td>Bromodomain PHD finger Transcription Factor</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>C/EBP</td>
<td>CCAAT-Enhancer-Binding Protein</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CHD</td>
<td>Chromodomain Helicase DNA-binding</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin ImmunoPrecipitation</td>
</tr>
<tr>
<td>CSB</td>
<td>Cockayne Syndrome group B</td>
</tr>
<tr>
<td>CTCF</td>
<td>CCCTC-binding Factor</td>
</tr>
<tr>
<td>DAVID</td>
<td>Database for Annotation, Visualisation and Integrated Discovery</td>
</tr>
<tr>
<td>DEGs</td>
<td>Differentially Expressed Genes</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's Modified Eagle Medium</td>
</tr>
<tr>
<td>DNA</td>
<td>DeoxyriboNucleic Acid</td>
</tr>
<tr>
<td>DPE</td>
<td>Downstream core Promoter Element</td>
</tr>
<tr>
<td>DTT</td>
<td>DiThioThreitol</td>
</tr>
<tr>
<td>ETS</td>
<td>Erythroblast Transformation Specific</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal Bovine Serum</td>
</tr>
<tr>
<td>FC</td>
<td>Fold Change</td>
</tr>
<tr>
<td>FDR</td>
<td>False Discovery Rate</td>
</tr>
<tr>
<td>GC</td>
<td>Guanine-Cytosine</td>
</tr>
</tbody>
</table>
**GREAT**  Genomic Regions Enrichment of Annotations Tool

**GTFs**  General Transcription Factors

**H3K27Ac**  Histone 3 Lysine 27 Acetylation

**H3K4me1**  Histone 3 Lysine 4 monomethylation

**HnRNPU**  Heterogeneous nuclear Ribonucleoprotein U

**HSS**  HAND-SANT-SLIDE

**HSV-1**  Herpes Simplex Virus 1

**Ifnb1**  Interferon beta 1

**IL**  Interleukin

**INO80**  Inositol requiring 80

**IP**  ImmunoPrecipitation

**IRF**  Interferon Regulatory Factor

**ISGs**  Interferon-Stimulated Genes

**ISWI**  Imitation SWI/SWI

**JAK**  Janus Kinase

**KO**  KnockOut

**LPS**  LipoPolySaccharide

**Luc**  Luciferase

**MACS**  Model-based Analysis of ChIP-Seq

**MNase**  Micrococcal Nuclease

**mRNA**  messenger RNA

**NF-kB**  Nuclear Factor kappa-light-chain-enhancer of activated B cells

**NFR**  Nucleosome-Free Region

**NGS**  Next-Generation Sequencing

**NURF**  Nuclear Remodelling Factor

**PBS**  Phosphate-Buffered Saline

**PCR**  Polymerase Chain Reaction

**Pen/Strep**  Penicillin/Streptomycin
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHD</td>
<td>Plant HomeoDomain</td>
</tr>
<tr>
<td>PIC</td>
<td>Pre-Initiation Complex</td>
</tr>
<tr>
<td>PMSF</td>
<td>PhenylMethylSulfonyl Fluoride</td>
</tr>
<tr>
<td>pol II</td>
<td>polymerase II</td>
</tr>
<tr>
<td>PRG/SRG</td>
<td>Primary/Secondary Response Gene</td>
</tr>
<tr>
<td>PSCAN</td>
<td>Promoter SCANning</td>
</tr>
<tr>
<td>RIPA</td>
<td>Radio ImmunoPrecipitation Assay buffer</td>
</tr>
<tr>
<td>RNA</td>
<td>RiboNucleic Acid</td>
</tr>
<tr>
<td>RPKM</td>
<td>Reads Per Kilobase Million</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>quantitative Reverse Transcription PCR</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulfate</td>
</tr>
<tr>
<td>seq</td>
<td>sequencing</td>
</tr>
<tr>
<td>shRNA</td>
<td>short hairpin RNA</td>
</tr>
<tr>
<td>SICER</td>
<td>Spatial clustering for Identification of ChIP-Enriched Regions</td>
</tr>
<tr>
<td>SMARCA</td>
<td>SWI/SNF-related Matrix-associated Actin-dependent Regulator of Chromatin subfamily A</td>
</tr>
<tr>
<td>SPRI</td>
<td>Solid-Phase Reversible Immobilisation</td>
</tr>
<tr>
<td>STAT</td>
<td>Signal Transducer and Activator of Transcription</td>
</tr>
<tr>
<td>SWI/SNF</td>
<td>SWItch/Sucrose Non-Fermentable</td>
</tr>
<tr>
<td>TAFs</td>
<td>TBP-associated factors</td>
</tr>
<tr>
<td>TBP</td>
<td>TATA-Binding Protein</td>
</tr>
<tr>
<td>TF</td>
<td>Transcription Factor</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-Like Receptor</td>
</tr>
<tr>
<td>TSS</td>
<td>Transcription Start Site</td>
</tr>
<tr>
<td>UT</td>
<td>UnTreated</td>
</tr>
<tr>
<td>UV</td>
<td>UltraViolet</td>
</tr>
<tr>
<td>WT</td>
<td>WildType</td>
</tr>
</tbody>
</table>
FIGURES INDEX

Figure 1. Genomic regulatory elements controlling RNA polymerase II transcription.

Figure 2. Transcriptional control of the inflammatory response.

Figure 3. Dynamic properties of nucleosomes.

Figure 4. Domain organisation of the catalytic subunits of the different remodeller families.

Figure 5. Model for the regulation of ISWI ATPase activity.

Figure 6. Characterisation of the ISWI complexes in bone marrow-derived macrophages.

Figure 7. Evaluation of the knockdown efficiency of the two SMARCA5-targeting shRNAs used in this study.

Figure 8. RNA-seq analysis reveals a number of differentially expressed genes (DEGs) upon SMARCA5 knockdown.

Figure 9. A fraction of SMARCA5-dependent genes is strongly associated with the interferon response.

Figure 10. Evaluation of knockdown efficiency of SMARCA5-targeting shRNA.

Figure 11. ATAC-seq analysis unveils genome-wide effects of SMARCA5 depletion on chromatin accessibility.

Figure 12. Impact of SMARCA5 depletion on enhancer activity as determined by H3K27ac ChIP-seq analysis.

Figure 13. Optimisation of ChIP-seq conditions for chromatin remodellers.

Figure 14. SMARCA5 recruitment to chromatin in BMDMs as determined by ChIP-seq analysis.
**Figure 15.** Impact of SMARCA5 depletion on CEBPB expression and recruitment to chromatin.

**Figure 16.** Evaluation of knockdown efficiency of the shRNAs targeting ISWI accessory subunits.

**Figure 17.** RNA-seq analysis upon depletion of ISWI non-catalytic subunits uncovers a major role of the BPTF subunit in the regulation of the inflammatory gene expression program.

**Figure 18.** Evaluation of SMARCA5 deletion in Smarca5^{flox/flox} BMDMs transduced with a Cre-expressing retroviral plasmid.

**Figure 19.** Global and specific effects of SMARCA5 deletion on gene expression of retrovirally transduced BMDMs.

**Figure 20.** Genome-wide effects of SMARCA5 deletion on chromatin accessibility as determined by ATAC-seq analysis of BMDMs subjected to retroviral transduction.
Table 1. The ISWI family.

Table 2. List of shRNAs used in this study.

Table 3. List of antibodies and dilutions used for Western Blot analysis.

Table 4. List of primers used in this study for qPCR analysis.

Table 5. Motif enrichment analysis on the promoters of genes downregulated upon SMARCA5 depletion.
ABSTRACT

The inflammatory response is driven by a highly accurate and kinetically complex transcriptional program that is controlled by the stimulus-regulated usage of thousands of cis-regulatory elements (i.e. enhancers and promoters). Stimulus-activated transcription factors (TFs) involved in this response, as well as their interplay with lineage-determining TFs, have been extensively characterised. However, the role of different families of co-regulators that are recruited by those TFs in most cases remains to be elucidated. In light of this, I set out to dissect the role of a specific family of co-regulators, namely the ISWI family of ATP-dependent chromatin remodellers, in modulating the transcriptional response to inflammatory stimuli such as lipopolysaccharide (LPS) in bone marrow-derived mouse macrophages. By combining transcriptomic and other genomic experiments and analyses, I have found that ISWI complexes present in innate immune cells are involved in the regulation of different components of the transcriptional program of murine macrophages. On one hand, I found that ISWI complexes are strongly associated with the regulation of the interferon response. By ChIP-seq analysis, I show that distinct ISWI subunits are recruited to regulatory regions bound by Interferon Regulatory Factors (IRFs), where they control the epigenomic landscape and thus regulate the interferon gene expression program. On the other hand, by ATAC-seq analysis I also identified a large set of genomic regions which are maintained in a repressed state in the presence of the ISWI catalytic subunit SMARCA5. However, these regions are not occupied by SMARCA5, indicating that in this case it is not the direct action of the ATPase which controls the chromatin state.
INTRODUCTION

1. Transcriptional regulation of the inflammatory response

The expression of tens of thousands of protein-coding genes in eukaryotes is primarily regulated at the level of transcription. Gene expression is dynamically regulated and finely tuned depending on the developmental stage and the differentiation state of each cell, as well as in response to a wide variety of extracellular signals. In macrophages, inflammatory stimuli are able to rapidly trigger the expression of hundreds of genes (Medzhitov & Horng, 2009b). Therefore, understanding the molecular basis for the control of complex transcriptional programmes appears relevant to many physiological processes, including the induction of the inflammatory response.

1.1 Overview of the transcription process

1.1.1 Transcriptional machinery

Transcription of protein-coding genes is carried out by RNA polymerase II (RNA pol II). The discovery, in 1969, of nuclear RNA polymerases represented a milestone in the field of eukaryotic transcription (Roeder & Rutter, 1969). RNA pol II is a multi-subunit complex with an enzymatic activity that catalyses the synthesis of mRNA from the DNA template. However, purified RNA polymerase II is not able to drive accurate transcription initiation by itself, rather it requires additional auxiliary factors which are commonly known as ‘basal’ or ‘general’ transcription factors (GTFs) (Roeder, 1996).

GTFs include transcription factor (TF) IIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. TFIID is a multi-subunit complex that is formed by TBP (TATA-binding protein) and fourteen TBP-associated factors (TAFs). These factors, together with RNA pol II, assemble at promoters into the transcription pre-initiation complex (PIC) in a tightly regulated cascade of events leading to initiation of transcription (Hampsey, 1998; Orphanides et al., 1996; Sainsbury et al., 2015), which is modulated by regulatory factors bound to cis-regulatory elements.
1.1.2 *Cis*-regulatory elements

There are several classes of *cis*-regulatory DNA sequences involved in the regulation of transcription by RNA pol II, including promoters, enhancers and insulator elements. All of these elements contain recognition motifs for multiple sequence-specific DNA-binding factors that regulate transcription (Figure 1).

Regulatory elements can be experimentally identified by ChIP-seq thanks to their particular signatures of histone modifications. For example, enhancers are characterised by high levels of monomethylation of histone H3 lysine 4 (H3K4me1), while promoters are mainly enriched in H3K4 trimethylation (H3K4me3) (Heintzman et al., 2007). In addition, acetylation of histone H3 lysine 27 (H3K27ac) has been shown to correlate with enhancer as well as promoter activity (Rada-Iglesias et al., 2011).

- **Promoters**

Among all genomic regulatory elements, the core promoter is of utmost importance since it contains the site of transcription initiation, where the RNA pol II transcriptional machinery will exert its action. It has been defined as “the minimal stretch of contiguous DNA sequence that is sufficient to direct accurate initiation of transcription by the RNA pol II machinery” (Butler & Kadonaga, 2002).

The core promoter can extend for up to 35-40 bp upstream or downstream from the transcription start site (TSS) and contains different sequence motifs at distinct positions relative to the +1 start nucleotide (Butler & Kadonaga, 2002; Juven-Gershon et al., 2006). The TATA box was the first core promoter element identified and it was initially believed to be a universal motif in every core promoter, since the first protein-coding genes that were isolated all contained the consensus TATAAA sequence 25-30 bp upstream of the TSS (Goldberg, 1980).

This simplistic view was challenged when many other motifs commonly found in core promoters were identified later on, including the downstream core promoter element (DPE) (Burke & Kadonaga, 1996), the initiator element (Inr) (Smale & Baltimore, 1989) or the
TFIIB recognition element (BRE) (Hoffmann et al., 1992), together with the observation that not all promoters contain a TATA box (Burke & Kadonaga, 1996; Pugh & Tjian, 1991). Today, it is widely accepted that the structure of core promoters can be very diverse and that such diversity contributes to an additional level of transcriptional regulation (Juven-Gershon et al., 2006; Smale, 2001).

Most core promoter elements serve as recognition sites for subunits of the TFIID and TFIIB complexes (Müller & Tora, 2004; Pugh & Tjian, 1991). Different TAFs appear to be responsible for selective recognition of specific core promoter elements and, thus, direct the assembly of structurally and functionally different PICs composed of largely overlapping but partially different sets of factors (Hampsey & Reinberg, 1997; Müller et al., 2007; Verrijzer et al., 1995).

The region surrounding the core promoter (from about -250 to +250), which contains multiple binding sites for sequence-specific TFs, is known as the proximal promoter. Collectively, we can refer to the promoter as the region containing the core promoter with the TSS as well as the proximal promoter region.

A particular class of promoters to be considered are those characterised by the presence of multiple TSSs randomly distributed over a GC-rich region of 100 bp or more (Kawaji et al., 2006). These are generally TATA-less, are usually associated with fewer consensus binding sites and are often depleted of nucleosomes. Therefore, PIC assembly is thought to occur in a non-specific manner at each TSS. As a result, initiation of transcription occurs at multiple dispersed sites. These broad promoters are often associated with ubiquitously or broadly expressed genes.

- **Enhancers**

Enhancers can be located many kilobases and up to one megabase away from the TSS and can either play an activating or a repressing role. They were first characterised in cell-based reporter assays based on their capacity to activate gene expression regardless of
their location or orientation relative to the promoter in a plasmid construct (Banerji et al., 1981, 1983).

Enhancer-bound factors are believed to communicate, directly or indirectly, with factors bound to the core promoter in order to regulate transcription even from long distances. Therefore, the activity of a given promoter can be under the influence of multiple enhancers, and at the same time a given enhancer might be able to stimulate a wide range of promoters (Bulger & Groudine, 2011). However, enhancer stimulatory effects are actually limited by promoter specificity and by the presence of insulator elements (Blackwood & Kadonaga, 1998). For instance, a study by Butler and Kadonaga identified enhancers which specifically regulated the activity of promoters that contained either a canonical TATA box or DPE elements (Butler & Kadonaga, 2001).

The human genome has been estimated to contain up to a million putative enhancer elements. Enhancer signatures can be identified at a subset of these sequences in a cell-type specific manner. Consequently, enhancer activity has been proposed to play a critical role in the regulation of cell-type specific gene expression (Heintzman et al., 2009; Heinz et al., 2015). However, the extent to which genomic regions with enhancer marks are bona fide functional enhancers remains to be determined.

- **Insulators**

Insulator (or boundary) elements are defined by their capacity to block the activity of enhancers on linked promoters if located between them (Kuhn & Geyer, 2003). These sequences are usually enriched in binding sites for the CCCTC-binding factor (CTCF). CTCF can bind insulator elements, where it acts as a boundary between genomic regions with different chromatin states and independent transcriptional activity.

One possible mechanism by which CTCF is able to exert this role is through interaction with cohesins (Rubio et al., 2008). This association has been proposed to contribute to changes in the nuclear organization and transcriptional regulation by bringing distal genomic regulatory regions into vicinity of each other, thus providing a possible
mechanistic link for long-range interactions of linearly separated DNA regions, such as between CTCF-bound sites or between enhancers and their target promoters.

Figure 1. Genomic regulatory elements controlling RNA polymerase II transcription. The horizontal line represents a region of the genome. The different colours of promoters and enhancers represent their different biological functions or tissue-specificity. The regulatory domain of each gene is delimited by insulator elements. Taken from (Blackwood & Kadonaga, 1998)

1.1.3 Role of transcription regulatory factors

- Transcription factors

Sequence-specific DNA-binding transcription factors (TFs) bind to genomic cis-regulatory regions, where in turn they recruit additional downstream factors and co-factors that mediate transcription. Therefore, sequence-specific TFs play a crucial role since they function as an interface between the genetic regulatory information contained at promoters and enhancers and the RNA pol II transcriptional machinery (Kadonaga, 2004).

Transcription factors are usually composed of different functional domains, including essential modules for binding to DNA and for activation or repression of transcription. Additional modules may also be present, such as regulatory modules (e.g. nuclear receptors) or modules for multimerization.

DNA binding modules have been extensively studied and many structurally different binding domains have been defined, including the helix-turn-helix or the zinc finger motifs. TFs from the same family tend to have similar DNA binding properties but different activation functions according to their activation domains, which might interact with different transcriptional regulators.
**Co-regulators**

A typical target of TF activation domains is the Mediator complex, which in mammals is composed of 26 subunits that have each been shown to interact with different TFs (Allen & Taatjes, 2015). Mediator interacts with RNA pol II and appears to globally regulate its activity, possibly by facilitating stable PIC assembly. The exact mechanisms involved are not completely understood but probably involve interactions of Mediator with multiple PIC factors. Moreover, possible interactions of Mediator with chromatin remodelling complexes might promote nucleosome displacement at regulatory regions, while the large size of the Mediator complex may also contribute to nuclear organisation by promoting the formation of long-range interactions between enhancers and promoters (Malik & Roeder, 2010).

In addition to the global regulator Mediator, multiple other co-regulators, such as the aforementioned chromatin remodellers, also play critical roles in the control of transcription, even if the requirement for some of them might be restricted to specific cell types or even particular gene subsets. A general feature of co-regulators, as explained above for the Mediator complex, is that they do not directly bind DNA but rather are usually recruited by DNA-bound TFs through protein-protein interactions, which also allow the downstream communication with the basal transcriptional machinery. Thus, they can play a direct role in the regulation of transcription by acting as a bridge between *cis*-regulatory regions and the transcriptional machinery. Alternatively, they can indirectly activate or repress transcription by mobilising nucleosomes (i.e. chromatin remodelling) or by covalently modifying histones and other proteins.

### 1.2 Gene regulation in murine macrophages

Primary macrophages, which can be differentiated from murine bone marrow, have long been used as a model system to study transcriptional regulation due to their ability to very rapidly induce massive epigenomic and gene expression changes upon exposure to danger signals. Using primary macrophages exposed *in vitro* to specific inflammatory
stimuli (e.g. LPS or interleukin 4 (IL4)), many aspects involved in the dynamic regulation of inflammatory genes have been characterised.

Genome-wide studies using ChIP-seq have permitted the identification of the repertoire of genomic regulatory elements controlling macrophage identity and inflammatory gene expression. From those studies it became clear that macrophage-specific cis-regulatory elements are enriched with binding sites for PU.1 (Ghisletti et al., 2010; Heinz et al., 2010), a member of the ETS family of TFs specifically expressed in cells of the hematopoietic system.

High levels of expression of PU.1 are required to induce the differentiation and maintain the identity of macrophages (DeKoter & Singh, 2000; Nerlov & Graf, 1998). Pioneer or master regulators are key transcriptions factors in the control of cell type-specific gene expression programs and are defined as indispensable for proper differentiation and functioning of cells. In macrophages, PU.1 is thus considered to be a master regulator. As such, it is not surprising that the generation of myeloid progenitors is almost completely abrogated in Pu.1−/− mice (Scott et al., 1994). Moreover, it has been shown that when ectopically expressed in fibroblasts, PU.1 is able to induce their trans-differentiation into cells with a macrophage-like phenotype by activating macrophage-specific enhancers and inducing a myeloid gene expression program (Feng et al., 2008; Ghisletti et al., 2010), which further supports its role as the master regulator of macrophage differentiation.

Classical pioneer TFs are able to recognise their specific binding sites in a context of closed chromatin, which they can make accessible for subsequent binding of other sequence-specific TFs. PU.1, however, seems to require cooperative binding of additional myeloid-specific lineage-determining TFs (e.g. C/EBP family), as shown by the co-occurrence of binding motifs at genomic sites occupied by PU.1 in vivo (Heinz et al., 2010). In fact, another study by Heinz et al. later demonstrated that mutations in binding motifs for PU.1 not only affected PU.1 binding itself but also resulted in a loss of nearby binding of C/EBPβ even if those motifs remained intact, and vice versa (Heinz et al., 2013).
Consistent with this, the defects observed in C/EBPβ-deficient mice prove that PU.1 is not the only TF that is necessary for the terminal differentiation and activation of macrophages (Cain et al., 2013; Tanaka et al., 1995). Furthermore, ectopic expression of either C/EBPα or C/EBPβ in differentiated B cells was shown to trigger their efficient reprogramming into macrophages in a series of sequential changes that also required endogenous PU.1 expression (Xie et al., 2004).

In addition to the synergies between lineage-determining TFs, it has recently been shown that PU.1 directly interacts with SWI/SNF chromatin remodelling complexes (Minderjahn et al., 2020). As a result, these complexes are recruited to PU.1-bound sites, where their activity is required to promote the opening of the chromatin by catalysing a reorganisation of nucleosomal spacing. So, the ability of PU.1 to shape the macrophage-specific regulatory landscape is dependent on collaborative interactions with other TFs and co-regulators.

It was first thought that the repertoire of macrophage-specific enhancers would be limited to those constitutively marked by H3K4me1 and bound by the lineage-determining TF PU.1 (Figure 2, left) (Ghisletti et al., 2010; Heinz et al., 2010). However, a new class of enhancers, which are occluded by nucleosomes and unmarked by enhancer-specific histone marks in the basal state but become activated in response to stimulation (i.e. enhancer marks and TF binding are acquired), have been identified and defined as latent or de novo enhancers (Figure 2, right) (Kaikkonen et al., 2013; Ostuni et al., 2013). Such an expansion of the enhancer repertoire is dependent on the functional cooperation between stimulus-activated and lineage-determining TFs. In the specific case of macrophages, PU.1 and stimulus-dependent TFs such as STAT1 or STAT6 are cooperatively recruited to latent enhancers activated after exposure to inflammatory stimuli.
Figure 2. Transcriptional control of the inflammatory response. (Left) Poised enhancers are defined during differentiation of macrophages by PU.1 binding and deposition of H3K4me1. Stimulus-activated TFs, such as NF-κB, can bind the nucleosome-depleted region maintained by PU.1. Activated enhancers acquire the H3K27ac mark. (Right) Latent enhancers are not detectable in differentiated cells until they become activated by stimulus-dependent TFs, such as STATs, which collaborate with PU.1 to reorganise local chromatin. Adapted from (Smale & Natoli, 2014)

Several families of stimulus-activated TFs can be induced upon exposure to inflammatory stimuli, including AP-1, IRFs, NF-κB or STATs (Glass & Natoli, 2016; Smale & Natoli, 2014). Different external stimuli will each activate specific receptors, which in turn activate downstream signalling pathways leading to the activation of different TFs in the nucleus. Stimulus-activated TFs will subsequently induce the robust expression of distinct subsets of genes encoding key functional programs (e.g. genes controlling cell migration, phagocytosis, inflammation, immune responses). LPS, for example, is an agonist of the TLR4 receptor, which is selectively coupled to the activation of IRF3 and the NF-κB (via the IKK kinase complex) and AP-1 families (via Jun kinases) (Doyle et al., 2002). The activation of these TFs in macrophages rapidly triggers the induction of an interferon-dependent immune response, which plays an essential role in the defence against viral infections and other pathogens.

Based on their kinetics of expression as well as in their requirements for new protein synthesis and chromatin remodelling, inducible genes can be classified into two different categories, namely primary and secondary response genes (PRGs and SRGs) where PRGs are those that do not require protein synthesis for induction (Hargreaves et al., 2009;
Ifnb1 is one of the most representative examples of a PRG, while many interferon-stimulated genes (ISGs) belong to the SRGs group.

The transcriptional induction of ISGs has been suggested to depend, in part, on the chromatin structure at regulatory elements, which are maintained in a repressed state to limit transcription in the absence of stimulation. Upon activation, however, several studies have shown that TFs of the STAT and IRF families interact with chromatin remodelling complexes. Chromatin remodellers, thus, contribute to the potent induction of interferon-target genes by facilitating the opening of occluded interferon-stimulated response elements, which are consensus DNA sequences that can be recognised by the well-conserved DNA-binding domains of IRF family members. (Cui et al., 2004; Huang et al., 2002).

2. The role of chromatin during transcription

Chromatin is the nucleoprotein complex in which DNA is packaged. However, the function of chromatin goes far beyond compacting genetic information so that it can fit inside the nucleus, as it has also proven essential to gene regulation (Li et al., 2007).

2.1 Nucleosomes: the fundamental unit of chromatin

It was first proposed by R. Kornberg that chromatin structure is based on a repeating unit of eight histone molecules and approximately 200 base pairs of DNA (Kornberg, 1974; Kornberg & Lorch, 1999). These units were termed nucleosomes and represent the most fundamental level of chromatin organization in all eukaryotic genomes. Each octamer is typically made up of two copies of the highly conserved histones H2A, H2B, H3 and H4, and 147 base pairs of DNA are wrapped nearly twice around this core into a superhelix (Luger et al., 1997). Nucleosome cores are spaced by a stretch of DNA of variable length.
known as ‘linker DNA’, forming an array of nucleosomes which can be further stabilised by binding of linker histone H1.

Assembly of nucleosomal arrays determines the formation of higher-order chromatin structures. However, nucleosomes are not only responsible for DNA packaging, but their intrinsic properties also allow for a dynamic regulation of such structures and they are therefore involved in the modulation of most nuclear processes (Li et al., 2007; Saha et al., 2006), from transcription to replication or DNA repair.

The first form of chromatin differentiation or specialization involves deposition of histone variants (Figure 3a) (Henikoff & Smith, 2015). Histone variants distinguish from canonical histones not only by their primary sequence but also by the fact that they are constitutively expressed and incorporated into particular regions of chromatin in a manner non-dependent on replication. Histone variants can have different properties and when deposited, they will specialise chromatin for a specific function by modifying the fundamental structure of nucleosomes and altering their stability. For example, the histone variant H2A.Z has been linked to a number of nuclear functions, some of which are even contradictory (Kamakaka & Biggins, 2005). However, this may be attributed to the rapid rate in which these variant nucleosomes seem to be exchanged. H2A.Z is usually enriched at nucleosomes flanking the TSS, where it might contribute to the recruitment of RNA pol II and subsequent gene activation (Talbert & Henikoff, 2010).

In addition, histones can also be covalently modified at lysine residues in their amino-terminal tail domains (Figure 3b), which might directly influence chromatin structure by altering histone-DNA and histone-histone contacts (Hayes & Hansen, 2001). These modifications include methylation, acetylation, ubiquitination, ADP-ribosylation and sumoylation, and can positively or negatively affect nuclear processes such as transcription (Wolffe & Hayes, 1999). It has also been proposed that distinct patterns of histone modifications might be read by a variety of effector proteins which might specify different downstream functions (Strahl & Allis, 2000).
A third way in which chromatin structure can be regulated is through the action of chromatin remodelling complexes (Figure 3c), which can modify histone-DNA interactions and therefore affect nucleosome positioning.

![Diagram of nucleosome dynamics](image)

**Figure 3. Dynamic properties of nucleosomes.** a. Canonical histones can be replaced by variant histones (in green) with unique tails that might bind specific regulatory proteins. b. Histone modifications can also recruit specific regulatory factors. c. Nucleosome repositioning allows the binding of regulatory factors to nucleosomal DNA. Taken from (Saha et al., 2006)

### 2.2 Nucleosome positioning at regulatory elements

Some TFs, such as the hepatocyte nuclear factor HNF3, have been shown to directly bind nucleosomal DNA (Hayes & Hansen, 2001). However, most DNA-binding factors can only access their target sequences if located in nucleosome-free regions. Therefore, nucleosome occupancy strongly influences the activity of cis-regulatory elements. Several genome-wide studies in the last few years have in fact confirmed that nucleosome density is typically lower at functional TF binding sites compared to coding regions (Bernstein et al., 2004; Yuan & Liu, 2008).
Yuan et al. studied nucleosome positioning over yeast promoters and discovered a nucleosome-free region (NFR) flanked by well-positioned nucleosomes on both sides (known as the -1 and +1 nucleosomes), followed by an array of nucleosomes that packages each gene (Yuan et al., 2005). -1 and +1 positioned nucleosomes seem to determine the location of the TSS by recruiting components of the transcriptional machinery to the NFR. Even if they are tightly positioned, -1 and +1 nucleosomes can undergo changes that affect their stability, including histone replacement or modification of histone tails, which will ultimately promote eviction upon PIC assembly.

A similar model was proposed by He et al. to explain patterns of nucleosome positioning at enhancers based on the study of nucleosome occupancy at androgen receptor binding sites in a human prostate cancer cell line (He et al., 2010). They proposed that TF binding sites are flanked by a pair of well-positioned nucleosomes separated by 250-450 bp, while the binding site itself seems to be occupied by a much less stable nucleosome likely containing the H2A.Z histone variant. This central nucleosome can therefore be easily displaced, thereby facilitating access of TFs to those regulatory sites upon activation of transcription.

DNA sequences enriched in GC dinucleotides seem to promote nucleosome occupancy, while nucleosome-depleted regions tend to be enriched in poly (dA.dT) elements (Bernstein et al., 2004). However, nucleosome positioning can not only be explained by the underlying DNA sequence, as it is also strongly dependent on the catalytic action of chromatin remodelling complexes (Jiang & Pugh, 2009).

2.3 Families of ATP-dependent chromatin remodellers

ATP-dependent chromatin remodelling complexes can modify chromatin structure by disrupting histone-DNA contacts, which might alter the conformation of nucleosomes or mobilise them to a different location. By doing so, they thus allow or impede recruitment of transcriptional regulators to particular loci and, as such, can be involved in either activation
or repression of transcription, as well as other nuclear processes (Kingston & Narlikar, 1999).

Many different chromatin remodelers have been identified in eukaryotic cells. All remodelers require the energy of ATP hydrolysis to exert their functions, and so their catalytic subunits have evolved to contain a conserved ATPase domain. The ATPase core domain of chromatin remodelers contains a DNA-duplex-destabilising subdomain coupled to a DNA-translocating motor subdomain. Additional domains in each catalytic subunit can provide binding specificity and regulatory activity (Figure 4).

![Figure 4. Domain organisation of the catalytic subunits of the different remodeler families.](image)

All chromatin remodelling complexes contain a catalytic subunit defined by a conserved core ATPase domain divided in two parts: DEXDc and HELICc domains. Each catalytic subunit is further characterised by the presence of additional domains that modulate the catalytic activity or provide binding specificity to either nucleosomes or regulatory proteins. Taken from (Hota & Bruneau, 2016)

Most chromatin remodelers are large complexes with multiple accessory subunits that might directly regulate the catalytic activity of the core ATPase or provide binding specificity, both for targeting and recruitment of additional factors. In fact, each family of remodelling complexes is specialised to modify chromatin structure in a distinct and substrate-specific manner and might therefore play different biological roles (Hargreaves & Crabtree, 2011). A brief summary of what is known about each family of chromatin remodelers is provided below.
- **SWI/SNF family**

  SWI/SNF complexes were first identified in yeast and Drosophila. Mammalian SWI/SNF complexes were subsequently characterised and were usually found to be composed of eight or nine subunits, which can vary according to cell type. Each complex contains a catalytic subunit that is a homologue of the yeast SWI2/SNF2 ATPase, and five additional core members, which are also conserved. SWI/SNF remodelers have been shown to disorder nucleosome positioning, which can make chromatin available to DNA-binding factors.

- **INO80 family**

  INO80 complexes can slide nucleosomes along DNA or even catalyse eviction or exchange of particular histones. INO80 remodelers have been shown to regulate DNA repair at double-strand breaks by promoting nucleosome eviction and facilitating recruitment of repair factors. The INO80 family has also been proposed to play a role in the regulation of telomere length during replication.

- **CHD family**

  While there are at least nine ATPases known in mammalian cells, CHD1 is the only ATPase of the family present in yeast. In addition to a conserved ATPase domain, they all have N-terminal tandem chromodomains and other domains that provide a specialised function. For example, CHD1 is recruited to H3K4me3 sites of active transcription, where it might facilitate transcriptional elongation and splicing. CHD3 and CHD4, on the other hand, are incorporated into the Nucleosome Remodelling and Deacetylase (NURD) complex, where they cooperate with histone deacetylases to repress transcription.

- **ISWI family**

  The ISWI family was first identified in Drosophila through characterization of the NURF complex (Tsukiyama & Wu, 1995). Each ISWI complex contains a conserved ATPase subunit but is specialised by unique accessory proteins. ISWI remodelers catalyse the
sliding of nucleosomes along DNA, regulating their phasing and thus creating ordered nucleosomal arrays with uniform spacing, which may serve to restrict DNA accessibility and facilitate gene silencing.

3. ISWI family of ATP-dependent chromatin remodellers

Most of what is known about the ISWI family comes from studies done in *S. cerevisiae* and in *Drosophila*. In fact, the first component of the family to be characterised was the NURF complex, which was purified from *Drosophila* embryo extracts (Tsukiyama & Wu, 1995). ISWI complexes are composed of a catalytic subunit and one to three additional subunits. The composition of the main mammalian complexes, as well as their counterparts in *S. cerevisiae* and *Drosophila*, is shown in Table 1.

<table>
<thead>
<tr>
<th>Species</th>
<th><em>S. cerevisiae</em></th>
<th><em>Drosophila</em></th>
<th>Mammalian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex</td>
<td>ISW1a</td>
<td>ISW1b</td>
<td>ISW2</td>
</tr>
<tr>
<td>Catalytic subunit</td>
<td>Isw1</td>
<td>Isw2</td>
<td>ISWI</td>
</tr>
<tr>
<td>Accessory subunits</td>
<td>loc3</td>
<td>loc2/4</td>
<td>Itc1</td>
</tr>
<tr>
<td></td>
<td>Dpb4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dls1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. The ISWI family. Main components of the different ISWI complexes in yeast, Drosophila and mammalian cells. Adapted from (Hargreaves & Crabtree, 2011).

As shown in Table 1, there are two isoforms of the ISWI ATPase in mammalian cells: SMARCA5 and SMARCA1, which are homologues of the yeast Isw1 and Isw2, and *Drosophila* ISWI. It was first believed that the different accessory subunits would preferentially interact with only one of the two ATPases, thus forming some of the stable complexes which have been more extensively characterised. However, the complexity of the ISWI family is greater than previously thought, as it was recently reported that every accessory subunit can associate and function with both ATPases. In a study by Oppikofer et al., endogenous SMARCA5 and SMARCA1 were both co-purified with each accessory.
subunit (Oppikofer et al., 2017). They further showed that the newly characterised ISWI complexes form stable and direct interactions, and that they are also functional and can mobilise nucleosomes.

ISWI ATPases have intrinsic chromatin remodelling activity, which can be specifically regulated by different accessory subunits. Moreover, accessory subunits assemble around ATPases to provide binding specificity to each complex (He et al., 2008). BPTF, for example, has been shown to interact with sequence-specific TFs bound to enhancers and is also able to recognise specific histone modifications through its bromodomain and PHD domains (Li et al., 2006; Richart et al., 2016).

3.1 Structure and regulation of the ISWI ATPase

The domain architecture of the ISWI ATPase, which seems to be rather conserved in eukaryotes, has been extensively studied. The catalytic core consists of 2 ATPase lobes, whose catalytic activity is inhibited at the basal state (i.e. in the absence of a nucleosome) by the auto-regulatory domains AutoN (located at the N-terminal) and NegC (located at the C-terminal) (Clapier & Cairns, 2012). At the C-terminal there is also a HAND-SANT-SLIDE (HSS) domain, which is able to associate, depending on the conformational state, with extranucleosomal DNA or with the nucleosome core (Hota et al., 2013).

In the inactive conformation (Figure 5, left), one of the lobes of the ATPase domain is still able to recognise nucleosomal DNA. The engagement with a nucleosome triggers a major conformational change of the ATPase to an active state (Figure 5, right), where the H4 tail of the nucleosome binds to the other ATPase lobe at the binding site for AutoN and thus prevents the auto-inhibition by this regulatory domain. In this extended conformation, the HSS binds the DNA flanking the nucleosome and is able to ‘measure’ the length of the linker DNA. The DNA length-sensing function provided by the HSS domain is required to promote proper nucleosome repositioning and ensure that evenly spaced nucleosome arrays are generated upon translocation of the DNA. The translocation of DNA occurs
subsequent to another conformational change upon which the HSS domain binds to the nucleosome core, where it has also been proposed to enhance the catalytic activity of the ATPase by relieving the intrinsic inhibition from the regulatory domain NegC (Clapier et al., 2017; Harrer et al., 2018; Yan et al., 2016).

Mammalian ISWI ATPases can function both as monomers and as dimers. However, several studies have shown that remodelling activity is more efficient upon dimerization (Armache et al., 2019; Leonard & Narlikar, 2015). In a dimer, each protomer can bind opposite sides of the nucleosome and the two protomers can communicate via the HSS domain.

![Figure 5. Model for the regulation of ISWI ATPase activity. Schematic showing the functional domains of ISWI ATPases and their interplay with nucleosomal epitopes that promote conformational changes leading to their activation. Taken from (Dao et al., 2020)](image)

### 3.2 Biological functions of the ISWI complexes

As explained before, many chromatin-related processes depend on the accessibility of DNA, which can be regulated by chromatin remodelling complexes. The ISWI family of chromatin remodellers, in particular, has been linked to diverse biological functions, including DNA repair and regulation of transcription (Erdel & Rippe, 2011).

#### 3.2.1 Role of the ISWI ATPases in DNA repair

The stable knockdown of ISWI ATPases makes cells hypersensitive to DNA damage. Moreover, ISWI complexes have been shown to accumulate at sites of DNA damage,
where they seem to play an essential role not only in nucleosome repositioning but also in the recruitment of downstream effector proteins that mediate the repair process (Aydin et al., 2014; Erdel & Rippe, 2011). For instance, the ISWI catalytic subunit SMARCA5 has been reported to facilitate the recruitment of the repair factor CSB to UV-induced lesions associated with stalled transcription. In a study by Aydin OZ et al., the interaction between SMARCA5 and CSB appeared to be required for efficient repair of UV-induced damage and to promote transcription recovery (Aydin et al., 2014). After initial recruitment to DNA damage sites, SMARCA5 was shown to re-localise to the periphery of the lesion, which would be consistent with actual chromatin remodelling being catalysed by SMARCA5.

There is increasing evidence to show that remodelling complexes of the ISWI family are implicated in multiple pathways of the DNA damage response, from the repair of UV lesions mentioned above to the repair of double-strand breaks (Toiber et al., 2013), even if the different mechanisms involved are not as yet completely understood. It has become clear that ISWI chromatin remodellers play an important role in preserving DNA integrity and preventing genomic instability, which could lead to senescence, apoptosis, or even tumorigenesis if not resolved properly.

3.2.2 Role of ISWI ATPases in the control of transcription

Several studies have pointed to a repressive role of ISWI ATPases in the regulation of transcription. SMARCA5 is implicated in the positioning of the +1 nucleosome, which in turn promotes promoter-proximal pausing of RNA pol II in early transcriptional elongation (Jimeno-González et al., 2015). This function of ISWI seems to be conserved from yeast and Drosophila, since nucleosome positioning had previously been shown to be affected upon loss of different ISWI factors (Gkikopoulos et al., 2011; Sala et al., 2011).

Consistent with their role in +1 nucleosome positioning, S. cerevisiae ISWI ATPases are highly enriched at NFRs, where they have been shown to bind extended stretches of DNA adjacent to particular TF binding sites (Zentner et al., 2013). However, in the same study they found that Isw1 (but not Isw2) occupancy at NFRs and gene bodies was positively
associated with nucleosome turnover and transcriptional elongation rate, showing that ISWI complexes can play different (sometimes even opposing) roles in the control of transcription. In this case, Isw1 remodelling activity appeared to help re-establish proper chromatin structure subsequent to transient nucleosome disruption caused by RNA pol II transit.

The idea that distinct ISWI complexes can have different roles in the organisation of chromatin is supported by recent findings from Giles et al. They studied the binding profiles of the catalytic subunits of the main ATP-dependent chromatin remodellers in human cells and were able to classify them into two functional groups based on their preference for binding chromatin with particular epigenomic features, such as specific histone modifications. The ISWI catalytic subunit SMARCA5 was found to be enriched at ‘actively marked’ chromatin, while SMARCA1 was preferentially bound to ‘repressively marked’ regions (Giles et al., 2019).

In mammalian cells, SMARCA5 is reported to be crucial for determining the regularity of the nucleosome arrays adjacent to CTCF binding sites (Clarkson et al., 2019), where it acts to promote and maintain CTCF binding (Dluhosova et al., 2014; Wiechens et al., 2016). Depletion of CTCF or SMARCA5 affects the transcription of an overlapping cohort of target genes, which supports the functional significance of the action of SMARCA5 at CTCF sites (Wiechens et al., 2016). In differentiating myeloid cells, CTCF and SMARCA5 are both recruited to an enhancer that controls PU.1 expression, where they exert an enhancer-blocking activity and cooperatively repress transcription. Consequently, SMARCA5 also has an effect on the expression levels of PU.1-target genes, such as Cebpa, which was found to be upregulated upon SMARCA5 knockdown (Dluhosova et al., 2014).

Interestingly, SMARCA5 has been shown to be able to directly regulate cytokine gene expression. In a study performed using stimulated murine T cells, SMARCA5 appeared to be required for the induction of IL-3 expression, while the expression of IL-2 and other
cytokines seemed to be repressed by SMARCA5 (Precht et al., 2010). They showed that SMARCA5 was bound to the promoters of the regulated genes, but no further mechanistic details were reported. These observations might be in part explained by additional studies which have shown that SMARCA5 and SMARCA1 contribute to nucleosome organisation not only at CTCF sites but also at binding sites for many other TFs, such as JUN (Wiechens et al., 2016). In line with this, Barisic and colleagues identified a whole class of TFs which specifically depend on the remodelling activity of SMARCA5 for binding to their target sites (Barisic et al., 2019). The effects of SMARCA5 on nucleosomal phasing around TF binding sites appeared particularly evident at distal regulatory regions (i.e. enhancers). Not surprisingly, then, they observed global changes in transcription upon knockout of SMARCA5 in mouse embryonic stem cells, where genes affected were mostly related to developmental and proliferative functions.
The aim of this study was to investigate the role of ISWI chromatin remodellers in the regulation of the inflammatory transcriptional program in murine macrophages. The motivation of my study was to contribute to a better understanding of the transcriptional basis of inflammation by providing insight into the role of a specific type of chromatin remodellers in controlling how genomic information flows from cis-regulatory elements to machineries effecting changes in gene expression.

The specific aims of the study are detailed below:

1. To characterise the ISWI complexes present in primary macrophages by analysing the expression and possible interactions between the different subunits.

2. To assess, using unbiased genome-wide approaches, the impact of the depletion of ISWI ATPases on the transcriptional activity and epigenomic landscape of primary macrophages in basal conditions and following LPS stimulation.

3. To understand the specific roles of each ISWI complex by carrying out transcriptomic analyses of macrophages depleted of the individual accessory subunits and measuring the effects on the transcriptional activity of the cells.

4. To further dissect the function and possible mechanisms of action of ISWI chromatin remodellers by determining their genome-wide occupancy.

5. To study the selectivity of ISWI complexes in the regulation of transcriptional responses to stimuli other than LPS, such as IL4.
MATERIALS AND METHODS

1. Mouse biology

1.1 Strains

Wildtype (WT) C57BL/6JCrI mice were purchased from Charles River.

Smarca5\textsuperscript{lox/lox} mice (Alvarez-Saavedra et al., 2014) were kindly provided by T. Stopka (Charles University, Prague).

Mice were housed in specific pathogen-free conditions according to institutional guidelines.

1.2 Genotyping

Genomic DNA was extracted using the Nucleospin Tissue Kit (Macherey-Nagel) from WT and Smarca5\textsuperscript{lox/lox} BMDMs transduced with retroviral pMSCV-Cre expression plasmid as outlined in the next section.

Genomic PCR analysis was used to confirm proper deletion in the Smarca5 gene. The primers used are listed below:

- \textit{Smarca5 flox-F} \quad ACTGAGGACTCTGTGAAACAGTCAAG
- \textit{Smarca5 flox-R} \quad TACACAACTAAGGCAGTGGGTTATAGTGC
- \textit{Smarca5 del5-F} \quad GTGCAAAGCCAGAGACGATGGTATG

2 Cell culture

2.1 Isolation, growth conditions and treatment of bone marrow-derived macrophages (BMDMs)

Primary macrophages were derived from the bone marrow of C57BL/6JCrI or Smarca5\textsuperscript{lox/lox} mice. Hematopoietic progenitor cells were isolated and plated for
differentiation in BM medium (Dulbecco’s Modified Eagle’s Medium (DMEM) supplemented with 20% endotoxin-free Fetal Bovine Serum from North American origin (FBS-NA), 30% L929-cell conditioned medium, 1% glutamine, 1% penicillin-streptomycin (Pen/Strep), 0,5% sodium pyruvate and 0,1% β-mercaptoethanol). BMDMs were differentiated for 6 days before being subjected to either treatment or lentiviral/retroviral infection.

For stimulation of the inflammatory response, BMDMs were treated with E. coli lipopolysaccharide (LPS) serotype EH100 (Enzo Life Sciences) or Interleukin-4 (IL4) (R&D Systems) at 10 ng/ml for either 2 hours or 4 hours.

2.2 Cell lines: growth conditions

293T (ATCC® CRL-3216™), Phoenix-ECO (ATCC® CRL-3214™), and RAW264.7 (ATCC® TIB-71™) cells were grown in DMEM supplemented with 10% FBS from South American origin, 1% glutamine and 1% Pen/Strep.

2.3 Plasmids

Five MISSION® shRNAs (Sigma Aldrich) targeting each ISWI subunit were tested. The most efficient were selected for further knockdown experiments and are listed below. A luciferase-targeting shRNA (shLUC) was used as control.

<table>
<thead>
<tr>
<th>shRNA</th>
<th>Clone</th>
</tr>
</thead>
<tbody>
<tr>
<td>shSMARCA5-1</td>
<td>TRCN0000084430</td>
</tr>
<tr>
<td>shSMARCA5-2</td>
<td>TRCN0000084431</td>
</tr>
<tr>
<td>shBPTF</td>
<td>TRCN000238663</td>
</tr>
<tr>
<td>shBAZ1A</td>
<td>TRCN000244322</td>
</tr>
<tr>
<td>shBAZ1B</td>
<td>TRCN0000108937</td>
</tr>
</tbody>
</table>

Table 2. List of shRNAs used in this study.
For retroviral transduction of Cre recombinase, the pMSCV-Cre vector kindly provided by B. Amati (IEO) was used.

2.1 Lentiviral infection

293T cells at 50% confluence were transiently transfected using calcium phosphate for production of infectious lentiviral particles. For each 10 cm plate the following mix was prepared (amounts were proportionally scaled down in case of transfection in smaller formats):

- 7,5 µg of lentiviral vector
- 6 µg of psPAX2 packaging plasmid (Addgene #12260)
- 2,5 µg of VSV-G envelope plasmid (Addgene #12259)
- 62,5 µl of CaCl₂
- H₂O up to 500 µl

The mix was added dropwise, while making bubbles, into 500 µl of 2X HEPES buffered saline, incubated for 10 min at room temperature and then added dropwise to the borders of each plate of 293T cells. 6-8 h after transfection, medium was replaced with 6 ml of fresh medium.

Viral supernatants were collected after 24 and 48 hours (new medium was added after collecting supernatants at 24 h), filtered through a 0,45 µm filter and purified or concentrated 2X (2 plates of 293T to infect 1 plate of BMDMs) by incubating overnight at 4 °C with the Lenti-X concentrator (Clontech) according to the manufacturer’s instructions. Overnight incubation was followed by a centrifugation at 1500 g for 45 min at 4 °C. Virus-containing pellets were then resuspended in BM medium and transferred to BMDMs at days 6 and 7 post-isolation. 8 h after the second round of infection, medium was replaced with fresh BM medium containing 4 µg/ml of puromycin for selection of infected cells.
Puromycin was refreshed 48 h later and BMDMs were maintained in culture for 24 h more before harvesting.

2.5 Retroviral infection

Phoenix-ECO producer cells at 50% of confluence were transiently transfected to generate infectious retroviruses. Cells were transfected using calcium phosphate as outlined before. The amounts of DNA used to transfect each 10 cm plate were: 10 µg for the retroviral plasmid (pMSCV) and 5 µg for the pCL-ECO packaging plasmid.

2 plates of Phoenix-ECO cells were transfected for each plate of BMDMs that had to be infected. 6-8 h after transfection, medium of producer cells was replaced with 8 ml of BM medium. Viral supernatants were collected 48 and 72 h later, filtered through a 0.45 µm filter and transferred to BMDMs in the presence of polybrene to increase transduction efficiency. Each plate of BMDMs was subjected to 4 rounds of infection. Rounds 1 and 3 were performed on the morning of days 1 and 2 post-isolation of BMDMs using viral supernatants from one plate of Phoenix-ECO, while 2nd and 4th rounds were performed in the afternoon using viral supernatants from another plate of producer cells. 3-4 hours after the last round of infection, medium was replaced with fresh BM medium supplemented with puromycin at 4 µg/ml for selection of infected cells. Puromycin was refreshed 48 h later and cells were kept in culture for another 24-48 hours before harvesting.

3 Molecular biology

3.1 Protein extraction and western blotting

Cells were harvested and washed twice in ice-cold 1X PBS, then lysed for total protein extraction in s300 buffer (50mM Tris-HCl pH 7.6, 0.2% Nonidet P-40 (NP-40), 300 nM NaCl, 10% glycerol) supplemented with a protease inhibitor cocktail (Roche), and cleared by centrifugation at 13000 rpm for 10 minutes. The protein concentration was determined by Bradford assay (Bio-Rad) following the manufacturer’s protocol.
Samples were subjected to electrophoresis in 6-10% polyacrylamide gels. Proteins were resolved under reducing conditions according to molecular weight and then transferred to nitrocellulose membranes, which were incubated with 5% non-fat milk in TBST buffer (1X Tris Buffer Saline, 0.1% Tween20) to block unspecific sites. The antibodies and corresponding dilution factors are listed below:

<table>
<thead>
<tr>
<th>Target</th>
<th>Antibody</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMARCA5</td>
<td>Bethyl A301-018A</td>
<td>1:1000</td>
</tr>
<tr>
<td>BPTF</td>
<td>Millipore ABE24</td>
<td>1:500</td>
</tr>
<tr>
<td>BAZ1A</td>
<td>Bethyl A301-318A</td>
<td>1:500</td>
</tr>
<tr>
<td>BAZ1B</td>
<td>Abcam ab50850</td>
<td>1:2000</td>
</tr>
<tr>
<td>CEBPB</td>
<td>Bethyl A302-738A</td>
<td>1:2000</td>
</tr>
<tr>
<td>STAT1</td>
<td>Cell Signaling 9172L</td>
<td>1:1000</td>
</tr>
<tr>
<td>pSTAT1 (Tyr701)</td>
<td>Cell Signaling 7649S</td>
<td>1:1000</td>
</tr>
<tr>
<td>Vinculin</td>
<td>Sigma V9131</td>
<td>1:10000</td>
</tr>
<tr>
<td>Tubulin</td>
<td>Santa Cruz sc32293</td>
<td>1:1000</td>
</tr>
</tbody>
</table>

Table 3. List of antibodies and dilutions used for Western Blot analysis.

Following hybridization with primary antibodies, membranes were incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies and subsequently developed with the ECL system using a Chemidoc (Bio-Rad) to acquire images.

3.2 Co-immunoprecipitation analysis

Immunoprecipitation was performed using nuclear extracts. To prepare nuclear lysates, cells were harvested and washed twice in ice-cold 1X PBS. Cell pellets were lysed in a 2-step protocol. Cell pellets were resuspended in a lysis buffer containing 50 mM Tris-HCl pH 8, 2 mM EDTA, 0.1% NP-40 and 10% glycerol, and supplemented with a protease
inhibitor cocktail (Roche). Lysates were incubated for 10 min at 4 °C and cleared by centrifugation at 3000 rpm for 5 min. The supernatant contained the cytosolic fraction. To extract the nuclear fraction, pellets were resuspended in 3-4 times their volume with a nuclear extraction buffer (250 mM NaCl, 50 mM Tris-HCl pH 8, 0,5 mM EDTA, 0,5 mM EGTA, 0,2% NP-40) supplemented with protease inhibitors and PMSF. Lysates were incubated for 10 min on ice and cleared by centrifugation at 13000 rpm for 10 min.

Protein G magnetic beads (Invitrogen) were pre-blocked with blocking buffer (0,5% BSA in 1X PBS) and bound to 5-10 µg of primary antibody by incubating for 6 h at 4 °C on a rotating platform. Beads were washed twice in blocking buffer and subsequently added to the previously quantified nuclear lysates for overnight incubation at 4 °C on a rotating wheel. The following day, beads were washed 6 times with the nuclear extraction buffer and immunoprecipitated proteins were eluted in SDS loading buffer (0,25% Bromophenol blue, 50% Glycerol, 10% SDS, 0,25 M Tris-HCl pH 6,8, 0,5M DTT) by boiling samples at 95 °C for 3 min and removing beads by incubation on a magnetic stand. Proteins were then analysed by Western Blot as previously described.

3.3 Chromatin immunoprecipitation

BMDMs were fixed with formaldehyde at either 1% (for H3K27ac and CEBPB ChIP) or 2% (for SMARCA5 and BPTF ChIP) for 10 min at room temperature, followed by quenching with 125 mM or 250 mM Tris-HCl pH 7,6 respectively. Fixed cells were washed with 1X PBS, harvested by scraping, pelleted, and then lysed to prepare nuclear extracts. Chromatin was fragmented either by sonication or Mnase digestion.

Lysates with fragmented chromatin were cleared by centrifugation at 13000 rpm for 10 min (4 °C) and incubated overnight with protein G Dynabeads (Invitrogen) previously blocked with 1X PBS/0,5% BSA and coupled with 5-10 µg of antibody. Following immunoprecipitation, beads were washed six times with a modified RIPA buffer (50 mM Hepes pH 7,6, 500 nM LiCl, 1 mM EDTA, 1% NP-40, 0,7% Na-Deoxycholate) using a
magnetic stand, and one final wash was carried out with 1X TE containing 50 mM NaCl. Immunoprecipitated chromatin was eluted in 1X TE/2% SDS and decrosslinked by overnight incubation at 65 °C. DNA was purified using the PCR purification kit (Qiagen).

Primers used for ChIP validation by qPCR analysis are reported in Table 4.

The following antibodies were used: SMARCA5 (Abcam ab3749), BPTF (Millipore ABE24), H3K27ac (Abcam ab4729), CEBPB (Bethyl A302-738A).

3.4 RNA extraction and cDNA synthesis

RNA was extracted using the Quick-RNA Miniprep kit (Zymo Research) according to the manufacturer’s instructions. 50µl of Dnase/Rnase-free water were used to elute the RNA.

RNA concentration was determined using a NanoDrop spectrophotometer (Thermo Fisher). Samples were used for either RT-qPCR or RNA-seq analysis.

To prepare cDNA for RT-qPCR analysis, 200-400 ng of total RNA was retrotranscribed. ImProm-II Reverse Transcriptase (Promega) was used for cDNA synthesis following the manufacturer’s protocol.

3.5 Quantitative real-time PCR

qPCR analyses were conducted on a Viia7 Real-Time PCR machine using SYBR Green Master Mix (Applied Biosystems, Thermo Fisher). Data analysis was performed on the Thermo Fisher Cloud platform.

Primers for mRNA expression were designed spanning an exon junction. mRNA levels were normalised to TBP expression using the ΔΔCt method.

Primer sequences both for expression and ChIP validation are provided below.

<table>
<thead>
<tr>
<th>Target</th>
<th>forward</th>
<th>reverse</th>
<th>Expression/ChIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBP</td>
<td>CTGGAATTGTACCACGCTT</td>
<td>ATGATGACTGCAGAACATCG</td>
<td>Expression</td>
</tr>
</tbody>
</table>
4 Next-generation sequencing

4.1 ATACseq

Prior to starting the protocol, Tn5 transposase (produced in-house) was loaded with 0.148 volumes of a 100 µM equimolar mixture of pre-annealed Tn5MEDS-A and Tn5MEDS-B sequencing adapters by incubating for 1 h at room temperature.

50x10^3 cells were lysed in 100 µl of ice-cold lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.1% NP-40). Nuclei were pelleted by centrifugation at 500 g for 20 minutes at 4 °C. Supernatants were then discarded and pellets of nuclei were resuspended in 25 µl of a tagmentation buffer (1 µl adapter-loaded Tn5 transposase and 5 µl of 5X transposase buffer (50 mM Tris-HCl pH 8.5 and 25 mM MgCl2)) and incubated for 1 h at 37 °C. Reactions were quenched by adding 9 µl of a buffer containing 2 µl of 5% SDS, 2 µl of Proteinase K (20 mg/ml) and 5 µl of cleanup buffer (900 mM NaCl, 30 mM EDTA), followed by an incubation at 40 °C for 30 min. Tagmented DNA was purified using SPRI beads at a ratio of 2X volumes.

Libraries were generated by amplifying fragments of tagmented DNA in a PCR reaction. At the same time, samples were barcoded with dual-index sequencing adapters. KAPA
HiFi HotStart ready mix was used (activated by incubating at 95 °C for 1 min) and 4 µl of sequencing primers for each reaction. PCR conditions were as follows:

- 72 °C 5 min
- 98 °C 2 min
- 98 °C 20 s
- 63 °C 30 s
- 72 °C 1 min
- 4 °C on hold
- 14 cycles

Subsequently, fragments shorter than 500 bp were isolated using SPRI beads at a ratio of 0,65X, followed by purification with SPRI beads at a ratio of 1,8X. DNA concentration from final libraries was measured with a Glomax luminometer (Promega) using the Quanti Fluor assay (Promega). Libraries were then diluted to 2 ng/µl and quality was analysed using a TapeStation (Agilent Technologies) before sequencing on an Illumina NextSeq 500 platform.

### 4.2 RNAseq

RNA quality was assessed on a TapeStation (Agilent Technologies). 4 ng of total RNA from each sample were used to prepare libraries following the SMART-seq2 protocol (adapted from Picelli et al., 2014). Briefly, polyadenylated mRNA was selected using oligo(dT) primers and subsequently retrotranscribed. The resulting cDNA was pre-amplified, purified with SPRI beads at a 1X ratio, and tagmented with a Tn5 transposase (in-house production) previously loaded with sequencing adapters as explained before. cDNA fragments obtained after tagmentation were amplified and ligated to dual-index sequencing primers in a PCR reaction.

Adapter-ligated libraries were purified with SPRI beads at a 1X ratio, quantified at a Glomax luminometer (Promega) using the QuantiFluor assay, and diluted to 2 ng/µl.
TapeStation (Agilent Technologies) was used for a quality check of the final cDNA libraries, which were then sequenced on an Illumina NextSeq 500 platform.

4.3 ChIPseq

ChIP was performed as described above. ChIP DNA was quantified by QuantiFluor assay (Promega) at a Glomax luminometer according to the manufacturer’s protocol. 1-5 ng of each sample was used to generate libraries for sequencing following standard methods (adapted from Blecher-Gonen et al., 2013). Briefly, samples were processed through subsequent enzymatic treatments of end-repair, dA-tailing, and ligation to sequencing adapters. Adapter-ligated libraries were amplified by PCR and subsequently purified using SPRI beads. Final DNA libraries were quantified and diluted to 2 ng/µl before sequencing on an Illumina NextSeq 500 platform.

5 Computational methods

5.1 RNAseq data analysis

Reads obtained from Illumina NextSeq 500 platform were quality filtered and subsequently mapped to the mouse mm10 (GRCm38) reference genome using TopHat v2.1.1 (Kim et al., 2013). Reads mapping on exons of annotated genes were counted with featureCounts v1.5.2 (Liao et al., 2014).

Differentially expressed genes (DEGs) were identified using the EdgeR R-package (Robinson et al., 2009) and filtered by FDR<0.01, Log₂FC>1 and RPKM>0.1 (in at least three biological replicates).

Functional enrichment analysis was performed using the DAVID tool v6.8 (D. W. Huang et al., 2009). A set of non-differentially-expressed genes was used as background. Top-ranking Biological Process categories associated with each group of DEGs were selected, and significance is shown as -log10 of the p-values.
Motif enrichment analysis on the promoters of DEGs was performed using the PSCAN tool (Zambelli et al., 2009). For this analysis, the promoters were considered as the regions 750 bp upstream and 250 bp downstream the TSS. The promoters of genes not identified as differentially expressed were used as background. Significantly enriched motifs identified were ranked according to their p-values.

5.2 ChIPseq data analysis

Reads obtained from Illumina NextSeq 500 platform were quality filtered, trimmed, and subsequently mapped to the mouse mm10 (GRCm38) reference genome using Bowtie2.3.4 (Langmead & Salzberg, 2012). Reads with low mapping quality and PCR duplicates were removed using SAMtools (Li et al., 2009). Peak calling was performed using SICER (Zang et al., 2009). Regions with an enrichment lower than 2-fold compared to input genomic DNA were discarded.

SMARCA5 ChIP-seq regions were subjected to functional enrichment analysis. Enriched regions (i.e. peaks identified by SICER) were analysed using GREAT v4.0.4 (McLean et al., 2010) with default parameters. The whole mm10 genome was used as background. Top-ranking categories were selected, and significance is shown as -log10 of the p-values.

A motif enrichment analysis was performed using the online tool PSCAN-ChIP v1.3 (Zambelli et al., 2013) with default settings and using a mixed background built with a random selection of genomic regions from different cell types. Significantly over-represented and under-represented motifs in the central 150 bp of SMARCA5-bound regions were identified and ranked according to their p-values.

5.3 ATACseq data analysis

Reads obtained from Illumina NextSeq 500 platform were quality filtered, trimmed, and subsequently mapped to the mouse mm10 (GRCm38) reference genome using Bowtie2.3.4 (Langmead & Salzberg, 2012). Reads with low mapping quality, PCR
duplicates and reads mapping to chrM were removed using SAMtools (Li et al., 2009). Peak calling was performed using MACS2 (Zhang et al., 2008). Peaks corresponding to regions blacklisted by the ENCODE Project Consortium were discarded (Dunham et al., 2012).

Differentially accessible regions in two biological replicates of depleted vs. control cells were identified using the EdgeR R-package (Robinson et al., 2009) and filtered by FDR<0.01, Log₂FC>1 and RPKM>0.5.

Functional enrichment analysis for differentially accessible regions was performed using the GREAT online tool v4.0.4 (McLean et al., 2010) with default settings. Non-changing regions were used as background. Top-ranking categories associated with regions of increased or decreased accessibility were selected, and significance is shown as -log10 of the p-values.

To identify over-represented motifs corresponding to known binding sites for transcription factors, a motif enrichment analysis was performed using PSCAN-ChIP (Zambelli et al., 2013) on the 200 bp around the summit of differentially accessible regions previously identified. Non-changing regions were used as background. Significantly enriched motifs were ranked according to their p-values.
RESULTS

1. Characterisation of the ISWI complexes in BMDMs

To characterise the different ISWI complexes present in murine macrophages, the expression at the RNA level of the following ISWI subunits was analysed both at the basal state and at different time-points after LPS stimulation (2 h and 4 h): Smarca1, Smarca5, Baz1a, Baz1b and Bptf. Analysis of RNA-seq data previously generated in the lab revealed that all ISWI subunits except Smarca1 were expressed in macrophages, some of which were dynamically regulated in response to LPS treatment (Figure 6A). More specifically, the expression of Baz1a appeared to be strongly induced at 2 h and high expression levels were maintained at 4 h post-induction. The expression of Baz1b, on the other hand, was transiently repressed at 2 h but basal levels seemed to be restored already at 4 h after LPS stimulation.

The above results were subsequently validated at the protein level by Western Blot analysis. In this case, a strong upregulation of the subunit BAZ1A was seen at 4 h post-stimulation with LPS, concomitant with a certain degree of repression of the BPTF subunit (Figure 6B).

The ISWI catalytic subunits SMARCA5 and SMARCA1 are reported to be interchangeable and form functional complexes with any of the ISWI accessory subunits (Oppikofer et al., 2017). To determine whether the BPTF subunit could still form a complex in macrophages even if SMARCA1 was not expressed, a possible physical interaction with SMARCA5 was explored. Co-immunoprecipitation (co-IP) was performed using a specific antibody against SMARCA5. As shown in Figure 6C, BPTF was co-immunoprecipitated with SMARCA5 as detected by Western Blot analysis. This confirms that the accessory subunit BPTF and the catalytic subunit SMARCA5 are part of the same complex, which might substitute for the NURF complex, thus making SMARCA5 the common catalytic subunit of all ISWI complexes in BMDMs.
Figure 6. Characterisation of the ISWI complexes in bone marrow-derived macrophages. (A) mRNA expression profile of ISWI subunits at the basal level and upon LPS stimulation (2 h and 4 h). Reads per kilobase million (RPKM) values were retrieved from RNA-seq data previously generated in the lab. (B) Western Blot analysis showing protein levels of ISWI subunits at different time-points. A-Tubulin levels are shown as loading control. (C) Western Blot analysis showing interaction between SMARCA5 and BPTF as identified by SMARCA5 IP in total protein lysates. IgG IP is shown as control.

2. Impact of SMARCA5 depletion on the transcriptional activity of macrophages

To understand the general role of ISWI chromatin remodellers in the regulation of the inflammatory gene expression program, I first focused on the study of the common catalytic subunit SMARCA5. shRNA-mediated knockdown by lentiviral transduction in BMDMs was used. Two distinct shRNAs showing the highest efficiency were selected out of a set of five and were subsequently used to knockdown SMARCA5 both in unstimulated cells and in cells subjected to LPS treatment (2 h and 4 h). BMDMs transduced with a Luciferase-targeting shRNA (shLuc) were used as control. Efficient depletion of SMARCA5 in three replicates was confirmed by RT-qPCR analysis (Figure 7).
Figure 7. Evaluation of the knockdown efficiency of the two SMARCA5-targeting shRNAs used in this study. RT-qPCR analysis showing effective depletion of Smarca5 at the mRNA level. Transcript levels were normalised against Tbp. Data are expressed as the mean±SD (n=3).

To determine the transcriptomic profile of control and SMARCA5-depleted cells, polyadenylated RNA-seq analysis was performed. Libraries for RNA-seq were prepared using SMART-seq technology and sequenced on an Illumina NextSeq 500 platform as previously described (Picelli et al., 2014). The obtained reads were subsequently mapped to the mouse genome and counted when falling on exons.

Differential expression analysis was performed by comparing three replicates of control and SMARCA5-depleted cells using the tool EdgeR (Robinson et al., 2009). This analysis was first carried out using data generated with the most efficient shRNA (shSMARCA5-1) (Figure 7). Results obtained from this first analysis are reported in Figure 8 and are explained in detail below. However, it should be noted that to exclude off-target effects these results were subsequently validated with additional analysis using the data produced with the second shRNA (shSMARCA5-2), even if the magnitude of part of the effects was reduced to some extent due to the lower efficiency of this shRNA.

A total of 924 genes were identified as differentially expressed with a log₂FC>1 and an FDR<0.01 in at least one of the time-points (0 h, 2 h, 4 h LPS). Differentially expressed genes (DEGs) were hierarchically clustered using the Pearson’s correlation distance and the complete-linkage method, and their transcriptional profiles were subsequently plotted in the heatmap shown in Figure 8A. The exact numbers of upregulated and downregulated genes identified at each time-point are specified in the table in Figure 8B. Notably, there
were more downregulated than upregulated genes upon loss of SMARCA5, particularly at the basal level.

To investigate whether specific biological functions were associated with either upregulated or downregulated genes, a gene ontology analysis on each specific cluster of DEGs was carried out using the DAVID tool (Figure 8A, right) (Huang et al., 2009). Interestingly, several categories related to interferon signalling were identified as the most significantly enriched among those genes downregulated, as illustrated by the presence of the specific category ‘response to interferon-beta’. The enrichment for the terms identified among upregulated genes, which included very heterogeneous categories such as ‘transport’ or ‘cofactor metabolic process’, was not as strong.

In addition, a motif enrichment analysis on the promoters of DEGs (from -750 to +250 of the TSS) was performed using the PSCAN tool (Zambelli et al., 2009). The promoters of genes whose expression was not found to be dependent on SMARCA5 were used as background for this analysis. The top-ranking motifs identified for each group of DEGs are shown in Figure 8C. Consistent with the results obtained in the gene ontology analysis, we found a significant enrichment in binding sites for many interferon regulatory factors (IRFs) at the promoters of downregulated genes. On the other hand, over-represented motifs among promoters of upregulated genes were not so obviously associated with the categories previously identified in the gene ontology analysis, and the significance was also much lower compared to the results on downregulated genes.
Figure 8. RNA-seq analysis reveals a number of differentially expressed genes (DEGs) upon SMARCA5 knockdown. (A) Left, heatmap showing hierarchically clustered DEGs (FDR<0.01). Right, top gene ontology categories associated with the genes in each cluster. (B) Table showing number of DEGs at each time-point. (C) Table showing the top significantly enriched motifs identified by PSCAN analysis on the promoters of each group of DEGs.

Taken together, the results outlined above all point to a role of SMARCA5 in the regulation of interferon-stimulated genes. Therefore, I set out to identify the players that might mediate the association of SMARCA5 with the interferon response. As shown in Figure 8C, binding sites for IRF8 were the most significantly enriched at the promoters of...
genes downregulated upon SMARCA5 depletion. Importantly, motifs for most other IRFs were also ranked at the top of the list with very low p-values, most likely due to the high degree of similarity between the binding sites for distinct IRFs (Table 5).

<table>
<thead>
<tr>
<th>Motif</th>
<th>Matrix</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF8</td>
<td></td>
<td>2.19E-27</td>
</tr>
<tr>
<td>IRF9</td>
<td></td>
<td>5.92E-26</td>
</tr>
<tr>
<td>STAT1</td>
<td></td>
<td>1.41E-25</td>
</tr>
<tr>
<td>IRF4</td>
<td></td>
<td>1.17E-24</td>
</tr>
<tr>
<td>STAT1::STAT2</td>
<td></td>
<td>1.27E-24</td>
</tr>
<tr>
<td>IRF3</td>
<td></td>
<td>3.23E-24</td>
</tr>
<tr>
<td>IRF5</td>
<td></td>
<td>1.82E-20</td>
</tr>
<tr>
<td>IRF7</td>
<td></td>
<td>8.05E-20</td>
</tr>
<tr>
<td>IRF2</td>
<td></td>
<td>1.91E-19</td>
</tr>
<tr>
<td>STAT2</td>
<td></td>
<td>2.36E-17</td>
</tr>
<tr>
<td>IRF1</td>
<td></td>
<td>6.04E-16</td>
</tr>
</tbody>
</table>

Table 5. Motif enrichment analysis on the promoters of genes downregulated upon SMARCA5 depletion.

We hypothesised that specific IRFs involved in the regulation of the interferon response mediated by SMARCA5 might also be differentially expressed upon SMARCA5 depletion. Thus, the levels of expression of each *Irf* gene in the RNA-seq dataset were carefully checked. As shown in Figure 9A, *Irf7* mRNA levels were very strongly downregulated in SMARCA5-depleted cells, while the expression of other *Irfs* was not significantly affected by knockdown of SMARCA5. The downregulation of *Irf7* was validated by RT-qPCR in two independent experiments (Figure 9B).

If the regulation of the interferon response by SMARCA5 were in part mediated by IRF7, one would expect that similar sets of genes would be impacted by a loss of either factor.
Opportunely, an RNA-seq dataset in which IRF7 had been knocked down had previously been generated in the Natoli lab. We took advantage of these data and crossed both datasets to identify overlapping genes that are both dependent on SMARCA5 and IRF7. In line with the results outlined above, 62% of genes downregulated upon SMARCA5 depletion were also IRF7-dependent genes (Figure 9C), while only 33% of the upregulated were dependent on IRF7.

Finally, we used publicly available data (Mancino et al., 2015; Tong et al., 2016) on the genomic occupancy of several IRFs (specifically IRF1, IRF3, IRF7 and IRF8) to check if the promoters of genes identified as commonly regulated by SMARCA5 and IRF7 overlapped with IRF binding in basal conditions. The number of promoters of genes commonly regulated by SMARCA5 and IRF7 which overlap with binding of each of the IRFs analysed, is shown in Figure 9D. Interestingly, over 80% of the promoters analysed were bound by at least one IRF. In contrast, there was very little overlap between upregulated genes and IRF binding (not shown). A representative example of an interferon-stimulated gene downregulated upon SMARCA5 depletion which overlaps with IRF1 and IRF8 binding at the promoter (Mancino et al., 2015), is shown in Figure 9E.

Collectively, these data show that interferon-stimulated genes are strongly dependent on SMARCA5, and that IRF7 may mediate the action of SMARCA5 on these genes.
Figure 9. A fraction of SMARCA5-dependent genes is strongly associated with the interferon response. (A) mRNA expression of all Irf genes in control and SMARCA5-depleted cells at the basal state (UT). RPKM values were retrieved from the RNA-seq experiments in Figure 7 and are plotted as mean±SD (n=3). (B) RT-qPCR analysis showing Irf7 mRNA downregulation upon SMARCA5 depletion. Transcript levels were normalised against Tbp. Data are expressed as the mean±SD (n=2). (C) Pie chart showing the proportion of SMARCA5-downregulated genes dependent on IRF7 as identified by RNA-seq. Data regarding IRF7-dependent genes had been previously generated in the lab. (D) Venn diagram showing the number of genes commonly regulated by SMARCA5 and IRF7 with IRF binding at their promoters identified by ChIP-seq analysis. (E) Snapshot showing a cluster of interferon genes downregulated upon SMARCA5 depletion and with IRF binding at their promoters.
3. Impact of SMARCA5 depletion on the epigenomic landscape of macrophages

The above data suggested a role of SMARCA5 in the regulation of interferon genes. This led to the hypothesis that SMARCA5 might be controlling the chromatin state at IRF binding sites, possibly through cooperation with IRF factors. To elucidate the mechanisms involved, I repeated lentiviral infections to study the impact of SMARCA5 depletion on the chromatin landscape using the most efficient shRNA. Since I had found limited effects at 4 h, I decided to include only the basal and 2 h time-points in the subsequent analyses. Efficient depletion of SMARCA5 in two independent experiments was confirmed by RT-qPCR (Figure 10A) and Western Blot analysis (Figure 10B).

Figure 10. Evaluation of knockdown efficiency of SMARCA5-targeting shRNA. (A) RT-qPCR analysis showing effective depletion of Smarca5 at the mRNA level. Transcript levels were normalised against Tbp. Data are expressed as the mean±SD (n=2). (B) Western Blot analysis showing SMARCA5 protein levels in control and depleted cells. Vinculin levels are shown as loading control.

ATAC-seq analysis was performed to identify changes in chromatin accessibility. As with RNA-seq analysis, ATAC-seq libraries were sequenced on an Illumina NextSeq 500 platform. The obtained reads were subsequently mapped to the mouse genome and counted with the MACS2 tool (Zhang et al., 2008) for peak calling. Differentially accessible sites were then identified by comparing the two replicates of control and SMARCA5-depleted cells using the tool EdgeR for statistical analysis (Robinson et al., 2009).

A total of 3377 differentially accessible sites upon SMARCA5 knockdown were identified (log2FC>1 and FDR<0.01), 2617 of which gained accessibility and 760 lost accessibility in
at least one of the time-points. As shown in Figure 11A, the overlap between regions identified at the basal level and at 2 h was in fact quite small. Figure 11B shows hierarchically clustered differentially accessible regions. In contrast to the effects of SMARCA5 depletion observed on the transcriptional activity of macrophages, where a large number of genes were found to be downregulated, accessibility was increased upon SMARCA5 knockdown in about three times more regions than it was decreased.

To investigate whether specific biological functions were associated with regions that either gained or lost accessibility, a gene ontology analysis was carried out using the GREAT tool, which assigns biological functions to sets of genomic regions based on the annotations of nearby genes (McLean et al., 2010). Several categories related to the interferon response were identified as the most significantly enriched among genes assigned to regions with decreased accessibility (Figure 11C, highlighted in red). The signatures associated to regions with increased accessibility showed similar enrichment but, again, appeared to refer to more general functions, including ‘response to LPS’ or ‘regulation of the inflammatory response’.

Subsequently, a motif enrichment analysis focused on ATAC-seq peaks (200 bp around the summit) was performed using the PSCAN-ChIP tool (Zambelli et al., 2013). A set of regions in which accessibility remained unchanged upon SMARCA5 knockdown was used as background for this analysis. Over-represented motifs identified in regions with increased or decreased accessibility are shown in Figure 11D. Regions that gained accessibility were significantly enriched in binding sites for factors of the C/EBP and AP-1 families, while the most significantly enriched motifs in regions that lost accessibility belonged to the POU family. However, following the results from the functional enrichment analysis, I was particularly interested in the enrichment of IRF motifs, which was also significant, as shown in Figure 11D, even if not ranked at the top of the list. These results support the idea that SMARCA5 can control the accessibility of chromatin at IRF binding sites.
Figure 11. ATAC-seq analysis unveils genome-wide effects of SMARCA5 depletion on chromatin accessibility. (A) Venn diagram showing the number of regions with increased (top) or decreased (bottom) chromatin accessibility (FDR<0.01). (B) Heatmap showing hierarchically clustered regions with a gain or loss in accessibility. (C) Top-ranking gene ontology categories associated with genes assigned to regions with increased (top) or decreased (bottom) accessibility as identified by GREAT analysis. Interferon-related signatures are highlighted in red. (D) Motif enrichment analysis on differentially accessible sites. (E) Pie charts showing the genomic distribution of sites with increased (left) or decreased (right) chromatin accessibility.

In addition, the genomic distribution of sites with increased or decreased accessibility was determined according to the following annotation criteria: promoter (+/- 2.5 kb), intragenic (> 2.5 kb) or extragenic (< -2.5 kb). As shown in Figure 11E, a similar small fraction of regions with increased or decreased accessibility overlapped with promoters. Differentially accessible sites more frequently mapped to intragenic regions, most likely pointing to changes in chromatin accessibility associated with the transcriptional wave. Such effects were more evident in regions that lost accessibility compared to regions with
increased accessibility, where the fraction of sites mapping to intragenic regions was smaller. However, it remains to be determined whether the differences in the genomic distribution of sites that gain or lose accessibility might be connected to the regulation of specific gene categories. Moreover, it will be interesting to see if regions with different annotations are enriched in binding sites for particular TFs which might not have been identified in the previous motif enrichment analysis, where as outlined above all regions with increased or decreased accessibility were considered without regard to how they are genomically distributed.

To further characterise the impact of SMARCA5 depletion on the chromatin landscape, H3K27Ac Mnase ChIP-seq was performed, which would provide information on the activity of enhancers. ChIP-seq libraries were prepared following standard methods used in the lab and sequenced on an Illumina NextSeq 500 platform. Obtained reads were mapped to the mouse genome.

A global analysis of H3K27Ac ChIP-seq data revealed genome-wide effects on enhancer activity. Regions with statistically significant changes in acetylation levels (log2FC>1 and FDR<0.01) either after LPS stimulation or SMARCA5 depletion were identified and hierarchically clustered, and signal intensity was subsequently plotted in the heatmap shown in Figure 12A. Metaprofiles showing average H3K27Ac signal for each condition in each of the seven clusters identified (A-G) were generated, and the two where a strongest impact of SMARCA5 depletion was observed are shown in Figure 12B. The activity of regions in cluster G (n=1967) appeared to be decreased in SMARCA5-depleted cells, while increased acetylation levels were found at regions in cluster E (n=867) especially after LPS stimulation. The effects of SMARCA5 knockdown at regions in other clusters were very mild. In fact, those clusters were identified based on the changes mostly due to LPS induction.

To investigate whether specific TFs might potentially regulate differentially acetylated regions, a motif enrichment analysis on each cluster of regions identified was carried out.
NFkB was the top motif identified in most clusters, showing the magnitude of the effects of LPS stimulation. Interestingly, regions in cluster G were significantly enriched in IRF binding sites, while one of the top motifs identified in cluster E was for the TF CEBPB (Figure 12C), consistent with the previous findings of over-represented motifs at sites with decreased or increased accessibility.

Figure 12. Impact of SMARCA5 depletion on enhancer activity as determined by H3K27ac ChIP-seq analysis. (A) Heatmap showing hierarchically clustered regions with changes in H3K27ac levels. (B,C) Line plots showing average profiles of H3K27ac signal (B) and motif enrichment analysis (C) from two clusters of regions induced by LPS and affected by SMARCA5 depletion. (D) Snapshot showing the overlap between accessibility and H3K27ac at a cluster of interferon genes.

These results suggested a strong correlation between changes in acetylation levels and chromatin accessibility. However, such possible correlation remains to be demonstrated with a formal analysis showing the overlap of both datasets. Several examples of regions where changes in acetylation levels appear to be in agreement with changes in chromatin accessibility have been found by visual inspection of the data. As a representative
example, a snapshot showing regions with reduced accessibility and acetylation at a cluster of interferon genes is shown in Figure 12D, where two specific regions with both decreased accessibility and acetylation are marked.

In summary, I have shown that SMARCA5 is required to maintain accessibility and acetylation levels at IRF binding regions, which would in part explain why the expression of interferon-stimulated genes is affected by the loss of SMARCA5. Moreover, I found that SMARCA5 plays a repressive role at many genomic regulatory regions, where the chromatin is opened and enhancers are activated upon knockdown of SMARCA5. A question that remained unanswered, and which became the focus of my next investigations, was whether all of these effects came from direct regulation by SMARCA5, or if there were other players involved.

4. SMARCA5 genome-wide occupancy in macrophages

To map genomic regions bound by SMARCA5, I performed ChIP-seq analysis. However, unlike transcription factors and histone marks, chromatin remodellers do not directly contact the DNA and they are notoriously difficult to ChIP. Therefore, new conditions for ChIP-seq had to be tested and optimised. Different methods of fixation and chromatin shearing were explored. Large amounts of cells were needed to test all conditions, so RAW264.7 cells were used for practical reasons. The optimal results were obtained with cells fixed with 2% formaldehyde and subsequently digested with Mnase, as shown in Figure 13.
Once the conditions had been optimised, I performed SMARCA5 ChIP-seq analysis in untreated and LPS-treated BMDMs. ChIP-seq libraries were prepared and sequenced as explained before. Preliminary inspection showed that SMARCA5 binds broad chromatin domains (>1kb). Using the tool SICER (Zang et al., 2009), which was specifically developed for calling broad peaks from ChIP-seq data, more than 19000 regions with an average size of 5 Kb were identified as bound by SMARCA5. The overlap between basal and LPS-stimulated binding was quite strong (>70%), so further analysis was mainly focused on binding at the basal state.

To determine whether the effects on chromatin accessibility previously outlined might come from direct regulation by SMARCA5, three groups of regions were selected based on how their accessibility is affected by the knockdown of SMARCA5 (down, not changing and up), and the intensity of SMARCA5 binding was plotted in 5 Kb regions centred on the summit of selected ATAC-seq peaks (Figure 14A). Surprisingly, regions with increased accessibility upon SMARCA5 depletion were rarely occupied by SMARCA5 in control cells, which might indicate that such effect is regulated through an indirect mechanism. On the other hand, regions that lose accessibility appeared to be constitutively bound by SMARCA5.

To elucidate the mechanisms involved, a motif enrichment analysis focused on the central 150 bp of SMARCA5-bound regions was performed using the PSCAN-ChIP tool.
Over-represented motifs are shown in Figure 14B. Interestingly, the IRF1 motif was identified as significantly enriched, even if it did not rank at the top of the list. In addition, under-represented motifs at regions bound by SMARCA5 were also identified in this analysis. Consistent with SMARCA5 being rarely bound to regions that gained accessibility in depleted cells (Figure 14A, right), which were previously shown to be enriched in binding sites for factors of the C/EBP and AP-1 families (Figure 11D), the same motifs were identified as significantly under-represented at regions bound by SMARCA5 (Figure 14B, bottom).

Following the identification of the IRF1 motif at regions bound by SMARCA5, I decided to explore whether these regions were also associated with specific functions related to the interferon response. To do so, SMARCA5-bound sites were assigned to their nearest genes and a gene ontology analysis was subsequently carried out using the GREAT tool (McLean et al., 2010). As shown in Figure 14C, several categories related to interferon signalling were identified as significantly associated with genes in the proximity of SMARCA5-bound regions.

Taken together, these results point to a direct role of SMARCA5 in the control of the chromatin state at IRF-bound regions, where SMARCA5 acts to maintain chromatin accessibility and acetylation, and thus regulates the expression of interferon-stimulated genes. Figure 14D shows a representative snapshot where binding of SMARCA5 can be seen at the promoters of two interferon genes and in correspondence of regions where there is also decreased accessibility and acetylation upon knockdown of SMARCA5.

In contrast, genomic regions where chromatin is maintained in a closed state in presence of SMARCA5, but which become accessible in depleted cells, do not seem to be directly controlled by SMARCA5. Instead, the effects observed might be mediated by factors of the C/EBP and AP-1 families.
Figure 14. SMARCA5 recruitment to chromatin in BMDMs as determined by ChIP-seq analysis. (A) Heatmap showing intensity of SMARCA5 ChIP-seq signal at ATAC-seq peaks +/- 2.5 kb. ATAC-seq regions with decreased, not changing or increased accessibility upon SMARCA5 knockdown were obtained from the dataset presented in Figure 10. (B) Motif enrichment analysis on genomic regions occupied by SMARCA5. (C) Bar plot showing top-ranking molecular signature pathways associated with genes previously assigned to genomic regions bound by SMARCA5 as identified by GREAT analysis. (D) Snapshot showing SMARCA5 binding in correspondence of regions that lose accessibility and acetylation upon SMARCA5 depletion.

5. CEBPB: a possible mediator of the action of SMARCA5

We hypothesised that any factor involved in the opening of chromatin upon SMARCA5 knockdown might also be differentially expressed in depleted cells. Cebpa has been reported to be upregulated upon SMARCA5 knockdown in mouse erythroleukemia cells...
In BMDMs, however, Cebpa expression levels were low and no differences were observed between depleted and control cells (not shown). Interestingly, the TF CEBPB was found to be upregulated upon SMARCA5 depletion. Cebpb mRNA levels were only slightly upregulated in unstimulated SMARCA5-depleted cells as determined by RT-qPCR analysis in two independent experiments (Figure 15A). However, a Western Blot analysis revealed a strong induction of CEBPB expression at the protein level in both untreated and stimulated cells (Figure 15B).

To determine if the binding profile of CEBPB was affected by SMARCA5 depletion, ChIP-seq analysis in control and SMARCA5-depleted cells was performed. Before preparing ChIP-seq libraries, the material from two ChIP experiments was analysed by qPCR. As shown in Figure 15C, CEBPB enrichment was detected at two constitutive target regions. The enrichment was higher upon SMARCA5 knockdown, which might simply be due to higher levels of protein being able to immunoprecipitate higher amounts of DNA.

A genome-wide analysis of the CEBPB ChIP-seq data identified regions with constitutive CEBPB binding as well as regions differentially occupied by CEBPB in control and SMARCA5-depleted cells. As expected, there was a higher enrichment of CEBPB signal at a number of genomic regions upon SMARCA5 depletion (log₂FC>0.5) (Figure 15D). The fact that also regions with decreased (log₂FC<-0.5) or equivalent (log₂FC<10.005) CEBPB occupancy were identified might suggest that the effects observed upon loss of SMARCA5 were indeed due to differential CEBPB binding and not only to differences in CEBPB expression.

To confirm the specificity of the ChIP-seq experiment, a motif enrichment analysis was performed. CEBPB was the top significantly enriched motif (Figure 15E). A motif enrichment analysis for regions that either gain or lose CEBPB binding might help explain the differences observed, but remains to be carried out.
Figure 15. Impact of SMARCA5 depletion on CEBPB expression and recruitment to chromatin. (A) RT-qPCR analysis showing mRNA expression of Cebpb in control and SMARCA5-depleted cells. Transcript levels were normalised against Tbp. Data are expressed as the mean±SD (n=2). (B) Western Blot analysis showing expression of CEBPB upon SMARCA5 knockdown. Vinculin levels are shown as loading control. (C) ChIP-qPCR analysis showing CEBPB enrichment at two constitutive target regions in control and SMARCA5-depleted cells. Data from two independent experiments are shown and plotted as mean±SD. (D) Box plot showing CEBPB average signal intensity in control and SMARCA5-depleted cells in regions with increased enrichment (log2FC>0.5), not changing (log2FC<0.005I) or with decreased enrichment (log2FC<-0.5). Data are expressed as normalised average read counts. (E) Top significantly enriched motif in genomic regions identified by CEBPB ChIP-seq analysis. (F) Snapshot showing CEBPB binding in control and SMARCA5-depleted cells in correspondence of representative ATAC-seq regions that either do not change or gain accessibility upon SMARCA5 knockdown.

It will be interesting to see whether regions with increased CEBPB occupancy correlate with sites that gain accessibility and acetylation upon SMARCA5 knockdown. Preliminary
inspection of the data suggested that some of the regions which gain accessibility upon SMARCA5 depletion might be de novo bound by CEBPB in depleted cells. An example is shown in Figure 15F, where a region with constitutive CEBPB binding can also be seen.

6. Effects mediated by ISWI accessory subunits on the transcriptomic profile of macrophages

The study of the common catalytic subunit SMARCA5 provided relevant information on the general role of ISWI chromatin remodelers. However, we wanted to further understand the specific role of each ISWI complex, which should be determined by the regulatory activity and binding specificity provided by ISWI accessory subunits. As such, to assess the specific role of each ISWI complex, shRNA-mediated knockdown was used once again to deplete each of the accessory subunits (BAZ1A, BAZ1B and BPTF) in BMDMs and subsequently transcriptional activity was analysed by RNA-seq. One shRNA targeting each subunit was selected out of a set of five. RNA-seq libraries were prepared using SMART-seq technology from three replicates with the highest knockdown efficiency, which was confirmed by RT-qPCR analysis (Figure 16).

**Figure 16. Evaluation of knockdown efficiency of the shRNAs targeting ISWI accessory subunits.** RT-qPCR analysis showing effective depletion of Baz1a, Baz1b and Bptf at the mRNA level. Transcript levels were normalised against Tbp. Data are expressed as the mean±SD (n=3).

Differential expression analysis was performed separately for each of the subunits by comparing the three replicates of control and depleted cells using the tool EdgeR for statistical analysis (Robinson et al., 2009). The total numbers of upregulated and
downregulated genes identified for each subunit at each time-point are reported in Figure 17A (log2FC>1 and FDR<0.01). BPTF had the highest impact, since the numbers of DEGs for the other subunits were much lower in comparison. Therefore, subsequent analyses were focused on further understanding the role of BPTF alone.

To investigate whether specific biological functions were associated with genes either upregulated or downregulated upon BPTF knockdown, a gene ontology analysis was carried out (Figure 17B). Downregulated genes were most significantly associated with the regulation of the cell cycle. Surprisingly, several categories related to the interferon response were identified as significantly enriched among upregulated genes, as clearly exemplified by the presence on the list of the specific signature ‘response to interferon-beta’. These findings are apparently in complete contrast with the results obtained upon knockdown of SMARCA5.

To identify genomic regions bound by BPTF, ChIP-seq analysis was carried out in untreated and LPS-treated BMDMs. ChIP-seq was performed using the same conditions as previously optimised for SMARCA5 (Figure 13). ChIP-seq libraries were prepared and sequenced as explained above.

Preliminary inspection of the ChIP-seq data showed that BPTF, similarly to SMARCA5, binds broad chromatin domains. In fact, BPTF binding seemed to largely overlap with SMARCA5. Using the tool SICER, BPTF-bound regions were identified. Interestingly, there was an overlap of around 70% between regions bound by BPTF and regions bound by SMARCA5. These results indicate that BPTF, like SMARCA5, might also occupy IRF-bound regions to regulate the expression of interferon-stimulated genes. As an example, Figure 17C shows a snapshot of an interferon gene downregulated upon SMARCA5 depletion and upregulated in BPTF-depleted cells, where the promoter seems to be occupied both by SMARCA5 and BPTF, even if binding of the latter is only observed after LPS stimulation.
Figure 17. RNA-seq analysis upon depletion of ISWI non-catalytic subunits uncovers a major role of the BPTF subunit in the regulation of the inflammatory gene expression program. (A) Table showing the number of DEGs identified upon knockdown of each ISWI accessory subunit in unstimulated macrophages and following LPS stimulation (FDR<0.01). (B) Bar plot showing top-ranking gene ontology categories associated with genes upregulated (top) or downregulated (bottom) upon BPTF knockdown. (C) Snapshot showing expression of an interferon-stimulated gene in SMARCA- and BPTF-depleted cells, as well as the overlap with SMARCA5 and BPTF binding in the same region.

7. Effects of SMARCA5 deletion in an alternative background: retrovirally transduced BMDMs

The above data all pointed to a role of ISWI chromatin remodellers in the regulation of interferon-stimulated genes. However, the interferon response is known to be induced to some extent by shRNAs delivered via lentiviral transduction (Kenworthy et al., 2009).
Therefore, I set out to explore the role of SMARCA5 using an alternative strategy. BMDMs from wildtype and Smarca5\(^{lox/lox}\) mice were retrovirally transduced with a Cre recombinase-expressing plasmid. As compared to lentiviral infection, macrophages are barely affected by retroviral infection. Infected cells were kept unstimulated or subsequently subjected to either LPS or IL4 treatment (2 h), since we wanted to assess the selectivity of ISWI complexes in the regulation of transcriptional responses to distinct stimuli.

Efficient recombination at the Smarca5 locus was confirmed by genomic PCR analysis (Figure 18A). Knockout efficiency was validated by RT-qPCR and Western Blot analysis: a complete deletion was observed both at the mRNA (Figure 18B) and at the protein level (Figure 18C).

![Figure 18. Evaluation of SMARCA5 deletion in Smarca5\(^{lox/lox}\) BMDMs transduced with a Cre-expressing retroviral plasmid. (A) PCR analysis on genomic DNA showing Cre-mediated recombination at the Smarca5 locus in floxed cells compared to WT BMDMs. (B) RT-qPCR analysis showing effective deletion of Smarca5 at the mRNA level. Transcript levels were normalised against Tbp. (C) Western Blot analysis showing SMARCA5 protein levels in WT and Smarca5-KO BMDMs. Vinculin levels are shown as loading control. Results from one representative experiment of three biological replicates are shown.]

To determine the transcriptomic profile of WT and KO cells, RNA-seq libraries were prepared from three biological replicates using SMART-seq technology as described above. Using the tool EdgeR (Robinson et al., 2009), differential expression analysis was performed by independently comparing the three replicates of WT and KO cells subjected to each of the treatments (UT, LPS, IL4). Differentially expressed genes (log\(_2\)FC>1 and FDR<0.01) either by treatment or by knockout of SMARCA5 were hierarchically clustered.
according to their transcriptional profile, and were subsequently plotted in the heatmap shown in Figure 19A.

If only genes affected by SMARCA5 deletion are taken into account, a total of 360 genes were identified as differentially expressed in at least one of the treatments. The specific numbers of upregulated and downregulated genes identified, as well as the overlap between different treatments, are shown in Figure 19B. In contrast to my previous findings in SMARCA5-depleted macrophages, the number of upregulated genes (Figure 19B, left) was higher than that of the downregulated ones (Figure 19B, right). Moreover, the magnitude of the effects observed in IL4-treated cells, in terms of number of DEGs identified, was lower compared to LPS.

To identify TFs that might mediate the transcriptional effects observed upon SMARCA5 deletion, a motif enrichment analysis focused on the promoters of DEGs in each of the thirteen clusters identified (A-M) was carried out. Only motifs enriched in clusters K, G and H were statistically significant. As shown in Figure 19C, NFkB was the top motif identified both in cluster K and G, most likely due to the magnitude of the effects of LPS stimulation. Interestingly, promoters of DEGs in clusters G and H, which are upregulated upon SMARCA5 deletion particularly after LPS treatment, were significantly enriched in binding sites for factors of the IRF family.

To validate the upregulation of interferon genes upon SMARCA5 knockout, I analysed the mRNA expression of Irf7 by RT-qPCR in three independent experiments (Figure 19D). These findings were completely opposite to the results previously obtained upon lentiviral-mediated knockdown of SMARCA5.
Figure 19. Global and specific effects of SMARCA5 deletion on gene expression of retrovirally transduced BMDMs. (A) Heatmap showing hierarchically clustered genes with changes in expression identified by RNA-seq analysis. (B) Venn diagrams showing the number of upregulated (left) or downregulated (right) genes identified by RNA-seq analysis as well as the overlap between the unstimulated, LPS-treated or IL-4-treated conditions (FDR<0.01) (C) Motif enrichment analysis on the promoters of genes in three of the clusters identified in (A). (D) RT-qPCR analysis showing Ifn7 mRNA expression in WT and KO BMDMs. Transcript levels were normalised to Tbp. Data are expressed as the mean±SD (n=3). (E) Western Blot analysis showing total and phosphorylated STAT1 (Tyr701) in WT and KO cells. Vinculin levels are shown as loading control.

Consistent with the observed upregulation of interferon-stimulated genes, a Western Blot analysis revealed a strong induction of the TF STAT1 at the protein level in both unstimulated and LPS-treated cells from two independent knockout experiments, as well as increased activity as determined by higher levels of tyrosine-phosphorylated STAT1 (Figure 19E). However, it remains to be determined whether this might be a result of the
upregulation of interferon response genes, or whether STAT1 is actually involved in mediating such upregulation.

To further characterise the role of SMARCA5 in this context, ATAC-seq libraries were prepared as previously explained. A genome-wide analysis of these data revealed global effects of SMARCA5 deletion on chromatin accessibility. Genomic regions with statistically significant changes in accessibility (log$_2$FC>1 and FDR<0.01) either after treatment or upon SMARCA5 deletion were identified, hierarchically clustered and subsequently plotted in a heatmap as shown in Figure 20A. Metaprofiles showing average ATAC-seq signal for each condition in each of the clusters identified were generated, and four of them are shown in Figure 20B. A strong impact of SMARCA5 deletion was observed in clusters B (n=720), D (n=428), A (n=224) and Q (n=195), where chromatin accessibility was significantly increased compared to WT cells.

To investigate whether specific TFs might potentially regulate differentially accessible sites, a motif enrichment analysis was performed focusing on the regions in each of the clusters identified. As shown in Figure 20C, regions in clusters B and D, which gained accessibility upon SMARCA5 deletion particularly after LPS stimulation, were most significantly enriched in binding sites for factors of the IRF family, consistent with the upregulation of interferon-stimulated genes previously observed in this setting but in contrast with the decreased accessibility observed at those sites upon SMARCA5 knockdown (Figure 11D). Interestingly, regions in clusters A and Q were significantly enriched in CEBPB binding motifs, which is in agreement with the previous findings from the ATAC-seq analysis in SMARCA5-depleted cells where the same motif was identified in regions that gained accessibility (Figure 11D).
Figure 20. Genome-wide effects of SMARCA5 deletion on chromatin accessibility as determined by ATAC-seq analysis of BMDMs subjected to retroviral transduction. (A) Heatmap showing hierarchically clustered regions with changes in chromatin accessibility identified by ATAC-seq analysis. (B) Line plots showing average profiles of ATAC-seq signal at four clusters of regions identified as affected by SMARCA5 deletion. (C) Motif enrichment analysis on differentially accessible sites from clusters shown in (B). (D) Snapshot showing expression of an interferon-stimulated gene in WT and KO BMDMs, as well as accessibility in the same cells determined by ATAC-seq analysis, and SMARCA5 binding in correspondence of the promoter of the same gene.

The fact that IRF motifs were identified as significantly enriched at differentially accessible sites suggested a correlation between changes in gene expression and changes in chromatin accessibility upon SMARCA5 knockout. A formal analysis to determine the possible overlap between both datasets remains to be carried out. However, an increase of chromatin accessibility upon knockout of SMARCA5 was observed by visual inspection of the data at several regions in the proximity of upregulated interferon genes.
An example of an interferon-stimulated gene upregulated by SMARCA5 deletion, and the promoter of which appeared to gain accessibility, is shown in Figure 20D. As previously shown, these regions overlap with SMARCA5 binding, which can also be seen in Figure 20D.
DISCUSSION

1. Role of ISWI complexes in the regulation of the interferon response

In this study, I have shown that ISWI complexes are involved in the regulation of the expression of interferon-stimulated genes (ISGs) in murine macrophages. In cells transduced with lentiviral shRNAs, depletion of the catalytic subunit SMARCA5 results in reduced expression of interferon genes. Importantly, I found that SMARCA5 is bound to IRF binding regions, where evidence suggests that it controls accessibility of the chromatin and acetylation levels. Consistent with my findings, a recent study by Cao L. et al. proposed that SMARCA5 mediates an antiviral response activated by the nuclear matrix protein HnRNPU, which they identified as a viral RNA sensor in the nucleus (Cao et al., 2019). Interestingly, they reported impaired induction of Ifnb1 in SMARCA5-KO BMDMs upon infection with HSV-1, as well as reduced activity at Ifnb1 enhancers as determined by H3K27Ac ChIP-qPCR.

It could be hypothesised that SMARCA5 is recruited to chromatin by one of the IRF factors to shape the epigenomic landscape at binding sites for transcription factors of the IRF family. One possible candidate is IRF7, since Irf7 is the top-ranked downregulated gene identified upon loss of SMARCA5. In addition, I provide evidence that genes downregulated upon SMARCA5 knockdown largely overlap with IRF7-dependent genes, which would support the idea that both factors cooperate to modulate the expression of ISGs. However, whether SMARCA5 is able to physically interact with IRF7 and whether the binding profile of SMARCA5 would be affected by the loss of IRF7 are questions that remain to be answered.

Here I have extensively studied the role of SMARCA5 as the common ISWI catalytic subunit in macrophages, since I demonstrated that SMARCA1 is not expressed in macrophages and I confirmed that BPTF is also able to interact with SMARCA5, as results
from a previous study performed in HeLa cells already indicated (Oppikofer et al., 2017). However, not all SMARCA5-containing complexes play the same roles. As a matter of fact, it is widely known that different ISWI complexes sometimes play opposing roles (Giles et al., 2019; Zentner et al., 2013). Therefore, I also attempted to characterise the specific roles of different ISWI complexes by analysing the impact of shRNA-mediated depletion of distinct ISWI accessory subunits on the transcriptional activity of macrophages.

Very limited effects were observed upon depletion of the BAZ1A and BAZ1B subunits, which indicates that ISWI complexes containing these accessory subunits may play redundant roles in the regulation of the inflammatory gene expression program. In favour of this, the expression profiles of BAZ1A and BAZ1B appeared to be complementary following LPS stimulation, as illustrated by the fact that the expression of Baz1a is strongly induced at 2h while Baz1b is repressed. Additionally, preliminary results indicated that the expression of both BAZ1A and BAZ1B is decreased upon SMARCA5 depletion, while the depletion of each accessory subunit does not affect the expression of other subunits. Further research is required to confirm the hypothesis that BAZ1A- and BAZ1B-containing complexes play redundant roles, as could be determined by assessing the impact of simultaneously depleting both subunits.

Depletion of the BPTF subunit, on the other hand, greatly affects the expression of ISGs. However, in contrast to the effects observed upon SMARCA5 depletion, loss of BPTF results in increased expression of interferon genes. These results are actually in agreement with findings from a study carried out in Drosophila, where deletion of NURF301 was shown to enhance the expression of a large set of JAK/STAT target genes (Kwon et al., 2008). Using ChIP-seq, I have further shown that BPTF binds the chromatin in a manner that largely overlaps with SMARCA5 genomic occupancy, including at binding sites for IRFs. Nevertheless, it remains to be determined whether the chromatin state at IRF binding regions is also impacted by the loss of BPTF, which would help explain how BPTF is able to antagonise the effects of SMARCA5 on ISGs.
A major concern for the interpretation of the results from this study was the fact that the interferon response is known to be induced by lentiviral vector-mediated transduction of shRNAs (Kenworthy et al., 2009). Consistently, I found increased *Irf7* mRNA expression in cells infected with a control luciferase-targeting shRNA compared to the levels of expression in uninfected cells. Considering that IRF7 is known to induce IFN-α/β (Honda et al., 2005), we can assume that effects of SMARCA5 depletion in lentivirally-transduced BMDMs were measured in a setting of basal and persistent activation of the interferon response.

Using an alternative strategy, which consisted in transducing BMDMs from "Smarca5lox/lox" mice with a Cre recombinase-expressing retroviral plasmid, deletion of SMARCA5 resulted in increased expression of ISGs and increased accessibility at IRF binding regions, in complete contrast to the phenotype observed in cells transduced with SMARCA5-targeting lentiviral shRNA. In this case, cells did not respond to the infection and the interferon response was only induced upon LPS stimulation. I hypothesise that these effects might be in part mediated by the TF STAT1, which is strongly upregulated in knockout cells and shows greater activity as determined by increased levels of its phosphorylated form.

Further studies are needed to determine the exact mechanisms underlying the regulation of the interferon response by ISWI complexes. However, we could speculate that SMARCA5 exerts its action in a context-dependent manner. In other words, the role of SMARCA5 may be different in 'naïve' cells compared to cells which have been previously activated, for example by a lentiviral infection, when re-challenged with inflammatory stimuli.

2. Repressive role of SMARCA5 at CEBPB-regulated regions

As outlined above, SMARCA5 appeared to play an activating role at regions regulating the interferon response in macrophages transduced with lentiviral shRNAs. However, in this same context, I also uncovered a major repressive role of SMARCA5. I have shown
that depletion of SMARCA5 results in increased accessibility at a large number of genomic sites, including many regions bound by CEBPB. In fact, CEBPB itself is strongly upregulated upon SMARCA5 depletion, which is in line with the results from a previous study in which the expression of Cebpα, another TF from the same family, was found to be upregulated upon SMARCA5 knockdown in mouse erythroleukemia cells (Dluhosova et al., 2014). The repression of CEBPB-regulated regions is consistent with a study which showed that CEBPB physically interacts with SMARCA5 and that, in its presence, the transactivation activity of CEBPB is suppressed (Steinberg et al., 2012).

The fact that the action of SMARCA5 has opposite effects on different genomic regions is not surprising. Several studies have previously reported opposing roles not only between different ISWI complexes, but also for one same complex. For example, the NURF complex in Drosophila has been proposed to act as a negative regulator on innate immune response genes, as mentioned above, while targets of nuclear receptors are apparently activated by NURF (Badenhorst et al., 2005; Kwon et al., 2008). It remains to be determined whether the repressive role of SMARCA5 observed in this work could be attributed to a specific ISWI complex.

In addition to connecting the effects observed to the function of a specific ISWI complex, further research is also needed to determine the mechanisms involved in the control of genomic regions that gain accessibility and CEBPB binding upon SMARCA5 depletion. Based on the finding that SMARCA5 is rarely bound to those regions, one may conclude that it is not the direct action of SMARCA5 which maintains them in a repressed state, as opposed to what happens at IRF binding regions. In fact, both effects might not be completely unrelated. A study by Kurotaki D. and colleagues showed that, in myeloid progenitors, the TF IRF8 physically interacts with CEBPA and prevents it from binding to chromatin and activating transcription of target genes (Kurotaki et al., 2014). Moreover, CEBPB and IRF1 have been shown to be part of the same complex (Rani et al., 2010). It is thus tempting to hypothesise that other factors from the IRF and C/EBP families might
also be able to interact. If IRF7 interacted with CEBPB preventing its binding to chromatin, the opening of the chromatin at CEBPB binding sites upon SMARCA5 knockdown could then be explained by the downregulation of IRF7 and subsequent release of CEBPB. To validate this hypothesis, we should first study the possible interactions both between these factors and possibly with SMARCA5 and then assess CEBPB occupancy in IRF7-depleted cells, which could be compared to the binding profile obtained in SMARCA5-depleted cells.

Lastly, preliminary results obtained in this study support the idea that the role of SMARCA5 at CEBPB-regulated regions might be conserved in macrophages subjected to retroviral transduction, as indicated by the significant over-representation of CEBPB binding motifs found at a fraction of the genomic sites that gain accessibility upon SMARCA5 knockout. However, further analyses need to be done in this setting to definitely confirm the conservation of the repressive role of SMARCA5 described above.

3. Future perspectives

In summary, this study has revealed different components of the transcriptional program of macrophages which seem to be regulated by ISWI complexes, or more specifically by SMARCA5 as the specific role of other ISWI subunits remains to be further characterised. However, further research is also needed to better understand the basis of the regulation by SMARCA5. Some pending questions have already been outlined above. One of the main aspects to explore will be the SMARCA5 interactome. It will be essential to identify interactions between SMARCA5 and transcription factors which can possibly mediate the action of ISWI complexes, for example IRFs, STATs or C/EBPs.

In addition, it will be interesting to investigate the selectivity of the action of SMARCA5 by assessing the impact of the loss of SMARCA5 on the response to stimuli other than LPS. In this regard, preliminary results reported in this study indicate that the expression of a substantial number of genes is affected by SMARCA5 deletion in IL4-treated cells, even if the number was lower compared to genes identified upon LPS stimulation.
Finally, the phenotypes observed should be further explored and characterised *in vivo*. To do so, I have bred the Smarca5\textsuperscript{lox/lox} mice described in this study with a Csf1r-Cre mouse (Qian et al., 2011) to obtain a macrophage-specific Smarca5-KO. Using these mice, we will be able to delete Smarca5 in all Csf1r-expressing cells (i.e. tissue resident macrophages) and study the systemic effects of the deletion in response to inflammatory stimuli *in vivo* (e.g. LPS). It will be interesting to investigate if Smarca5-deleted macrophages from different tissues respond differently to inflammation.
CONCLUSIONS

1. The expression of ISWI accessory subunits in primary macrophages is dynamically modulated in response to LPS stimulation.

2. The BPTF subunit is in a complex with SMARCA5, which is the only ISWI catalytic subunit expressed in macrophages.

3. ISWI complexes are recruited to chromatin to control the interferon gene expression program in macrophages, possibly through cooperation with factors of the IRF and STAT families.

In cells transduced with lentiviral shRNAs:

- Depletion of SMARCA5 results in the downregulation of many genes which largely overlap with IRF7-dependent genes, while BPTF depletion causes the upregulation of a similar set of genes.

- SMARCA5 is required to maintain chromatin accessibility and acetylation levels at IRF binding regions.

In Smarca5lox/lox cells transduced with a retroviral Cre recombinase-expressing plasmid:

- Interferon-stimulated genes are upregulated upon SMARCA5 knockout and the genomic regions in their vicinity appear to gain accessibility.

4. CEBPB may mediate the opening and activation of a large set of genomic regions which gain accessibility upon loss of SMARCA5.


https://doi.org/10.1093/nar/gkz908


I’d like to start by thanking Gioacchino. Thank you for giving me the opportunity to work in your lab, for trusting me with this project, and for providing an outstanding scientific training in these early steps of my career. Doing a PhD with the guidance of such a brilliant scientist has been truly inspiring.

Many thanks also to my external supervisor Irina Udalova for carefully reading all my yearly reports and always sending useful comments. I wish we could have meet in person, but the discussions we had on Skype were also very helpful.

Next, I need to thank the whole Natoli lab. I feel very lucky to have been able to work with such an excellent group of people from whom I have learnt so much. I wish I could thank you one by one, but I don’t want to extend myself too much. You should just know that, one way or another, you have all contributed to the success of this thesis, whether it was with valuable technical expertise, interesting discussions, or sometimes even with much needed moral support.

I do want to say special thanks to Marta S. and Sara. Thank you for everything you taught me, for every time you gave me good advice and, most importantly, for always being available to discuss paranormal activity in my experiments. You always helped me stay calm in the middle of the desperation of experiments not working, and I could not thank you both enough for that.

Finally, I’d like to thank Marta R. for being the best bench/desk/PhD mate I could have asked for. Nobody else has ever gone into a -80 for me (and probably never will)! Thank you for always being there for me, through all the ups and downs of this crazy adventure the PhD has been.

A tots els de casa (ja sabeu qui sou si arribeu a llegir això): gràcies per haver-me fet sempre costat desde la distància, per cada missatge o trucada de suport i per haver-me animat en els moments més difícils. Aquesta tesi també és una mica vostra.