PhD degree in Molecular Medicine (curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples “Federico II”

Settore disciplinare: Bio/11

3D GENOMIC ORGANIZATION OF MOUSE MACROPHAGES

Silvia Masella

IEO, Milan

Matricola n. R09400

Supervisor: Dr. Gioacchino Natoli

IEO, Milan

Anno accademico 2013-2014

A mamma e papà,

da sempre il mio sostegno più grande.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

3C: Chromosome Conformation Capture

4C: Circular Chromosome Conformation Capture

5C: Carbon-Copy Chromosome Conformation Capture

Ac: Acetylation

ACH: Active Chromatin Hub

AP-1: Activator Protein 1

BMDMs: Bone Marrow Derived Macrophages

BSA: Bovine Serum Albumin

cDNA: complementary DNA

ChIP: Chromatin Immunoprecipitation

ChIP-seq: ChIP-sequencing

CTCF: CCCTC-binding factor (zinc finger protein)

Ctrl: Control

ddH_2O: double-distilled water

dNTP: deoxyNucleotide Tri-Phosphate

eRNAs: enhancer RNAs

ES cell: Embryonic Stem Cell

FBS: Fetal Bovine Serum

H2: Histone2

H3: Histone3

H4: Histone4

H3K4: Lysine 4 on Histone3

H3K27: Lysine27 on Histone3

H4K5/8/12: Lysine5/8/12 on Histone4

Hi-C: Genome-Wide Chromosome Conformation Capture

HT: High-Throughput

ICR: Imprinting Control Region

IKK: Inhibitor of KappaB Kinase

Il: Interleukin

iPS cell: Induced Pluripotent Stem cell

IRF: Interferon-Regulatory Factor

KD: Knock-Down

LCR : Locus Control Region

LMA: Ligation-Mediated Amplification

LPS : Lipopolysaccaride

Luc: (Firefly) Liciferase

MEF: Mouse Embryonic Fobroblast

MHC-II: Major Histocompatibility Complex Class II

-me1: trimethylated lysine residue

-me3: trimethylated lysine residue

MHC-II: Major Histocompatibility Complex Class II

mRNA: messengerRNA

NF-kB: Nuclear Factor-kappa B

NGS: Next-Generation Sequencing

NPC: Neural Progenitor Cell

O.N.: Over Night

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

PolII: Polymerase II

PRGs : Primary Response Genes

qPCR: quantitative PCR

RPM: Revolutions Per Minute

shRNA: short hairpin RNA

SRGs: Secondary Response Genes

STAT: Signal Transducer and Activator of Transcription

TADs: Topological Domains or Topological Associating Domains

TFs: Transcription Factors

TLRs: Toll Like Receptors

TNFα: Tumor Necrosis Factor alpha

TSS: Transcription Start Site

WB: Western Blot

WCE: Whole Cell Extract

FIGURES INDEX

TABLES INDEX

ABSTRACT

The spatial organization of the genome and its biological function are intimately linked. It is becoming evident that transcription regulation often involves multiple long-range regulatory elements and it's influenced both by the genomic environment and by the shape of the genome. Recent studies based on Chromosome Conformation Capture-derived techniques, showed that eukaryotic cells organize their chromosomes into topological domains that are largely invariant among cell types and where the majority of looping interactions between regulatory elements take place (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). The principles defining the relationships between these elements and distal target genes remain poorly understood. Previous studies lack either the spatial resolution or the temporal coverage to observe possible dynamic changes in chromatin contacts between promoters and their corresponding distal regulatory elements during gene activation.

Here we exploited high-resolution 4C and 5C techniques to elucidate principles of 3D organization of the macrophage genome, in both basal conditions and after macrophage activation, dissecting the specific role of the macrophage master regulator PU.1 in the formation of the general chromosome topology.

Our findings indicate that the global organization of chromatin contacts is to a large extent unaffected by macrophage activation, which only partially impacts the looping between specific regulatory elements. Our analysis also reveals an high cell-type specificity of macrophage promoter-enhancer interactions, which is not dependent on the presence of PU.1.

INTRODUCTION

1. INFLAMMATION

Inflammation is a fundamental response to disruption of cellular and tissue homeostasis, such as infection or tissue injury, with many important physiological roles, including host defense and tissue remodeling or repair (Hotamisligil, 2006).

Inflammation can be considered a self-limiting process. Excessive and/or sustained inflammation is a common cause of diseases, including acute (sepsis) and chronic inflammatory disorders (autoimmune diseases, atherosclerosis, metabolic syndrome and several types of cancers) (Medzhitov, 2008). During the inflammatory response multiple mechanisms operate at different levels, including the control of gene expression in individual cells, the regulation of signaling pathways and the alteration of immune cell composition in tissues (Medzhitov and Horng, 2009).

Different cell types are recruited to the inflammatory sites and activated through cell-specific mechanism. For example, in the early phase of inflammation, activated tissue-resident macrophages release chemokines (such as Cxcl8, Cxcl5 and Cxcl1) which recruit neutrophils to the damaged tissue. After the extravasation and tissue infiltration, neutrophils release granule proteins able to recruit and activate inflammatory monocytes which in turn differentiate into macrophages (Ley et al., 2007).

Macrophages are essential components of the innate immunity and important mediators of the inflammatory response, since they are the major producers of inflammatory cytokines, mediating the host defense processes against invading pathogens and also driving host defense, tissue remodeling/repair and homeostasis (Qiao et al., 2013).

Macrophages residing in different tissues are extremely heterogeneous from the point of view of gene expression and functions properties. They have a differential expression of CD14 and CD16 receptors and they vary in the chemokine-receptor expression profile. Macrophages also show different sizes and nuclear morphologies and they differ in the degree of granularity. Some examples are the Langerhans cells of the epidermis, Küpffer

cells in the liver and alveolar macrophages in the lung (Gordon and Taylor, 2005). Within the same tissue, macrophages can undergo phenotypic and functional changes upon exposure to distinct micro-environmental stimuli (Lawrence and Natoli, 2011).

Based on their response to external stimuli, macrophages can be classified in different subtypes. In a simplified, yet broadly used classification, the Classical (M1) macrophages produce high amounts of proinflammatory cytokines and are elicited by interferon γ. The wound-healing (Alternatively activated or M2) macrophages are induced by Th2 cytokines (IL4 and IL13). However, the spectrum of macrophage activation states is in fact much broader (Murray et al., 2014).

Macrophages express a set of Pattern Recognition Receptors, namely receptors for invariant structures (molecular patterns) of microbes (such as endotoxin), which include both intracellular and transmembrane Toll Like Receptors (TLRs). These receptors mediate host immune defense by detecting the presence of microbial molecules and activating downstream intracellular signaling events (Takeuchi and Akira, 2010).

One example is the binding of the Gram- bacteria endotoxin (Lipopolysaccaride, LPS) to the TLR4 receptor which activates a signaling cascade that causes the activation of a set of transcription factors (TFs), such as members of the Nuclear Factor-kappa B (NF-kB) Activator Protein 1 (AP-1) and Interferon-Regulatory Factor (IRF) families. Upon activation, TLR4 recruits MyD88 and TIRAP and a complex involving TRAF6 and the IRAK kinases is subsequently formed. TRAF6 is an E3 ligase that catalyzes the formation of a K63-linked polyubiquitin chain on TRAF6 itself, eventually leading to the activation of a complex composed of TAK1, TAB1, and TAB2/3. TAK1 activates the IkB kinase which in turn activate the IKK complex. Phosphorylated IkB is degraded and releases NF-kB, which can translocate into the nucleus and induce the expression of cytokine genes. TAK1 also activates a MAP kinase cascade resulting in the activation of the JNKs (cJun N-terminal kinases) and the phosphorylation of cJun/AP-1, which also contributes to the induction of cytokine genes (Figure 1) (Takeuchi and Akira, 2010).

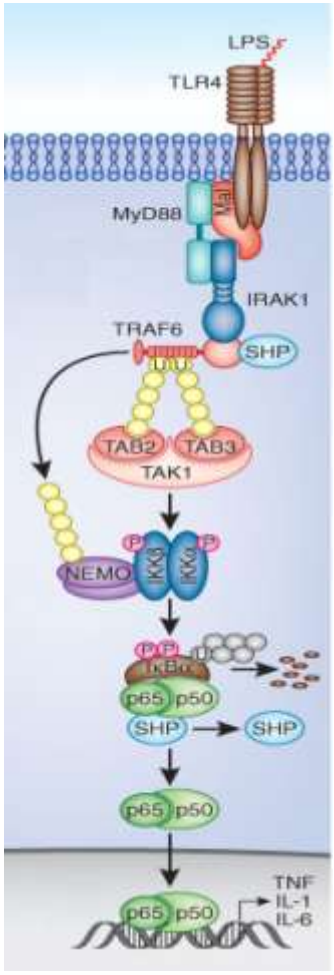

Figure 1 - Activation of the NF-Kb pathway by TLR4
Ligand binding by TLRs leads to the recruitment of receptor-specific adapters and induces activation of NF-kB. Adapted from (Beyaert, 2011).

2. TRANSCRIPTIONAL REGULATION OF THE INFLAMMATORY RESPONSE

Macrophage stimulation by inflammatory agents (notably LPS) leads to the activation of different TFs, whose recruitment to target gene promoters appears to be influenced by pre-existing chromatin features, such as positioned nucleosomes (Smale, 2010; Takeuchi and Akira, 2010). TFs activated by inflammatory stimuli and concurring to the activation of the inflammatory gene expression program have distinct binding specificities and include AP-1, IRFs, NF-kBs and STAT family members. The preferential or relative activation of specific subsets of these TFs depends on the nature of the activated receptor(s) (Smale, 2010; Takeuchi and Akira, 2010).

Several experimental data demonstrated that the transcriptional regulation of the inflammatory response differs at primary response genes (PRGs), that are defined as genes not requiring new protein synthesis for activation (Herschman, 1991) and are usually rapidly activated after stimulus, and secondary response genes (SRGs, genes that require new protein synthesis) that display delayed activation kinetics (Fowler et al., 2011).

The promoters of most PRGs are G+C rich and often associated with a CpG island (Ramirez-Carrozzi et al., 2009). The very high G+C content interferes with nucleosome assembly (Barozzi, 2014), thus making them prone to rapid activation in the absence of chromatin remodelers. CpG island promoters show high level of H3K4me3, histone acetylation and RNA Pol II already in unstimulated macrophages, suggesting that their permissive chromatin configuration enables a low level of constitutive transcriptional activity mainly generating unspliced transcripts. An increase in H3K4me3 and H4K5/8/12Ac occurs upon stimulation and is associated with an increase in Ser2-phosphorylated RNA Pol II, productive elongation and generation of normally spliced transcripts (Hargreaves et al., 2009; Medzhitov and Horng, 2009).

The promoters of SRGs, such as the Nitric oxide syntase 2 (*Nos2*) and the Interleukin 6 (*Il6*) genes, as well as of some PRGs with delayed activation kinetics such as *Ccl5*, show low basal levels of H3K4me3 and H3/H4Ac (De Santa et al., 2009; Escoubet-Lozach et al., 2011). SRG promoters are also characterized by a comparatively lower G+C content and a sequence context that favors nucleosome occupancy, explaining the requirement for a nucleosome remodeling step triggered by the SWI/SNF chromatin-remodeling for their activation (Hargreaves et al., 2009; Ramirez-Carrozzi et al., 2009).

A recent genome-scale analysis of nascent transcripts in LPS-induced macrophages revealed that CpG islands are in fact also present at some SRG promoters which are not constitutively transcribed probably because of the lack/inactivity in the basal state of TFs required for their transcription (Bhatt et al., 2012). However, non-CpG island genes differ

from CpG island-containing genes because they are induced by a larger magnitude after stimulation (Bhatt et al., 2012), possibly because of their tighter control in the basal state.

2.1 Distal *cis*-acting regulatory elements: enhancers

One of the most obvious features of inflammatory responses is that genes activated by identical stimuli differ extensively among cell types, even though induction of these genes depends on inflammatory TFs (such as NF-kB and AP-1 family members) that are ubiquitously expressed. This context-specificity likely depends on the existence of a cell type-specific repertoire of functional *cis*-regulatory elements, namely enhancers (Natoli, 2010).

Recent genome-wide studies revealed that genomic regions annotated as enhancers based on their chromatin signature, extensively diverge in different cell types, suggesting their relevance in cell type-specific transcriptional outputs (Ernst et al., 2011; Heintzman et al., 2009; Pennacchio et al., 2007; Shen et al., 2012; Xi et al., 2007).

Enhancers are short *cis*-regulatory sequences that control transcription of cognate genes acting over different distances (up to about 1Mb) (Bulger and Groudine, 2011). For example, multiple elements distributed over an 800kb gene desert participate in the activation of the *Hoxd* gene cluster, as demonstrated by a combination of genetic studies and biochemical analyses of 3D interactions (Montavon et al., 2011). Another example is provided by the limb-specific enhancer of the *Shh* gene (encoding the Sonic hedgehog morphogen), which is separated from its target by a 1Mb genomic interval (Sagai et al., 2005).

Enhancers are characterized by a specific chromatin signature consisting of high levels of H3K4me1 and comparatively much lower or absent levels of H3K4me3, which is instead associated with active or poised promoters (Barski et al., 2007; Heintzman et al., 2009; Heintzman et al., 2007; Zhou et al., 2011). Enhancers show distinct activity states defined by a combination of additional marks.

Active enhancers are bound by histone acetyltransferases, such as p300 and CPB, which induce H3 and H4 acetylation. Moreover they are often associated with RNA Pol II, which generates a variety of non-coding transcripts known as enhancer RNAs (eRNAs). Moreover enhancer nucleosomes contain histone variants, mainly H2A.Z and H3.3 (Barski et al., 2007; Calo and Wysocka, 2013; Ghisletti et al., 2010; Heintzman et al., 2009; Visel et al., 2009; Zlatanova and Thakar, 2008). H2A.Z deposition may create domains of nucleosomal instability that may facilitate initial TF binding events (Hu et al., 2013; Spitz and Furlong, 2012).

Poised enhancers are characterized by the absence of H3K27ac and RNA Pol II, while *repressed* enhancers may show high levels of H3K27me3, no H3K27ac and the binding of transcriptional corepressors (Creyghton et al., 2010; Rada-Iglesias et al., 2011).

Several studies showed that environmental stimuli induce dynamic changes in chromatin marks. For example, the treatment of macrophages with LPS causes the loss or gain of p300 binding, H3K27Ac and H4Ac at thousands of enhancers, thus leading to the transition between different activity profiles (Chen et al., 2012; Garber et al., 2012; Ghisletti et al., 2010; Ostuni et al., 2013).

Current models suggest that the enhancer repertoire characteristic of a given cell is organized during development by a specific set of lineage-determining TFs which create a specific regulatory landscape controlling the transcriptional output. Environmental stimulation can also induce the *de novo* deposition of enhancer chromatin marks at previously silent sites (Ostuni et al., 2013). This kind of regions were defined as *latent* enhancers and are characterized by the absence in the basal state of specific histone modifications and TFs binding. After stimulation these elements are bound by TF and acquire H3K4me1 and H3K27Ac with relatively slow kinetics. The formation of stimulus-dependent latent enhancers has been observed in macrophages subjected to a subset of unrelated stimuli such as TLR agonists and pro/anti-inflammatory cytokines, suggesting

that cells may respond to environmental changes also expanding their *cis*-regulatory repertoire (Ostuni et al., 2013).

2.2 Enhancer-Promoter Interactions

Long-range interactions between enhancer elements and target genes are mediated by direct physical interactions enabled by three-dimensional chromatin folding (Dostie and Dekker, 2007). Specifically, enhancer-promoter contacts require the looping out of the intervening sequences that separate them in the linear genome (Figure 2) (Bulger and Groudine, 2011; de Laat et al., 2008; Deng et al., 2014).

The establishment of close proximity between enhancers and target genes through chromatin loops has been confirmed by studies of nuclear architecture based on Chromosome Conformation Capture (3C) and its high-throughput variants (Dekker et al., 2002). For example the formation of contacts between the β-globin gene and its Locus Control Region (LCR) was shown to temporally correlate with transcriptional activation (Palstra et al., 2003). Another example is provided by the mammalian α-globin genes which are controlled by distal enhancer elements looped onto the α-globin promoters (Vernimmen et al., 2007; Vernimmen et al., 2009).

Chromatin loops between enhancers and promoters form over various distances in *cis* (more than 100Kb and up to 1Mb) and much more uncommonly in *trans (Sanyal et al., 2012).*

The precise method of loop formation is not yet understood. Two different mechanisms have been proposed to explain how loops between enhancers and promoters are established: the random collision within the nucleus (Figure 2a) and the tracking model (Figure 2b) in which RNA polymerase II (Pol II) complexes migrate along the chromatin fiber until they reach a promoter (Hatzis and Talianidis, 2002; Wang et al., 2005). While the Pol II complex moves toward the promoter, the chromatin between these two elements loops out. A few studies reported that Pol II interacts with both promoters and enhancers,

contributing to the formation of the loop (Szutorisz et al., 2005). In addition to Pol II, other proteins such as CTCF (a sequence-specific DNA binding proteins) and cohesins participate in the formation and stabilization of chromatin loops (Bellomy and Record, 1990). In various experimental systems CTCF has insulator activity, namely the ability to block the communication between an enhancer and a promoter if bound to DNA in between them (Burcin et al., 1997; Filippova et al., 1996; Klenova et al., 1993; Kohne et al., 1993; Lobanenkov et al., 1990). However, others demonstrated that CTCF could also serve as a transcriptional activator in a different sequence context (Vostrov and Quitschke, 1997). In mammalian genomes, CTCF has a complex distribution and binds many enhancers, promoters and even regions located in between an enhancer and its cognate promoter, suggesting a more complex panel of functions (Kim et al., 2007). The location of CTCF-binding sites largely differs from that of TFs. About half of the CTCF-binding sites are located far away from the TSS defining insulators and boundaries for gene clusters. CTCF sites are also located within genes, probably segregating alternative promoters within a single gene and thus contributing to an alternative usage of the promoter (Kim et al., 2007).

A role of CTCF and cohesins in the formation of chromatin loops has been demonstrated by 3C-based approaches. For example, 3C studies characterizing the *H19/Igf2* imprinting control region (ICR) suggested that CTCF regulates the activity of *H19* and *Igf2* genes by mediating intrachromosomal looping interactions (Engel et al., 2008; Yoon et al., 2007). Using 3C, De Laat and colleagues demonstrated that, during tissue-specific activation of selected globin genes, CTCF-bound regulatory sequences throughout the *β*-globin locus come into spatial proximity, thus forming an 'active chromatin hub' (ACH) (Palstra et al., 2003; Splinter et al., 2006). 3C analyses also demonstrated that CTCF-mediated long-range interactions occur at a subset of genes within the human major histocompatibility complex class II (MHC-II) locus. In particular, CTCF binds the *XL9* intergenic enhancer and two divergent promoters upstream of *HLA-DRB1* and *HLA-DQA1* genes, thus forming a loop

(Majumder et al., 2008). Furthermore, cohesins and CTCF sites appear to be associated with loops surrounding promoter-enhancers modules, while enhancers and promoters interactions are mediated by CTCF-free cohesion sites (DeMare et al., 2013; Kagey et al., 2010; Phillips-Cremins et al., 2013). Finally, studies with artificial enhancer-promoter arrays also demonstrated that protein-protein interactions between TFs bound to regions that are far away in the linear sequence contribute to the formation of loops between enhancers and promoters. (Nolis et al., 2009).

The probability of enhancer-promoter interactions through loop formation depends on many variables such as the flexibility of the looped chromatin, the occupancy of the sites by cognate TFs and the distance between the interacting elements. Chromatin remodelers and modifiers recruited to enhancers and promoters by TFs can modify biophysical features of the surrounding chromatin, thus facilitating DNA looping (Li et al., 2006). Furthermore, enhancers and promoters highly occupied by TFs are more prone to establish chromatin loops (Nolis et al., 2009). Finally, due to the physical properties of the chromatin fiber, the probability of an interaction is higher when the enhancer-promoter distance is shorter (Marenduzzo et al., 2007).

Whether the enhancer-promoter loop is a cause or an effect of transcription activation is still poorly understood. Recent studies support a causative role. For example, loop formation between the *β-globin* LCR and gene in erythroid cells occurs before gene activation suggesting that looping is not a result, but a cause of gene activity (Ragoczy et al., 2003). Furthermore, it has been demonstrated that forced chromatin looping between the *γ-globin* promoter and the LCR is sufficient to activate γ-globin expression in adult human erythroid cells (Deng et al., 2014).

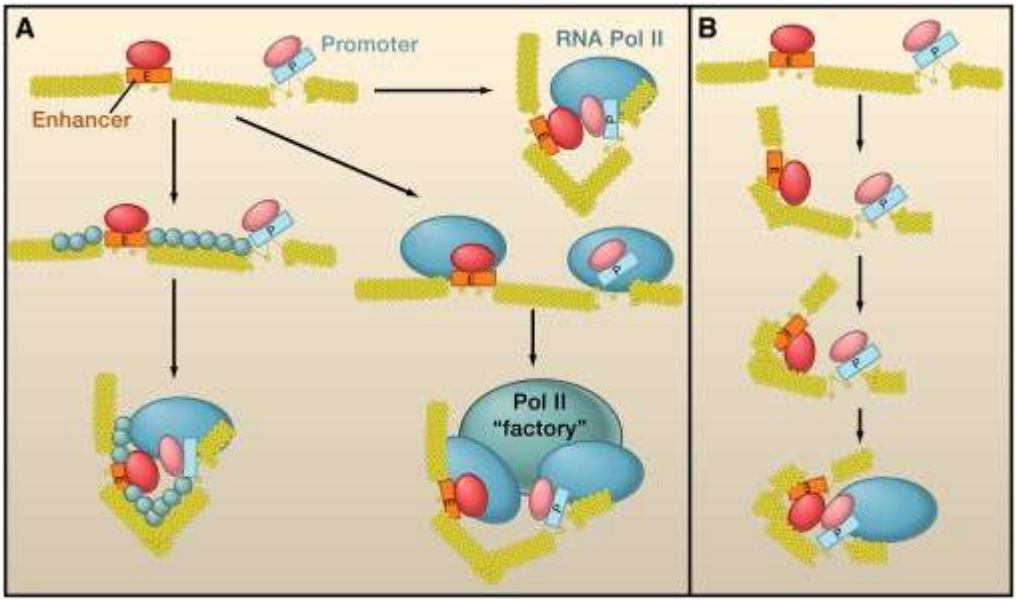

Figure 2 - A simplified view of enhancer-promoter looping interactions
(a) Transcription is activated through the interaction between cognate transcription factors that can interact through random collisions in the nucleus (horizontal arrow), thanks to the help of additional DNA binding proteins (blue circles) that bring the enhancer and promoter into proximity (vertical arrows). Otherwise both enhancer and promoter can interact with RNA polymerase II (diagonal arrow). (b) Tracking model. The enhancer-bound Pol II complex (red ovals) moves along the DNA until it finds the promoter (Pink oval). Adapted from Michael Groudine, Mark, Bulger Functional and Mechanistic Diversity of Distal Transcription Enhancers. Cell. 2011, VOLUME 144.

2.3 The myeloid TF PU.1 and the establishment of the enhancer repertoire in macrophages

The establishment of cell-type specific repertoires of enhancers is driven by the specific set of TFs active in a given cell type. Examples of these TFs include PU.1 and C/EBPβ in macrophages, B cells and dendritic cells (Garber et al., 2012; Ghisletti et al., 2010; Heinz et al., 2010), GATA-1 in erythroid differentiation (Ohneda and Yamamoto, 2002), and KLF4, SOX2 and OCT4 in human induced pluripotent stem (iPS) cells (Soufi et al., 2012). The hematopoietic system-restricted Ets-family TF PU.1 is important for myeloid and B lymphocyte lineage development as demonstrated by the absence of mature myeloid and B cells in PU.1$^{-/-}$ mice (Scott et al., 1994). The regulation of PU.1 protein concentration has a large impact on B cells and myeloid cells development, with high levels favoring macrophage differentiation (PU.1 concentration in B cells is ten-fold lower than in macrophages) (DeKoter and Singh, 2000).

Furthermore, PU.1 overexpression in fibroblasts induces their trans-differentiation into macrophage-like cells, while its absence blocks terminal macrophage differentiation from myeloid precursors indicating that its activity is necessary and sufficient to specify macrophage identity (Ghisletti et al., 2010; Heinz et al., 2010).

It has been demonstrated that PU.1 binds to the vast majority of macrophage enhancers and transcription start sites (TSS), and that it directly regulates both the deposition of enhancer-specific chromatin marks (particularly H3K4me1) and the displacements of nucleosomes, thus generating accessible stretches of DNA sequences. Nucleosome displacement by PU.1 enables the binding of other TFs that would otherwise be unable to invade nucleosomal DNA (Barozzi et al., 2014; Ghisletti et al., 2010; Heinz et al., 2010). This ability to bind sites in a nucleosomal context, thus making them accessible, is considered the peculiar property of a specific subset of TFs called *pioneer* TFs (Zaret and Carroll, 2011).

These accessible sites can be occupied in a constitutive manner by *primer* TFs which prepare possible inducible responses, probably by maintaining the chromatin in an accessible state. Once activated by stimulation, *effector* TFs coordinate the expression of specific groups of genes binding to specific subset of already accessible regulatory elements (Biddie et al., 2011; Garber et al., 2012; Natoli, 2010).

A study using an high-throughput chromatin immunoprecipitation method (HT-ChIP) in mouse dendritic cells treated with LPS at different time points, showed that these accessible sites can be occupied in a constitutive manner by *primer* TFs, such as such as Junb, Irf4 and Atf3, which prepare possible inducible responses, probably by maintaining the chromatin in an accessible state. Once activated by stimulation, *effector* TFs (such as NF-kB, IRF, AP-1 and STATs family members) coordinate the expression of specific groups of genes binding to specific subset of already accessible regulatory elements and leading to the induction of specific proinflammatory genes (Fugure 3) (Biddie et al., 2011; Garber et al., 2012; Natoli, 2010).

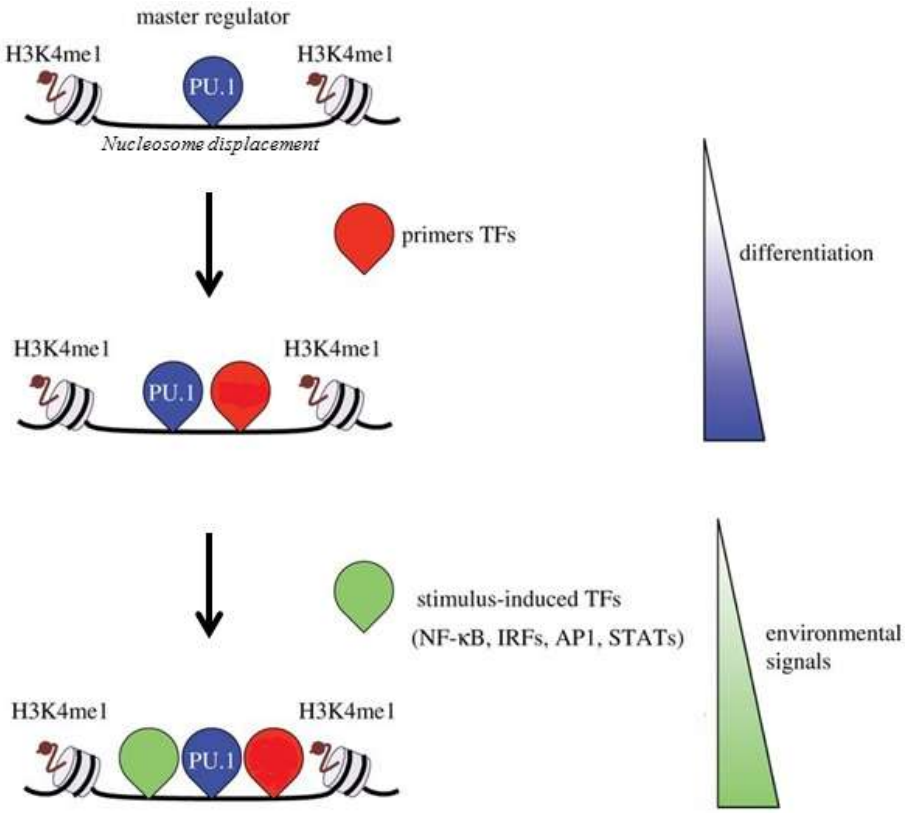

Figure 3 - The master regulator PU.1 organizes the enhancer landscape in macrophages
PU.1 is able to bind regulatory elements displacing nucleosomes and promoting the deposition of H3K4me1. This allows the binding of *primer* TFs, namely TFs which prime for activation regions associated with stimulus-dependent gene induction. Stimulus-dependent TFs are then recruited after stimulation (e.g. LPS).

The unique distribution of PU.1 in macrophages suggests that it could directly promote the looping of distant enhancers onto cognate TSSs or the reciprocal looping between distant regulatory elements, thus favoring long-distance interactions essential for gene regulation and organizing the genomic landscape of macrophages in both the 1-D and 3-D space (Ghisletti et al., 2010; Heinz et al., 2010). This hypothesis is supported by recent findings showing that in Irf8+ dendritic cell progenitors, high PU.1 levels are necessary for inducing Irf8 expression by remodeling its local chromatin interactions to loop a distant *cis*-enhancer into physical proximity to the *Irf8* promoter (Schonheit et al., 2013). Furthermore, in hematopoietic stem cells PU.1 autoregulates its expression by binding at its own upstream regulatory region (URE) and mediating the formation of a chromosomal loop (Staber et al., 2013).

3. 3C-BASED METHODS

In the recent years, Chromosome Conformation Capture (3C) and related methods (such as 4C, 5C and Hi-C) allowed defining some basic principles of the 3D genomic organization. All 3C methods are based on the principle that the spatial proximity of distant genomic regions is transformed into a ligation event that covalently links sequences that are separated in the linear genome but close to each other in the 3D space. Intact nuclei are crosslinked using formaldehyde, cut with a restriction enzyme and subsequent ligated (Dekker et al., 2002; Miele and Dekker, 2008). Using this procedure, DNA fragments that are far away in the linear genome, but adjacent in the 3D space, can be ligated to each other. Ligation events can be measured by the amplification of PCR products obtained with primers on the two sides of the junctions obtained after the ligation (Figure 4a). 3C and 5C allow identifying interacting elements between selected regions of the genome. 3C allows testing only a few interactions selected from candidate genomic regions (Dekker et al., 2002), while the 5C method enables large-scale analysis of broad loci (up to a few Mbs) (Dostie and Dekker, 2007; Ferraiuolo et al., 2012). The 4C technology allows screening the entire genome in an unbiased manner for DNA segments that physically interact with a single DNA fragment of choice (Simonis et al., 2006a; Zhao et al., 2006), while the Hi-C method is a genome-wide adaptation of 3C which provides a true all-by-all genome-wide interaction map with a resolution dependent on the depth of sequencing (Dekker et al., 2013; Lieberman-Aiden et al., 2009). ChIA-PET and ChIP-loop utilize an immunoprecipitation step to enrich only for chromatin contacts mediated by a specific protein in a genome wide manner (Figure 4b).

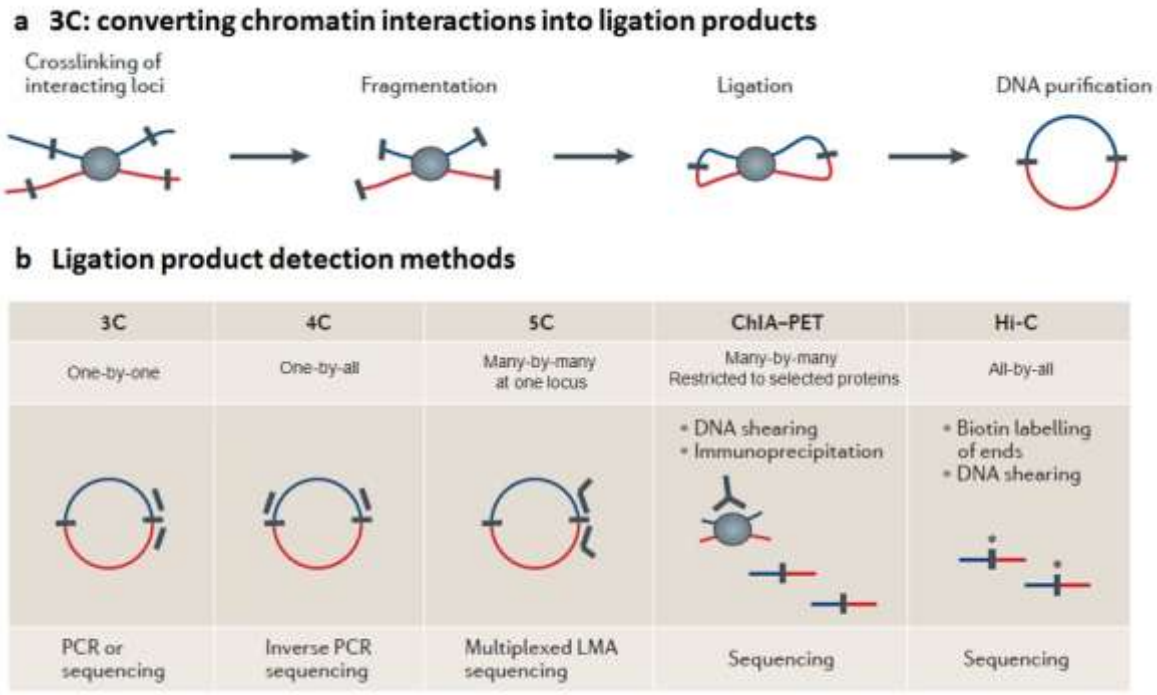

Figure 4 - Schematic representation of 3C-based methods
(a) All the 3C-based methods rely on the conversion of chromatin interactions into ligation products. First, cells are treated with formaldehyde in order to crosslink chromatin segments in spatial proximity. Then chromatin is fragmented using restriction enzymes or sonication. The fragments obtained are ligated and finally the DNA is purified and analyzed. (b) 3C-based methods can detect chromatin physical interactions focusing on different regions of the genome. In the classical 3C locus-specific primers are used to detect one at the time, by PCR, single ligation products. 4C allows the generation of genome –wide interaction profiles for single loci through inverse PCR. 5C uses 3C and hybrid capture approaches to map all the interactions occurring between two sets of large loci (for example promoters and enhancers) within a given region of the genome. Hi-C generates an all-by-all genome-wide interaction map which resolution is dependent on the sequencing depth. ChIA-PET allows the genome-wide analysis of chromatin interactions occurring between sites bound by a protein of interest. Adapted from (Dekker et al., 2013).

3.1 The Carbon Copy Chromosome Conformation capture (5C)

The 3C-carbon copy (5C) technology can detect millions of chromosomal interactions in parallel, increasing the throughput of a 3C. It combines 3C with ligation-mediated amplification (LMA). Ligation events are detected by microarray hybridization or high-throughput DNA sequencing (Dostie and Dekker, 2007).

In the 5C technique, a conventional 3C library is copied into a 5C library using LMA with a mixture of hundreds of 5C primers annealing head-to-head to the 3C ligation junctions. Primer pairs annealed next to each other are then ligated with a NAD dependent DNA ligase, thus generating a 5C library that is subsequently amplified with universal PCR primers (T3 and T7 primers) annealing to the universal tails of the 5C primers (Figure 5).

Therefore, the starting 3C library establishes which 5C ligation products are generated and how frequently they occur (Dostie and Dekker, 2007).

The resolution of a 5C depends on the spacing between neighboring oligonucleotides on the linear genome and it can never reach the one of a 4C or a HiC because not all the ends of the restriction fragments allow the design of a 5C oligonucleotide. However, 5C provides a matrix of interactions frequencies for many pairs of sites in a given locus enabling the construction of the average 3D conformation of a genomic region of interest (de Wit and de Laat, 2012).

The 5C primer design scheme can be adapted to address different kinds of questions. In an *alternating scheme,* forward and reverse 5C primers are designed alternatively on consecutive restriction fragments. This scheme is useful to obtain a general idea of the 3D organization of one locus or to assess the presence of long-range looping interactions in an unbiased manner. In an *anchored scheme,* 5C primers are designed at specific genomic elements such as TSS or enhancers. This allows investigating the interaction pattern of a specific class of elements with other elements or with the surrounding genomic regions. In the *mixed scheme* both alternating and anchored designed primers are included to examine chromatin architecture (Ferraiuolo et al., 2012).

Up to now, the 5C technology has been used to study many regions such as the human α- and β-globin loci (Bau et al., 2011; Dostie et al., 2006), the human *HOXA-D* gene clusters (Ferraiuolo et al., 2010; Fraser et al., 2009; Wang et al., 2011) and a region on the X-inactivation centre in mouse (Nora et al., 2012) .

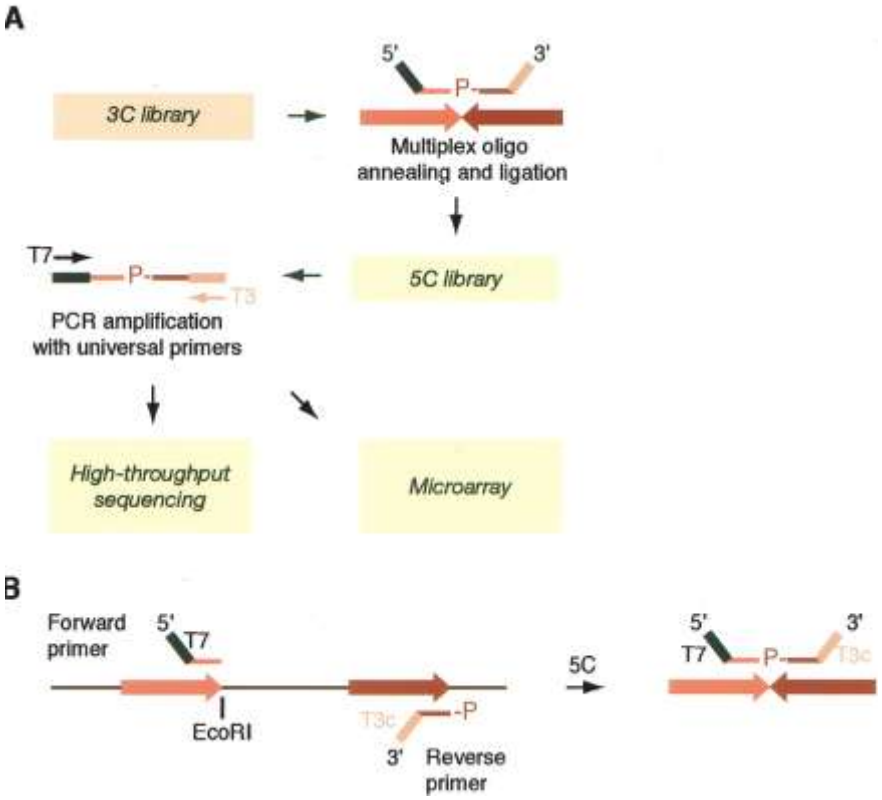

Figure 5 - Schematic representation of the 3C-carbon copy (5C) methodology
(a) A conventional 3C is converted into a 5C library by annealing and ligating 5C oligonucleotides in a multiplex setting. 5C libraries are then amplified with universal primers and analyzed on a microarray or by quantitative sequencing. (b) 5C primer design. Forward 5C primers include half of the restriction site of choice and anneal to the sense strand of the 3′-end of restriction fragments. Forward primers have a common 5′-end tail with the T7 promoter sequence. Reverse 5C primers anneal to the antisense strand of the 3′-end of restriction fragments and including half of the restriction site of choice. Reverse primers show a common 3′-end tail containing the complementary T3 sequence (T3c). Reverse primers are also phosphorylated at the 5′-end. 5C forward and reverse primers anneal to the 3C ligation products in a head-to-head manner. Adapted from (Dostie et al., 2006).

3.2 The Circular Chromosome Conformation Capture (4C)

The Circular Chromosome Conformation Capture (4C) has been developed to analyze all the DNA fragments interacting with a selected region called *viewpoint* and it combines the 3C technology with microarrays or next-generation sequencing (NGS) (Simonis et al., 2006b; Splinter et al., 2011).

In 4C a classical 3C library is digested a second time with a 4bp cutter restriction enzyme. The fragments obtained are ligated to form small DNA circles containing the 3C ligation junctions. Then, viewpoint-specific primers are used in a PCR reaction in order to specifically amplify all the fragments interacting with the chromosomal site of choice. The

PCR product is subsequently analyzed by microarrays or next-generation sequencing (NGS). The viewpoint-specific primer (the "reading primer") is designed to hybridize near the primary restriction site of interest, while the reverse primer anneals to the secondary restriction site closer to the viewpoint (Figure 6) (Simonis et al., 2006b; Simonis et al., 2007).

The theoretical resolution of a 4C depends on the cut frequency of the first restriction enzyme. The classical 4C strategies are based on enzymes recognizing a 6-nucleotide (nt) sequence, cutting in average once every few kilobases. However, the fragments obtained this way are larger than the average regulatory sequences, which are usually several hundred base pairs long. For this reason more recent 4C strategies use 4bp cutters that cut in average every ca. 250 bp, thus in principle allowing the detection of contacts between a gene and its regulatory sequences (Lower et al., 2009).

The 4C technology represents the principal method to investigate the DNA contact profile of individual genomic sites with larger regions elsewhere on the chromosome (*cis*-interactions) or on different chromosomes (*trans*-interactions). So far, the 4C technology has been applied in different studies, for example to study the DNA interaction of the active α-globin and β-globin loci (Schoenfelder et al., 2010b) or to define the topology of the inactive and active X chromosome (Wutz et al., 2002).

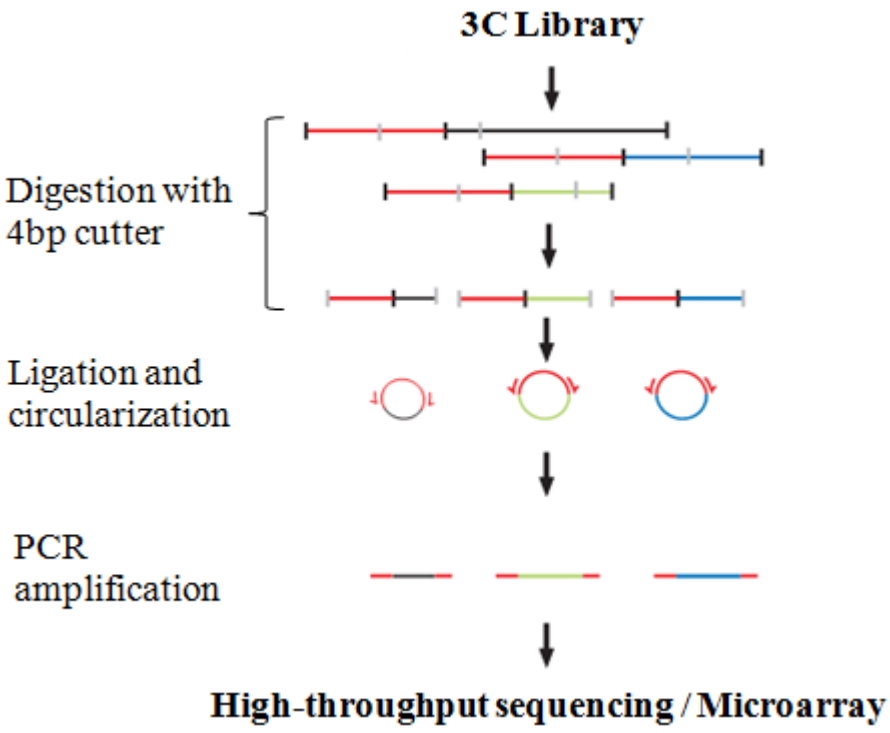

Figure 6 - Schematic overview of the 4C strategy
A 3C library is digested with a second restriction enzyme. The obtained fragments are religated to form small DNA circles that are amplified by inverse PCR using bait-specific primers facing outward. Black vertical lines represent first ligation junctions, while gray vertical lines represent secondary ligation Junctions. Adapted from (Simonis et al., 2007).

4. THE HIERARCHICAL ORGANIZATION OF THE GENOME

To fully understand how the genome works, it is important to investigate not only the information encoded in its sequence, but also the way this is physically and structurally organized within chromosomes. Assays showing the preferential sensitivity of active genes to DNase I first demonstrated the importance of spatial organization for the regulation of gene expression (Weintraub and Groudine, 1976).

Over the last century, increasing interest on this topic has been triggered by the observation that mutual positioning of genes and regulatory sequences in the nucleus might contribute to transcriptional control. In this regard detailed studies using various improved imaging techniques in individual cells and probing the folding of chromosomes across cell populations with 3C-based technologies have investigated chromosome organization at increasing resolution and details, revealing that chromatin is non-randomly organized in eukaryotic nuclei (Miele and Dekker, 2008; Misteli, 2007). Notably, it has been shown that

although general principles regulating chromosome folding apply to all cells, the spatial organization of the genome varies among cells so that individual nuclei display a distinct set of interactions (Croft et al., 1999; Orlova et al., 2012; Parada et al., 2003; Strickfaden et al., 2010).

4.1 Chromosome territories, "A" - "B" compartments and TADS

Chromosome painting and genome-wide Hi-C analyses demonstrated that within the nucleus of interphase cells individual chromosomes occupy distinct chromosome territories (Figure 7a) (Lichter et al., 1988; Lieberman-Aiden et al., 2009; Pinkel et al., 1988; Zhang et al., 2012). Chromosome territories can intermingle in some areas allowing physical contacts and possibly functional interactions between loci located on different chromosomes (Branco and Pombo, 2006; Misteli, 2007). However, loci located on the same chromosome are more likely to interact with each other than two loci located on different chromosomes (Lieberman-Aiden et al., 2009; Zhang et al., 2012). Furthermore, chromosome of similar size and gene density interact far more frequently in the nuclear space: short and gene dense chromosomes tend to localize near the center of the nucleus, while transcriptionally inactive, long chromosomes tend to associate to each other at the nuclear periphery (Croft et al., 1999; Guelen et al., 2008; Zhang et al., 2012).

Euchromatic gene-rich areas and heterochromatic gene-poor chromosomal regions belonging to the same chromosome are spatially separated in different sub-nuclear regions (Fraser and Bickmore, 2007; Misteli, 2007; Naumova and Dekker, 2010). Hi-C data revealed that the somatic cell genome is divided into region of open and closed chromatin named A and B compartment, respectively (Figure 7b) (Lieberman-Aiden et al., 2009; Zhang et al., 2012). The A compartment contains gene-rich, transcriptionally active and DNase I hypersensitive regions, while loci found in B compartments are relatively gene poor, transcriptionally silent and less sensitive to DNase I (Lieberman-Aiden et al., 2009; Zhang et al., 2012). These observations suggest a general trend in which loci with similar

genomic content and chromatin status tend to be proximal to each other and keep their distance from loci with different status.

Recent studies based on high-resolution Hi-C and 5C demonstrated that, within larger A and B compartments, human, mouse and Drosophila chromosomes are segmented into Mb-sized *topological domains* or *topological associating domains* (TADs) (Figure 7c) (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). TADs span >90% of mammalian genomes, are smaller than A/B compartments (their median size being 800Kb) and are characterized by genomic fragments with a higher tendency to interact with each other than with loci located in adjacent domains (Dixon et al., 2012). In contrast to A and B compartments that are related to cell-type-specific gene expression, gross domain organization appears stable across different cell types and conserved in evolution. However, regions within each domain may be dynamic and probably represent the context in which cell-type-specific regulatory events (Figure 7e) (such as enhancer-promoter interactions) take place (Dixon et al., 2012; Nora et al., 2012). Within TADs, at a sub-Mb scale, two elements represent additional features of the genome: smaller sub-domains termed sub-TADS, and intra-chromosomal 'looping' interactions (Figure 7d) (Jin et al., 2013; Phillips-Cremins et al., 2013; Sanyal et al., 2012; Zhang et al., 2013b). Although the majority of these interactions occur within TADs, less common inter-TAD loops have also been identified (Jin et al., 2013).

TADs are separated by genetically defined boundary elements. Nora et al. (2012) demonstrated that deleting a boundary region between two TADs in the X-Chromosome inactivation center led to partial fusion of the two flanking TADs (Nora et al., 2012). Mechanisms establishing TADs are still not clear. However, genome-wide analysis of boundary regions indicated that these spatial modules are statistically enriched in peculiar features such as TSS, and binding sites for the CTCF protein (Phillips and Corces, 2009). However, CTCF also frequently binds sites within TADs suggesting that CTCF binding is not sufficient to create a boundary (Dixon et al., 2012; Nora et al., 2012). Indeed, the

integration of high-resolution 5C maps with genome-wide maps of protein occupancy revealed that the 80% of looping interactions in ES cells were mediated by a combination of CTCF, Mediator and cohesin (Phillips-Cremins et al., 2013). These and other studies predict that long-range interactions (From 500 kb to 1 Mb) are anchored by CTCF/Cohesin, while within and between sub-TADs (at a length scale < 300 kb), Mediator and cohesin control short-range looping interactions between enhancers and promoters (Apostolou et al., 2013; Kagey et al., 2010; Phillips-Cremins et al., 2013; Zhang et al., 2013a). Recently, enrichment of clusters of pluripotent transcription factors in very long range interactions (>5 Mb) have been discovered by Hi-C analysis in ES cells (de Wit et al., 2013). The absence of cohesin enrichment in these analyses suggested that TFs may assist chromatin interaction between TADs independently from architectural proteins.

Thus far, TADs or TAD-like structures have not been described in bacteria, yeast or plants (Duan et al., 2010; Moissiard et al., 2012; Umbarger et al., 2011).

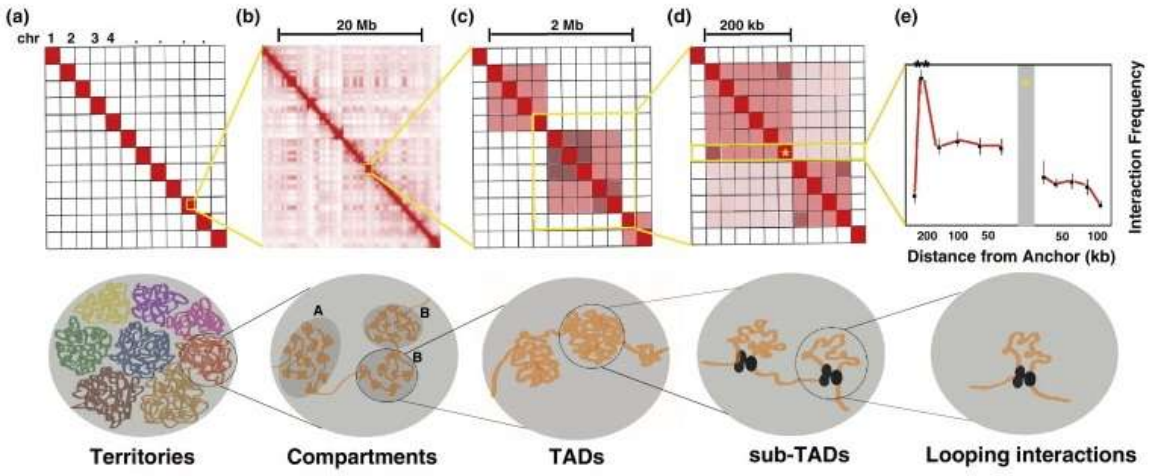

Figure 7 - The hierarchical organization of the genome
(a) Chromosome Territories. (b) A and B compartments are contained inside chromosome territories. (c) TADs within compartments. (d) TADs contain sub-TADs. (e) Intrachromosomal looping interactions take place within and between sub-TADs and TADs. Heatmaps show chromatin interaction data. The signal ranges from lowest (white) to highest (dark red). Adapted from (Phillips-Cremins, 2014).

4.2 The role of Mb-sized domains in gene expression

Each domain contains an average of tens of genes and hundreds of enhancers, and most specific long-range interactions between these regulatory elements are found between the boundaries of TADs, as showed by recent Hi-C experiments (Shen et al., 2012). This observation suggests that TADs probably regulate gene expression providing a locally constrained volume in which looping interactions between genes and regulatory elements occur. Consistent with this hypothesis, the *Shh* enhancer previously discussed is located 1Mb away from the *Shh* gene, but nevertheless it is contained within the same TAD (Sagai et al., 2005). It has been proposed that TADs may function to restrict promoters to their current or future regulatory landscapes thus limiting gene access to a certain set of distal regulatory elements. Consistent with this idea is the organization of the mouse *HoxD* cluster which is located at the boundary between two TADs in ES cells. The *HoxD* genes seem to switch their intra-chromosomal contacts from one domain to the other one during the process of early and late limb patterning (Andrey et al., 2013). It is possible that constraining promoters to a small number of regulatory elements leads to the co-regulation of genes located in the same TAD at specific times during development. Indeed, studies on the X-chromosome inactivation center revealed that, during cell differentiation, genes located in the same TAD tend to be coordinately expressed, possibly because they have access to the same set of regulatory elements (Nora et al., 2012). However, there are many cases across the genome in which co-regulated genes are not delimited by TADs, leaving the role of these spatial modules an open question.

4.3 Organization of intra-chromosomal looping interactions at the sub-MB scale

While gross TADs structure appears invariant across different cell types and conserved across evolution, the internal organization within each domain appears to be highly dynamic (Dixon et al., 2012; Nora et al., 2012). For example it has been shown that chromatin is widely reorganized at the sub-Mb scale during the differentiation of ES cells

in NPCs (Phillips-Cremins et al., 2013). Similar changes in chromatin architecture have been observed also during reprogramming: enhancer-promoter interactions of pluripotent genes identified in ES cells are lost in MEFs and eventually re-acquired upon full reprogramming to induced pluripotent stem (iPS) cells (Apostolou et al., 2013; Denholtz et al., 2013; Wei et al., 2013; Zhang et al., 2013a).

Chromatin architecture plays an essential role for gene expression as demonstrated in ES cells where the disruption of enhancer-promoter looping interactions brings to a consistent deregulation of the pluripotent genes Oct4 and Nanog (Apostolou et al., 2013; Levasseur et al., 2008; Zhang et al., 2013a). It is still unknown if enhancer looping is the cause or an effect of transcription. Recently Hi-C studies in IMR90 fibroblasts stimulated with TNF-alpha (TNFα) showed that looping interactions between TNFα–responsive enhancers and promoters exist prior to activation (Jin et al., 2013). However, the temporal resolution of this experiment was limited as it included just one time point (1h) at which many TNF-induced signaling events (e.g. NF-kB activation) have already been switched off and many genes undergone post-induction repression. Similarly, Noonan and colleagues reported that subset of intra-chromosomal contacts located around selected lineage-specific genes in the developing limb bud were already present in ES cells (DeMare et al., 2013). These observations lead to think that a subset of chromatin contacts is already established prior to transcriptional activation and allows rapid gene activation in response to external signals.

5. AIMS OF THE WORK

Although most genomic regulatory elements control transcription of neighboring genes in *cis*-, rules of enhancer-promoter communication and mechanisms enabling gene control by distal enhancers are still largely undefined (Williams et al., 2010).

The main mechanism by which regulatory elements communicate with their cognate target genes is through chromatin looping, which brings into close spatial proximity elements that are widely spaced in the linear genome (Bulger and Groudine, 2011; de Laat et al., 2008; Deng et al., 2014). Looping events show three main characteristics. They are mostly invariant in global organization but many of the them are cell type specific; they are not exclusive (namely, genes can interact with multiple distal elements and *vice versa*) and they frequently occur within TADs while inter domain interactions are less frequent (Shen et al., 2012).

TADs are spatial models in which the genomic DNA of metazoans is organized. Domain organization appears stable across different cell types and conserved across evolution, but regions within each domain that take part in cell-type-specific regulation may undergo dynamic interaction events (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012).

In this context, the impact of perturbations on looping interactions remains only very partially characterized.

Therefore we set out to exploit 4C and 5C techniques to investigate the principles of 3D organization of the macrophage genome in basal conditions and after activation. We also aimed at dissecting the specific role of the macrophage master regulator PU.1 in the 3D organization of genomic loci containing inducible genes.

Specifically, we first determined the macrophage genomic organization at a selected locus using 5C. Then, we employed high-resolution 4C to characterize the impact of inflammatory stimuli on looping interactions of selected inducible genes and to investigate the cell-type specificity of chromatin contacts in the macrophage genome, also defining the effect of reduced levels of PU.1 on 3D organization.

We chose macrophages as model system for several reasons: (1) macrophages have been extensively characterized by previous studies; (2) the master regulators of macrophage identity are known; (3) they constitute a dynamic system where massive transcriptional and epigenomic changes can be induced by external stimuli. For example, we can mimic the innate immune response to bacteria in vitro stimulating macrophages with LPS.

I found that macrophage activation doesn't have any major impact on the global chromatin interaction landscape of a selected locus, but it affects the looping between genes and regulatory elements at only a subset of inducible genes. Furthermore the general topology of the explored loci in macrophages was not affected by the depletion of the master regulator PU.1.

MATERIALS AND METHODS

1. CELL GROWTH CONDITIONS

Bone marrow cells were isolated from C57B6/Jhsd mice and plated in 10 cm dishes for 6 days in 10 ml of BM-medium (Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 20% low-endotoxin fetal bovine serum (FBS), 30% L929 conditioned medium, 1% glutamine, 1% penicillin-streptomicin, 0.5% sodium pyruvate and 0.1% β-mercaptoethanol.). The cells were cultured for 6 days before harvesting and for 5 days when used for infection. Stimulations were carried out at day 6 with LPS from E.Coli serotype EH100 (Alexis) at 10ng/ml.

E14 ES cells were grown in Glasgow Minimum Essential Medium (GMEM) supplemented with 20% FBS, 1:1000 β-mercaptoethanol, 1% glutamine, 1% penicillin-streptomicin, 1% non-essential amino acids, 1% sodium pyruvate and 0.3% leukaemia inhibitory factor (LIF).

NHI-3T3 mouse embryonic fibroblasts were cultured in DMEM supplemented with 10% FBS, 1% penicillin-streptomicin and 1% glutamine.

2. LPS TREATMENT

Each treatment has been performed adding 10ng/ml of LPS from E. coli, Serotype EH100 (Life Sciences) directly into the cell culture media. The efficacy of the treatment was assessed amplifying known LPS inducible genes by qPCR.

3. FORMALDEHYDE CROSSLINKING

BMDM were fixed at RT for 10’ adding formaldehyde from a 36.5% HCHO stock (SIGMA F8775) directly to the culture medium to a final concentration of 1% or 2% (As required for the experiment). Crosslinking was stopped by addition of Tris-HCl pH 7.6 (125mM as final concentration). Cells were collected with a cell scraper, pelleted and stored at -80°C.

E14 ES cells and NIH-3T3 cells were first trypsinized, resuspended in 45 mL of respective growth medium and then the crosslinking was performed as described above.

4. PROTEIN EXTRACTION AND WESTERN BLOT

Cells were harvested using scrapers and collected by one centrifugation at 2000 RPM for 5minutes at 4°C. Whole cell lysate was obtained with a lysis buffer "Buffer 1"(250mM NaCl, NP40 0,2%, Tris-HCl pH8 50mM, EDTA 0,5mM and EGTA 0,5mM) and centrifuged at 13000 RPM for 5 minutes at 4°C. Protease and phosphatase inhibitor were added to lysis buffers.

The proteins obtained were separated according to their molecular weight by electrophoresis in a polyacrylamide gel and transferred into a Protran nitrocellulose filter of 0.45 microns. After the blocking of the non specific sites by incubation in TBST buffer (20mM Tris-HCl pH7.5, 500mM NaCl, 0.1% Tween-20) supplemented with 5% milk, the filter were subjected to hybridization with specific antibodies.

Antibodies used:

Anti PU.1: home made

5. RNA AND cDNA PREPARATION

RNA was extracted using the RNeasy Mini Kit (Qiagen) according to manufacturer instruction. 0.5 ug of total RNA was used for cDNA synthesis (using the ImProm-II Reverse Transcriptase, Promega). RNA was quantified by ND-1000 spectrophotometer (NanoDrop Technologies) and its quality was assessed by measuring A260/A280 and A260/A230ratios.

Complementary DNA (cDNA) was obtained by reverse transcription with the following protocol: RNA was denatured at 65°C for 10min and immediately placed on ice; 0.5µg of denatured RNA were mixed with 1µl 10x Reaction buffer, 2.5µM random primers, 0.5mM dNTP and 0.5µl M-MuLV (Fynnzymes) reverse transcriptase, in a total volume of 10µl.

Enzymatic reactions were performed in a PCR machine: 10 min 25°C (annealing), 60 min 42°C (elongation) and 5 min 75°C (enzyme inactivation). Samples were then diluted in double-distilled water (ddH2O) to a final volume of 100µl.

5µl of the obtained cDNA was generally used as template for qPCR expression analyses. Quantification was performed on Nanodrop.

6. QUANTITATIVE RT-PCR (qPCR)

Real time PCR was performed using the SYBR Green detection system on 7500 Applied Biosystem machine. Briefly, 10µl of SYBR Green Master Mix (2X) (Applied biosystems) were mixed with 0.1µM gene-specific primers, 5µl of cDNA and ddH2O to a final volume of 20µl. Accumulation of fluorescent products was monitored for 40 PCR cycles. Each PCR reaction generated only one specific amplicon, as revealed by the melting temperature profile of final products (dissociation curve). QPCR detection system was updated with the Fast SYBR Green Master Mix during the course of the presented work and the PCR cycle settings adjusted following manufacturer instruction. Error bars in QPCR graphs represent standard deviation of three biological independent replicates unless otherwise stated.

Type of analysis	Type of primer	Forward	Reverse
qPCR	Pu.1	ATGCACGTCCTCGATACTCC	GCTGGGGACAAGGTTTGATA
	CD68	TCCCCACCTGTCTCTCTCAT	TTGCATTTCCACAGCAGAAG
	Emr1	GAGTGCACCCAAGATCCATT	TGGCATTGCTGTATCTGCTC
	Lyz2	GAATGGAATGGCTGGCTACT	CACTGCAATTGATCCCACAG

Table 1 – qPCR primers

7. RETROVIRAL INFECTIONS

Recombinant retroviruses were produced by transient transfection of ecotropic Phoenix cells. In brief, Phoenix cells were trypsinized and plated at 1x106 cells/10-cm plate 48h before transfection. Calcium phosphate were performed with 10ug of the retroviral plasmid of interest. The medium of Phoenix cells was replaced with 10ml of medium with chloroquine. Plasmid DNA was added to 5ul of pCL-Eco packaging vector, 61ul of 2M

CaCl2 and water up to 500ul. Then 500ul of HEPES Buffer2x were dripped into the mix. The mixture was dispensed on the phoenix cells by dripping. After 16h of incubation, the medium was replaced with 5ml of target cells medium. After 24h, NIH-3T3 were subjected for two days to four cycles of infection (2 cycles per day) with the supernatant of Phoenix cells supplemented with Polybrene (8ug/ml). NIH-3T3 cell were subsequently kept under selection for three days (Puromycin 1.5ug/ml). Target gene expression was validated by Western Blot, ChIP and qPCR.

Plasmids:

The MSCV (Murine Stem Cell Virus) is vector optimized for introducing and expressing target genes in different cell lines.

Pu.1 coding sequence :

ATGTTACAGGCGTGCAAAATGGAAGGGTTTTCCCTCACCGCCCCTCCATCGGATGACTTGGTTACTTACG
ATTCAGAGCTATACCAACGTCCAATGCATGACTACTACTCCTTCGTGGGCAGCGATGGAGAAAGCCATAG
CGATCACTACTGGGATTTCTCCGCACACCATGTCCACAACAACGAGTTTGAGAACTTCCCTGAGAACCAC
TTCACAGAGCTGCAGAGTGTGCAGCCCCCGCAGCTACAGCAGCTCTATCGCCACATGGAGCTGGAACAGA
TGCACGTCCTCGATACTCCCATGGTGCCACCCCACACCGGCCTCAGTCACCAGGTTTCCTACATGCCCCG
GATGTGCTTCCCTTATCAAACCTTGTCCCCAGCCCACCAGCAGAGCTCAGATGAGGAGGAGGGTGAGAGG
CAGAGCCCTCCCCTGGAGGTGTCTGATGGAGAAGCTGATGGCTTGGAGCCTGGGCCAGGTCTTCTGCACG
GGGAGACAGGCAGCAAGAAAAAGATTCGCCTGTACCAGTTCCTGCTGGACCTGCTGCGCAGCGGCGACAT
GAAGGACAGCATCTGGTGGGTGGACAAGGACAAAGGTACCTTCCAGTTCTCGTCCAAGCACAAGGAGGCG
CTGGCGCACCGCTGGGGCATCCAGAAGGGCAACCGCAAGAAGATGACCTACCAGAAGATGGCGCGCGCGC
TGCGCAACTACGGCAAGACAGGCGAGGTGAAGAAAGTCAAGAAGAAGCTCACCTACCAGTTCAGCGGCGA
GGTGCTGGGCCGTGGGGGCCTGGCCGAGCGGCGCCTCCCGCCCCACTGA

8. LENTIVIRAL INFECTIONS

Recombinant lentiviruses were produced by transient transfection of 50% confluent 293T cells cultured in 10ml of cell growth media (DMEM, 10% FBS south American, 1% penicillin-streptomycin, 1% glutamine). Calcium phosphate were performed as described

above, but with 8ug of the lentiviral plasmid of interest together with 6ug of psPAX8 plasmid and 3ug of pMD2 plasmid. After 6h of incubation, the medium was replaced with 5ml of target cells medium. 24h later, mouse macrophages were infected with the lentiviral vectors pLKO.1ShPu.1 or pLKO.1ShLuciferase. Cells underwent two cycles of infection over two days starting 5 days after plating and were collected at day 9, 48h post selection with 4ug/ml puromycin. Target gene knockdown was validated by Western Blot, qPCR and ChIP.

Plasmids:

pLKO is a lentiviral vector that directs the expression of short-hairpin RNA (shRNA) sequences. When packaged into a retrovirus, it allows for stable knock-down of target genes through RNA-interefrence.

Sh sequence for Pu.1: GAGCTATACCAACGTCCAATG

9. CHROMATIN IMMUNOPRECIPITATION (ChIP)

For ChIP experiments fixation with formaldehyde (2% as final concentration) and sonication were performed as previously described (Ghisletti et al., 2010). ChIP lysates were generated from 1 x 107 cells and using 10ug of the antibody for PU.1 (Santa Cruz sc-352). The antibody was prebound overnight to 100ml of G protein-coupled para-magnetic beads (Dynabeads) in PBS/BSA 0.5%. Beads were added to lysates and incubated overnight. Beads were then washed six times in this buffer: 50 mM HEPES [pH 7.6], 500 mM LiCl, 1 mM EDTA, 1% NP-40, and 0.7% Na-deoxycholate and once in TE containing 50 mM NaCl. DNA was eluted in TE containing 2% SDS and crosslinks reversed by incubation overnight at 65 C. DNA purification were performed using Qiaquick columns (QIAGEN) and quantified with PicoGreen (Invitrogen). ChIP validation by qPCR has been done using purified DNA for amplification on an Applied Biosystems 7500 Fast Real-time PCR system (SYBR Green, Applied Biosystems). Primers used for ChIP-qPCR are in reported in Table 2**.**

Type of analysis	Type of primer	Forward	Reverse
ChIP-qPCR	PU.1 Positive control 1	GCCTCAATGTATGGGCTTCT	CCCTTTCCCTTAGGAACCAC
	PU.1 Positive control 2	GGAAGTGAGAAAGCCAAGCA	AACCCATGGTGGTCCAGATA
	PU.1 Positive control 3	GGCGTTGGTAAAGACAGCAT	GCTGGAGGCCTCTGTGTAAC
	PU.1 Positive control 4	AGGTCCTAATGGGGTTTTGG	TGCCCTGAACTTTGATGTG
	PU.1 Positive control 5		
	PU.1 Negative control 1	TTTTCCAGGCAAAGCAGATT	ATGTATGGGCACAAGCACAA
	PU.1 Negative control 2	GCAGTAAAAGGTCGCCAGTC	AAGCACAGCCTCGTTCTCTC
	PU.1 Negative control 3	CCCCAGCCAACATGAGTTAC	TCAGGGGAAGCAACAGATCG

Table 2 – ChIP-qPCR primers
List of primers used in this work. Type of analysis for which each primer couple has been used is indicated (top left column). PU.1 positive control primers have been designed on PU.1-rich regions according to published dataset (GEO accession n. GSM594586). PU.1 negative control primers have been designed on PU.1-poor regions, around the TSS of Neg2, Hoxa7 and Hoxa11 respectively.

9.1 ChIP-seq data analysis

After quality filtering, 51 nt long single-end reads were aligned onto the mm9 release of the murine genome using Bowtie v0.12.7 (Langmead et al., 2009). Only unique alignments were retained, allowing up to two mismatches compared to the reference genome. Peak calling was performed using MACS v1.4 (Zhang et al., 2008). Cell type specific inputs were used as controls. In order to visualize the raw profiles on the Genome Browser (Meyer et al., 2013), wiggle files were generated with MACS v1.4 and converted to bigWig.

bigWig files of H3K4me1 and H3K27ac datasets in mESC were downloaded from GEO (Barrett et al., 2013) (GSM1003750 and GSM1000126 records respectively).

10. CARBON-COPY CHROMOSOME CONFORMATION CAPTURE (5C)

10.1 Preparation of 5C templates

5C was performed on mouse macrophages following a previously described protocol (Dostie and Dekker, 2007) with modifications. In brief, 3C templates were obtained crosslinking cells with 1% formaldehyde for 10 min at room temperature. $1x10^7$ cells were

lysed in 500 µl of lysis buffer (10 mM Tris-HCl [pH8.0]; 10 mM NaCl; 0.2% NP-40, 1× protease inhibitor) for 15 min on ice and disrupted with 15 strokes of p1000 pipette. After centrifugation, nuclei were resuspended in 500 µl of digestion buffer and pierced adding SDS (0.1% as final concentration) for 10 min at 65°C. SDS was then quenched with Triton-X100 (1% as final concentration). DNA was then digested by the 6bp cutter HindIII (800U) at 37°C O.N. After inactivation by 1.6% SDS at 65°C for 20 min, samples were diluted in 7.5 ml of 1× ligation buffer and 3000 U (NEB Units) of T4 ligase and incubated at 16°C for 4 hr. Ligated chromatin was digested by proteinase K O.N. to remove cross-links and purified by phenol-chloroform extraction. Finally, DNA was ethanol precipitated. 3 femtomoles of 5C primers were annealed to the junctions of the 3C material O.N. at 48°C, joined with 10U of NAD-dependent ligase for 1h, and then amplified by PCR for 25 cycles using T3 and T7 universal primers, thus generating a quantitative "carbon copy" of a part of the initial 3C library, which has been subsequently analyzed by high-throughput DNA sequencing (30 million-100bp paired-end reads on an Illumina Hi-seq 2000 platform).

10.2 Region of interest and primers design

The 5C experiment was designed to interrogate looping interactions between HindIII fragments containing transcription start sites (TSSs) and any other HindIII restriction fragment (distal fragments) in the selected region of interest ~5Mb interval (80,141,160-85,160,410 on mouse chr 11).

Forward and reverse 5C primers were designed with an anchored scheme using the my5C software (http://3dg.umassmed.edu/my5Cprimers/5C.php). Multiplex 5C libraries were produced by mixing 171 reverse primers annealing to the TSS of all genes in the locus of interest (ca. 3 restriction fragments per TSS), 581 forward primers annealing to all other restriction fragments and 21 reverse primers together with 20 forward primers

corresponding to random restriction fragments on a gene desert region (Chr 14). This design allows simultaneously measuring up to 99,351 possible contacts.

10.3 5C data analysis

After quality filtering, 101 nt long paired-end reads were trimmed (4 bases at the 5' and 50 bases at the 3') using the *fastx_trimmer* tool (which is part of the FASTX-Toolkit, http://hannonlab.cshl.edu/fastx_toolkit/). Novoalign (http://novocraft.com) was user to align the trimmed reads to the primer pool. Considering all the possible forward-reverse pairs, pair-wise interactions were summarized as a matrix using a custom C++ script. Considering a first pilot untreated sample, the distributions of counts for the forward and reverse primers showed that a small fraction of primers accounted for a major fraction of the reads. Primers showing total counts in the top 10 and 20 percent (for the forward and the reverse, respectively) were then excluded from the primers' pool in the next experiments.

The HiTC R package (Servant et al., 2012) was then used for further analyses and plotting. Data normalization was performed, through two steps: 1) *cis* counts on chr11 were linearly scaled to 1e7 reads; 2) linear multiple regression of the *cis* counts on the chr11 over the average *trans* counts (chr11-chr14) of the forward and reverse primers' for each interacting pair was performed in order to obtain *trans*-corrected measurements. More in details, the log2 of the *cis* counts was fitted in function of the sum of the log2 of the forward and reverse average *trans* counts. After regression, each *cis* count was replaced by the corresponding residual, which was then transformed from log to linear. *lm* R function was used to perform linear regression.

Binned matrices were obtained using the *binningC* function of the HiTC package, using a 50 kbps bin and setting *bnorm=T* and *method=median*. Global correlations among samples were performed using normalized, binned *cis* data on chr11. Pair-wise Spearman's rank

correlations coefficients were calculated with the function *cor* of R and plot using the package *corrplot*.

11. HIGH RESOLUTION CIRCULAR CHROMOSOME CONFORMATION CAPTURE (HiRes-4C)

11.1 Preparation of 4C templates

4C templates were prepared as previously described (Splinter et al., 2012). Briefly, from cells suspensions chromatin was cross-linked with 2% formaldehyde for 10 min at room temperature. $1x10^7$ cells were lysed in 500 µl of lysis buffer (50 mM Tris-HCl [pH7.5], 150 mM NaCl, 5mM EDTA, 0.5% NP-40, 1% TX-100, 1× protease inhibitor) for 15 min on ice and disrupted with 15 strokes of p1000 pipette. After centrifugation, nuclei were resuspended in 500 µl of digestion buffer and pierced adding SDS (0.3% as final concentration) for 10 min at 65°C. SDS was then quenched with Triton-X100 (1% as final concentration). DNA was then digested O.N by the 4bp cutter DpnII (800U) at 37°C. After inactivation by 1.6% SDS at 65°C for 20 min, samples were diluted in 7.5 ml of 1× ligation buffer and 3000 U (NEB Units) of T4 ligase and incubated at 16°C O.N. Ligated chromatin was digested by proteinase K O.N. to remove cross-links and purified by phenol-chloroform extraction. Finally, DNA was ethanol precipitated.

The second digestion was performed with 50 units of Csp61 (another 4-bp cutter enzyme recognizing a different sequence than the primary enzyme) and followed again by intramolecular ligation at 16°C O.N. in 14ml of 1x ligation buffer and 6000 U (NEB Units) of T4 ligase. After DNA purification, 200 ng of the obtained 4C template were used to perform 8 individual PCR reactions, which were then pooled together and purified for HT sequencing. The PCR products were purified first with the High Pure PCR Product Purification Kit (Roche cat. no. 11732676001) and then with the QIAquick PCR Purification Kit (QIAGEN cat. no. 28104). The first kit allows the separation of PCR products larger than 120bp from the adaptor-containing primers (which are respectively ~75 nt and ~40 nt long). The second kit increase the purity of PCR products.

11.2 4C primer design

Adaptor sequences necessary for Illumina single-read sequencing (Hi-seq 2000) were added to 4C primer pairs as additional 5′ overhangs. Final sequencing reads were composed of the 4C primer sequence (20 nt for each viewpoint) followed by 16 nucleotides that identify a captured sequence. The reading primer hybridizes to (and ends at) the 3′ side of the first cutting site. The 18-27bp long nonreading primers were designed 120 bp far from the secondary restriction site. In this way only primary ligation events are analyzed and most captured sequences can be identified unambiguously. Primers used for PCR are in reported in Table 3**.**

Jun	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGACAGC TGGAAAGTGATC
	RV	CAAGCAGAAGACGGCATACGAGGCAAAATATCTTAAGGCTTC
Nfkbia	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGGGTT TTCAAAAAGATC
	RV	CAAGCAGAAGACGGCATACGATCTAGTTTAACCGGCAAAAG
Ccl5	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTTTCAGGCC CCTCTTAGGATC
	RV	CAAGCAGAAGACGGCATACGATGCAACACAAGAGCAACTAC
Cxcl10	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTACGCCAAA TTTAGCCAGATC
	RV	CAAGCAGAAGACGGCATACGACAAGGCCATTTAATTAACGA
Il12b	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGCTTTC ACCCAGGGATC
	RV	CAAGCAGAAGACGGCATACGACTAAAGGCCATCACAGGTAG
Nos2	FW	ATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTGACAAGGAA GGCCCAGGATC
	RV	CAAGCAGAAGACGGCATACGATGGGGTCACTAATACAGGAG
Il12a	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTACGACAGG TTCCCTATGATC
	RV	CAAGCAGAAGACGGCATACGACCCTACTTTGCTCTTGAGAA
P2ry2	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTTTGGA GATGGTTGGATC
	RV	CAAGCAGAAGACGGCATACGATGAGAGCGTATTCTCCAGAT
Tcfec	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTATAACCAA GAATAATTGATC
	RV	CAAGCAGAAGACGGCATACGATTAGCATGTGTTCTTTTCCC
Csf2	FW	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCTATAG TGTCACCTGATC
	RV	CAAGCAGAAGACGGCATACGACAGCCAGAAGTCTCTCCTTA

Table 3 – 4C primers
List of primers used to amplify 4C templates. The reading primer hybridizes to the TSS of the gene.

11.3 4C data analysis

Sequencing was performed on Illumina Hi-seq 2000 platform. The mouse mm9 genome was used as the reference genome for mapping 4C sequence captures.

Data from single viewpoints was de-multiplexed using *fastq-multx* from the ea-utils suite (https://code.google.com/p/ea-utils/).

Low quality reads were then discarded and the 4cseqpipe (van de Werken et al., 2012), a recently published pipeline for 4C data analysis, was used to map the reads, compute statistics and calculate and plot the local interaction profile for each experiment. Nearcis

plots were computed for windows of +/- 400 or +/- 250 kbps from the viewpoints using *median* as *stat_type* and *5,000* as *trend_resolution*.

Before computing any correlation or cumulative plot, nearcis data was summarized in 2-kbps bins (by summing up the signal of the fragments in each bin). Local partial correlations were computed separately on each side of the viewpoint, starting with the 10 most proximal windows and iteratively adding one further window at a time. Spearman's rank correlations coefficients were calculated with the *cor.test* function of R.

RESULTS

1. 5C ANALYSIS OF A SELECTED GENOMIC REGION

1.1 Macrophage genomic organization at the *Ccl5* gene locus

To investigate the global 3D organization of the macrophage genome in basal and stimulated conditions, we first determined the looping interactions between enhancers and gene promoters using the 5C technique, which allows a parallel analysis of interactions between many selected DNA fragments in a selected region of the genome (Dostie and Dekker, 2007). The genomic locus analyzed is a 5 Mbp region of chromosome 11 that surrounds the *Ccl5* gene. We selected this locus for three main reasons. First, it includes several chemokine genes activated with distinct kinetics by inflammatory stimulation, such as *Ccl3, Ccl4, Ccl5* and *Ccl9*. Second, it contains both several LPS-inducible genes, and genes unaffected by stimulation, including transcriptionally active and silent genes (Figure 8a and b, which shows the nascent, chromatin-associated transcripts at the central region of this locus). Third, the *Ccl5* gene, together with the chromatin signature at its surrounding region, was extensively characterized in our laboratory (De Santa et al., 2010; De Santa et al., 2009).

Specifically, we analyzed the interactions between 171 TSS-containing restriction fragments (Reverse primers) and 581 distal restriction fragments (Forward primers) covering this locus (Figure 8a). Our design included 20 forward primers corresponding to random restriction fragments at a gene desert located on a different chromosome (Chr 14). Interaction frequencies measured between the locus of interest on chromosome 11 and the gene desert region on chromosome 14 allowed us to evaluate the level of non-specific *trans*-interactions between chromatin fragments not located in spatial proximity and unlikely to have any functional connection.

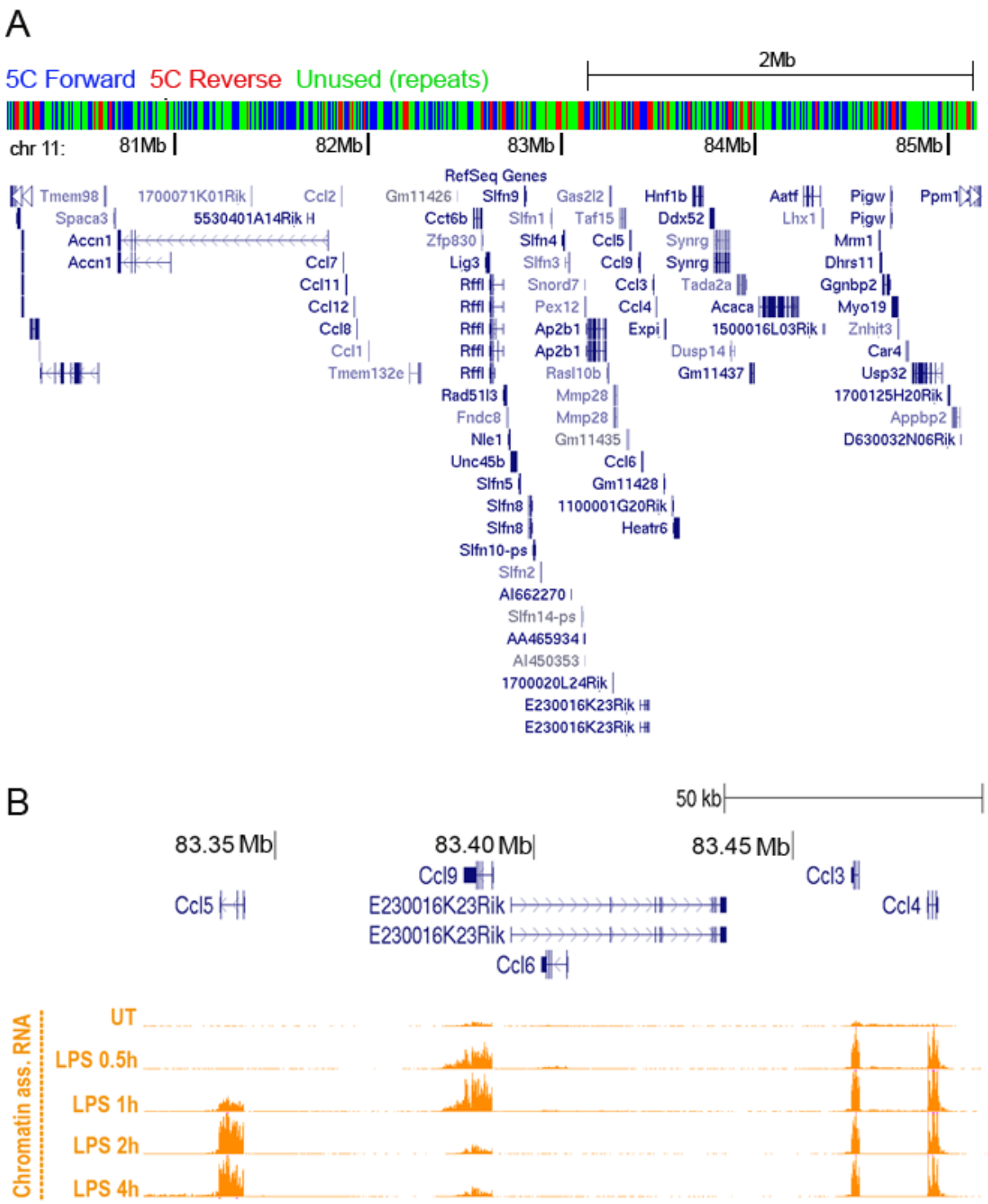

Figure 8 - Approach to identify looping interactions: 5C design
(a) Distribution of 5C –Forward and 5C-Reverse HindIII restriction fragments across the 5Mb analyzed showing positions of RefSeq genes surrounding the Ccl5 gene. Reverse 5C primers were designed for HindIII fragments that contain a TSS (Red) and forward primers for all other "distal" HindIII fragments (Blue). (b) Representative region in the locus of interest showing chromatin associated RNA in untreated or LPS-induced mouse macrophages.

The template 3C library used for 5C analysis was prepared according to the conventional 3C protocol (Naumova et al., 2012). We assessed the quality of the 3C library by semi-quantitative PCR with primer pairs against distant fragments in the region of interest and

designed to specifically amplify ligation junctions. The 3C library displays a single PCR product on agarose gel that increases at higher volumes and progressively disappears at greater dilutions indicating that unique contacts are present in the library and that their intensities increase with increasing library concentration (Figure 9a).

The 5C approach quantifies head-to-head 3C ligation products by ligating pairs of 5C primers annealed on the junctions of the 3C library. Ligated 5C primers are then amplified with T3 and T7 universal primers and used for NGS or microarray detection. The efficiency of annealing of the 5C primers to the 3C ligation products and the efficiency of PCR amplification of ligated 5C primers varies for different 5C primer pools. For this reason, these two steps of the technique required a long and detailed troubleshooting. The annealing procedure was performed using different amounts of 5C primers in order to find the amount necessary to obtain a 5C library. A 'no ligase' control was included to assess the level of ligase-independent background. We also included a control in which 5C was performed in the absence of the 3C library to check for nonspecific ligation of 5C primers. Furthermore, a 'no 5C primers' control allowed us to verify the absence of T3 and T7 primer dimers in the amplified samples. We found that 3 femtomoles of 5C primers were sufficient to obtain a 5C library. However, the appearance of the 5C library was accompanied by detectable levels of ligase-independent background noise and nonspecific ligation products in the control samples (Figure 9b). In order to solve this problem, we determined the minimum number of PCR cycles necessary to minimize amplification biases but sufficient to obtain a detectable amount of 5C libraries (Figure 9c).

The long-range interaction map representing TSS–distal fragment interactions along and between the 5Mb surrounding the Ccl5 gene is shown in Figures 9d and e.

Figure 9d displays the first 5C heatmap we generated. Columns represent reverse (RV) primers designed on all the TSSs contained in this locus, while rows represent forward (FW) primers annealing to the distal fragments. This initial map showed a high background noise with some RV primers making contact with all the FW primers and *vice versa*. We

also found some primers that did not generate any product (completely white lines and columns). However, the expected diagonal representing *cis* interactions occurring between regions located in close proximity along the linear genome was detected for the control region on chromosome 14 (dotted lines, bottom right) and for a few segments in the region of interest on chromosome 11 (black arrows).

In order to improve the quality of the map, we next split into deciles all the 5C primers for the number of detected contacts and eliminated the primers of the first upper deciles. Then, we generated a new 5C library using the restricted set of 5C primers.

The 5C map generated with this new pool of 5C primers (Figure 9e) was much cleaner than the previous one, with many fewer non-specific contacts, and showed the expected general features of spatial chromatin organization. As expected, we observed an inverse relationship between genomic distance and interaction frequency within the considered region. The number of chromatin contacts was higher for pairs of fragments closer in the linear genome. Our map also detected 5C signals outside of the diagonal and representing distant *cis*-interactions (black arrows).

The comparison between 5C and HiC data obtained in our lab (data not published) showed an overall similarity in the global topological features of the region considered, albeit with obvious differences in resolution (Figure 9f). Despite the presence of blind spots in the map due to the anchored scheme used for the design of the 5C primer set, our map readily detected TADs (black boxes) in the locus of interest.

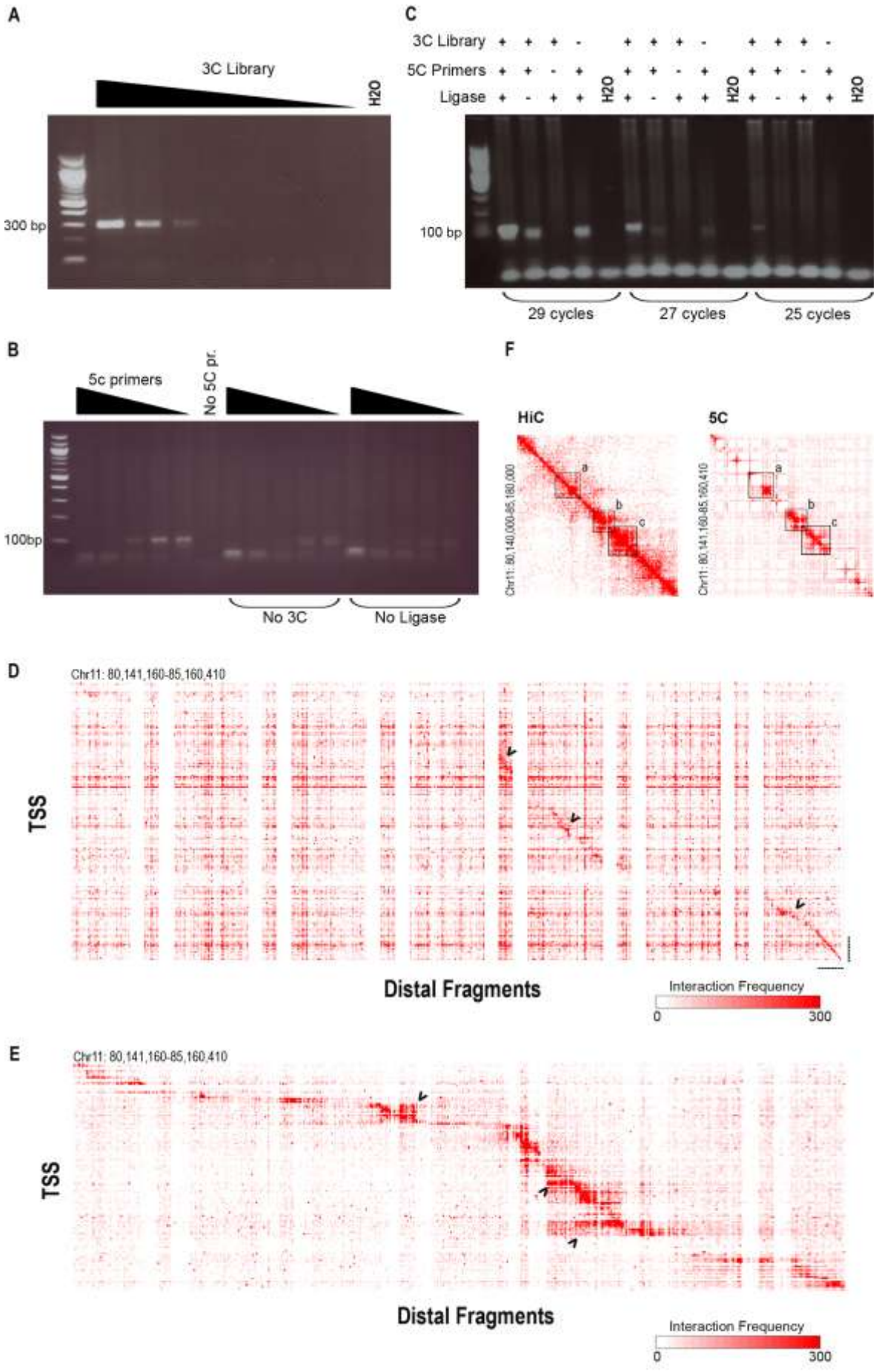
A
3C Library
H2O
300 bp
C
3C Library + + + - + + + - + + + -
5C Primers + + - + + + - + + + - +
Ligase + - + + + - + + + - + +
H2O
100 bp
29 cycles
27 cycles
25 cycles
B
5c primers
No 5C pr.
100bp
No 3C
No Ligase
F
HiC
5C
Chr11: 80,140,000-85,160,000
Chr11: 80,141,160-85,160,410
a
b
c
D
Chr11: 80,141,160-85,160,410
TSS
Distal Fragments
Interaction Frequency
0
300
E
Chr11: 80,141,160-85,160,410
TSS
Distal Fragments
Interaction Frequency
0
300

Figure 9 - 5C Setup
(a) PCR products from a 3C library titration. The 3C library was titrated with a couple of primers recognizing distal restriction fragments in the region of interest. (b) 5C primers titration over a range of 1.5-25 femtomoles as final annealing concentration. (c) 5C library amplification. The 5C library was amplified over a range of cycle numbers to identify the cycle number allowing the loss of aspecific products. (d and e) 5C interactions maps before (d) and after (e) primers screening and elimination. Heatmaps show all interrogated TSS-distal fragments interactions in the 5Mb region surrounding the Ccl5 gene (Chr.11) and in the gene desert region on chromosome 14 (Dotted line). Fragments are visualized in their genomic order using a linear scaling to 10^7 reads. The color of each square is a measure of the interaction frequency between FW and RV primers. The diagonal represents frequent cis interactions between regions located in close proximity along the linear genome. 5C signals away from the diagonal represent long-range looping interactions. (f) HiC and 5C matrices with counts binned in 50kbp windows for visualization.

1.2 5C shows a general conservation of chromatin contacts during the LPS response

We next investigated how TSS–distal fragment interactions relate to gene expression dynamics during the LPS response. Activation of primary mouse macrophages by inflammatory stimuli leads to a robust, highly reproducible and kinetically complex gene expression program that includes primary response genes (usually peaking at 15-30 min after stimulation) and secondary response genes (whose transcription starts in most cases not earlier than 1h post-stimulation). 5C libraries were generated from mouse macrophages treated with bacterial endotoxin (lipopolysaccharide, LPS) from 0.5h to 4h. During this time window a large number of genes in the locus of interest were activated, albeit with distinct kinetics (Figure 8b). Heatmaps of raw 5C data showed an overall high similarity among all the time points at each individual locus, suggesting that region-specific primers amplified each region in a robust and consistent manner. The comparison of 5C interaction maps from untreated and LPS-treated cells revealed a general conservation of chromatin contacts during stimulation (Figure 10a). Spearman's rank correlations calculated on the counts normalized by the number of primers in the bin was very similar among all the time points (in the range of 0.69-0.76 except for the LPS 30min which was slightly less correlated), suggesting that gene activation induced by external stimuli did not impact the global chromatin interaction landscape (Fig. 10b) and that most of the contacts between genes and distal *cis*-regulatory elements in this region are pre-formed.

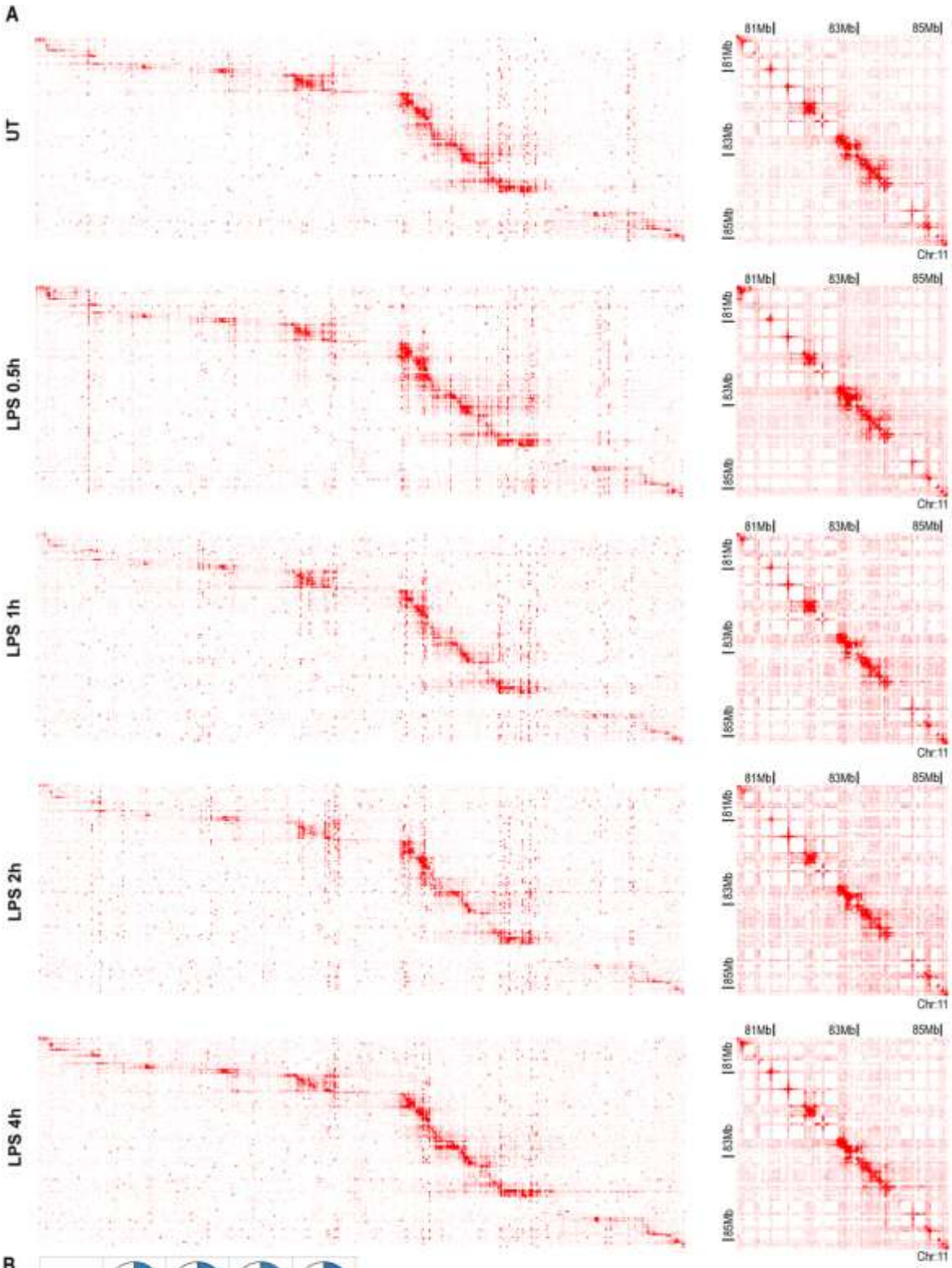
A
UT
LPS 0.5h
LPS 1h
LPS 2h
LPS 4h
81Mb
83Mb
85Mb
Chr:11
B
UT
LPS 30min
LPS 1h
LPS 2h
LPS 4h
0.69
0.7
0.62
0.69
0.62
0.71
0.74
0.69
0.72
0.7

Figure 10 - LPS treatment.
Comparison of 5C data from untreated or LPS-stimulated mouse macrophages reveals general conservation of chromatin contacts during stimulation. (a) Left column shows raw 5C interaction maps with normalized counts (regression on the counts in *trans*). Right column shows 5C matrices with counts binned in 50kbp windows for visualization. Counts are normalized by the number of primers in the bin. (b) Spearman's rank correlation excluding not available values (i.e. the white holes in the heatmaps).

1.3 Local and dynamic chromatin rearrangements at selected loci upon LPS stimulation

As discussed above, the 5C primers pool was designed using an anchored scheme, with all the RV primers designed on the TSS-containing restriction fragments. This allowed us to perform a more detailed analysis investigating the impact of LPS stimulation on the interaction pattern of a specific TSS with the surrounding genomic region. In particular, we initially focused our attention on the TSS of the *Ccl5* gene, for which our design included three different 5C primers. We considered a genomic window of approximately 400Kb around each RV primer and we calculated the number of contacts (raw counts) with the FW primers in that region. We excluded from the analysis a $\pm$ 10 Kb region from each RV primer. This 'blind spot' is motivated by the physical properties of the chromatin fiber, which determine an extremely high number of counts in the region adjacent to the viewpoint. As indicated by chromatin associated RNAs, LPS treatment led to the activation of *Ccl5* with a comparatively slow kinetics. Gene induction started at 1 h post stimulation and reached the maximum peak at about 2h (Figure 11a). A TSS-centered analysis of *cis*-interactions showed a decrease of chromatin interactions downstream the TSS (black arrows) with the intensity of signal loss correlating with the maximum peak of activation. Reciprocally, gene activation correlated with an increase in chromatin interactions upstream the TSS. This region contains several H3K4me1 peaks (Figure 11a) and undergoes inducible acetylation and transcription (De Santa et al. 2010), thus likely containing a cluster of enhancers (Figure 11a and b). The 5C data showed qualitatively similar data among the three RV primers used. Taken together, these results indicate that after gene activation induced by external stimuli local dynamic rearrangements can be

detected in the interaction pattern of a specific TSS with its regulatory elements, albeit the global chromatin interaction landscape of the locus remains unaffected. Other genes in the same region, however, did not show a consistent and reproducible reconfiguration of chromatin contacts in response to LPS stimulation.

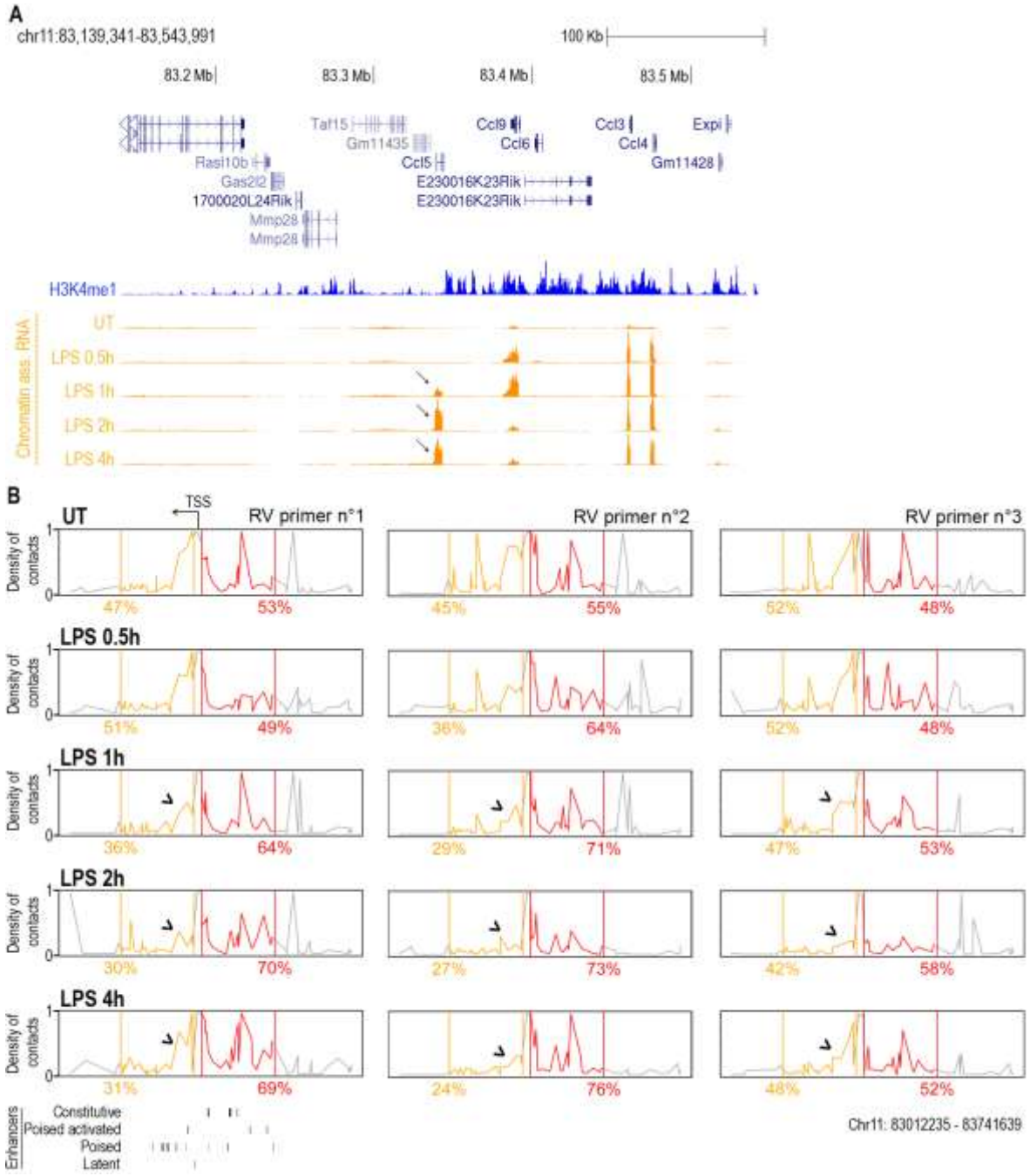

Figure 11 - TSS-centered 5C analysis
(a) A 400 Kb window around the TSS of the Ccl5 gene showing H3K4me1 and chromatin associated RNA in untreated or LPS-stimulated mouse macrophages. (b) 5C analysis targeted on three different RV primers annealing to the TSS-containing restriction fragments of the Ccl5 gene. Counts are normalized using the counts in *trans*. The orange window includes a region whose range is [-200Kb, -10Kb] from each RV primer, while the red one has a range of [+10Kb, + 200Kb] from each RV primer. The black arrow indicates transcription direction.

2. 4C ANALYSIS CHARACTERIZATION OF 3D GENOMIC ORGANIZATION IN MACROPHAGES ACTIVATED BY LPS

To further investigate if and how chromatin looping changes upon perturbation, we exploited a modified 4C technique as a complementary and higher resolution approach, which should theoretically allow mapping 3D interactions with a resolution in the order of a few hundreds base pairs. The high resolution 4C technology (Hi-Res 4C) is a variant of the standard 4C from which it differs because of the digestion of the 3C template with a frequently cutting primary restriction enzyme (Dpn II) recognizing a 4bp site, followed by a ligation to form circles, and eventually inverse PCR to amplify captured primers (Simonis et al., 2009; Simonis et al., 2007). The initial steps of 3C were performed as described previously (Dekker et al., 2002), yielding ligation products between DpnII fragments. Moreover, in the 4C technique, interactions of a single locus with the entire genome can be tested (Simonis et al., 2006a; Zhao et al., 2006). Hi-Res 4C-seq data were compared to nascent RNA-seq data, in order to correlate promoter–enhancer interactions with transcriptional changes. In particular, here we used a 4C-seq multiscale analysis to quantify the intensities of *cis* interactions between the viewpoint positioned near the TSS of selected genes and a 500kb genomic region surrounding it. To maximize the temporal resolution of the analysis, 4C libraries were generated in biological triplicates from mouse macrophages treated with bacterial endotoxin and sampled at 15min, 30min, 1h, 2h, 4h, 8h and 24h. We found good global correlations among replicates (Spearman's rank correlation coefficient in the range of 0.6-0.7 for intra-chromosomal interactions), suggesting that the 4C-seq data obtained were consistent and reproducible.

We first characterized *cis* interactions involving the TSS of the *Ccl5* gene. Figure 12 shows the 4C-seq profile of *Ccl5* together with chromatin associated RNA, H3K4me1 and H3K27ac data in a 500Kb region surrounding the *Ccl5* TSS. As described above, LPS activated Ccl5 starting from 1 h post stimulation and reaching the maximum expression at 2 h. The general interaction profile of the gene appeared symmetric with respect to the

viewpoint, with chromatin contacts both upstream and downstream of the promoter and no clear preference for the upstream H3K4me1-positive enhancer cluster.

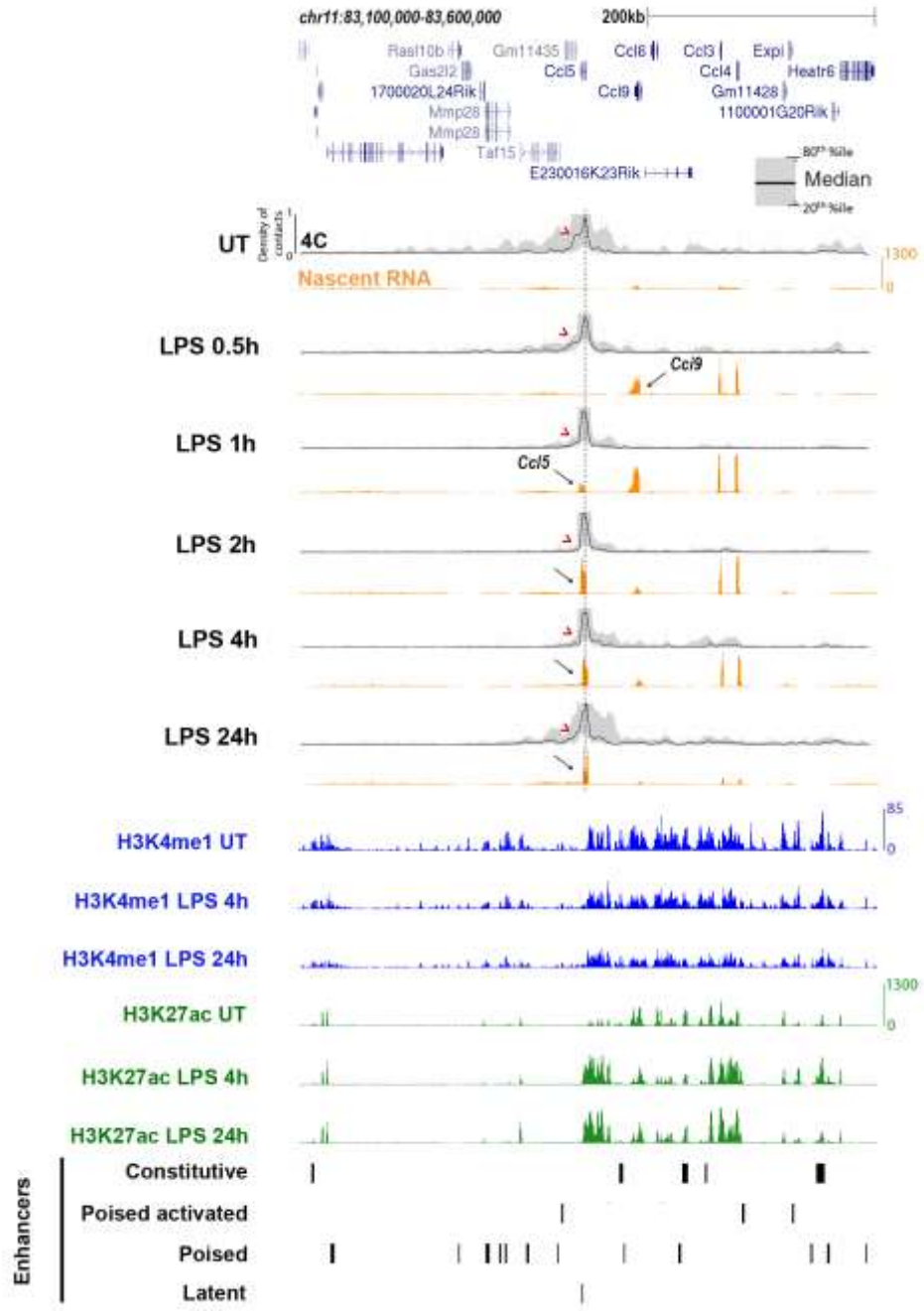

Figure 12 - Representative TSS-genome interactions for Ccl5
4C contact profile for Ccl5 using a 500kb window size. Dashed gray line indicates viewpoint position representing the TSS of the gene. Chromosomal context of the viewpoint is shown with chromatin associated RNA, active enhnancers (H3K27ac in the presence of H3K4me1) and distal regulatory elements (H3K4me1) histone modification data in untreated or LPS induced mouse macrophages.

Ccl5 activation was associated with small but clearly detectable dynamic differences in 3D interactions that were consistent with the 5C data shown above. Specifically, interactions

downstream of the TSS were strongly attenuated in concomitance with the maximum peak of activation. As a consequence, in response to stimulation the TSS of *Ccl5* preferentially interacted with the active and transcribed enhancers (De Santa 2010) located upstream of the TSS (even though the overall interaction frequency with these regions was not obviously increased).

We then evaluated the partial correlations among different time points in a spatial window of +/- 1Mbp around the viewpoint. Considering each side of the viewpoint separately, they were divided into equal bins of 2Kbp. Starting with the 10 bins nearest to the viewpoint, one bin at the time was then added and a partial correlation computed. Correlations were computed with respect to an untreated sample chosen as reference. As expected, the correlations were generally very high considering the region nearby the viewpoint, but decreased more rapidly after stimulation (Figure 13a). This loss of correlation at late time points corroborates the previous observation that a reconfiguration of chromatin interactions within this locus occurs after stimulation.

To obtain a more quantitative assessment of the observed changes, we measured the number of chromatin contacts between the TSS of *Ccl5* and the other genomic fragments, considering a window of $\pm$ 200Kb from the viewpoint (Figure 13b). Gene activation led to a decrease of chromatin interactions downstream the TSS together with a concomitant relative increase of contacts in the upstream enhancer cluster.

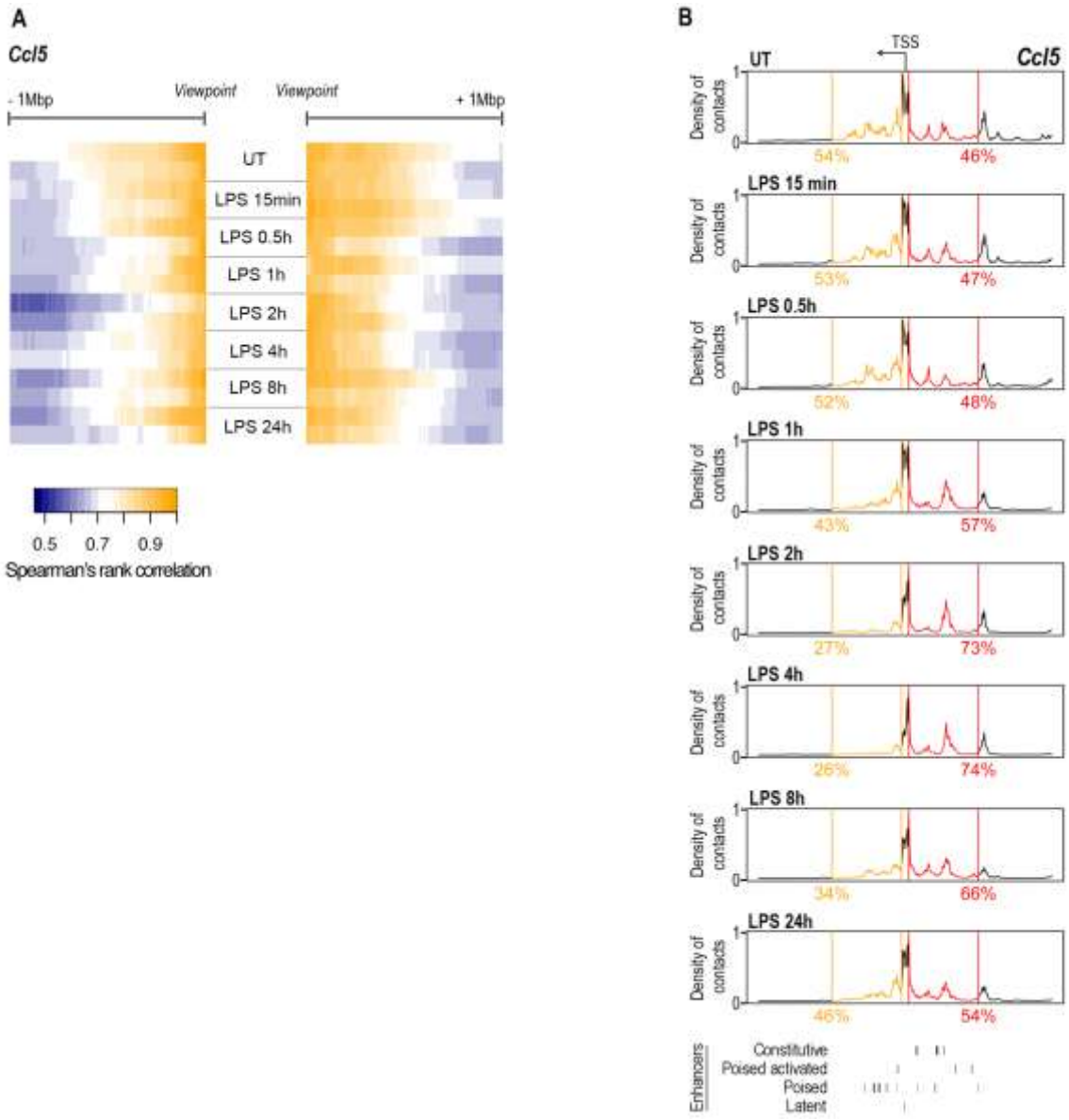

Figure 13 - 4C analysis of chromatin contacts during gene activation for Ccl5
(a) Partial correlations among time points. Spearman's rank correlation coefficient is calculated for a spatial window of +/- 1Mbp starting from the viewpoint and subdivided into 2Kbp-bins. (b) Percentage of chromatin contacts (smooted counts) around the TSS in a genomic window of approximately 400Kb. The orange window includes a region whose range is [-200Kb, -10Kb] from each anchor, while the red one has a range of [+10Kb, + 200Kb] from each anchor. A $\pm$ 10 Kb region from the TSS was considered as a blind spot and it was excluded from the analysis. The black arrow indicates transcription direction.

We than analyzed a panel of inducible genes expressed in myeloid cells and showing different pattern of induction after LPS stimulation (Bhatt et al., 2012).

The *Cxcl10* gene is another chemokine gene family member that responds to LPS with a slower kinetics than *Ccl5*, starting from 1 hour after the treatment (chromatin associated RNAs are shown in Figure 14). During stimulation, chromatin contacts involving the TSS of this gene showed a behavior similar to the one observed at *Ccl5*. *Cxcl10* transcription was not associated with drastic changes in the general 4C profile, confirming the idea that

the majority of chromatin contacts was largely preformed. However, after transcriptional activation, the pre-existing chromatin contacts downstream of the TSS (red arrows) were reduced, as indicated by the shape of the 4C profile that became sharper and more symmetric with respect to the viewpoint.

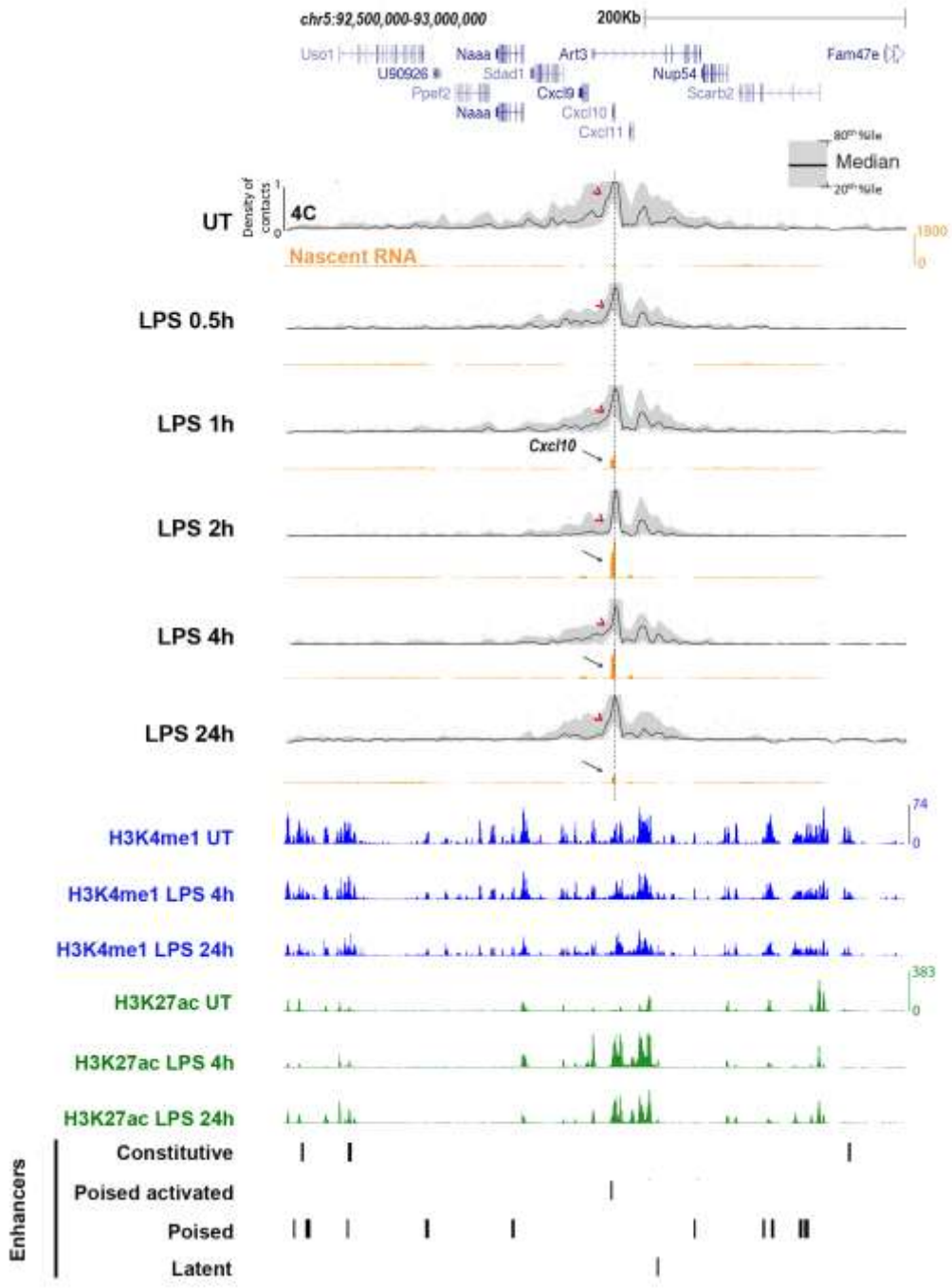

Figure 14 - Representative TSS-genome interactions for Cxcl10
4C contact profile for Cxcl10 using a 500kb window size. Dashed gray line indicates viewpoint position representing the TSS of the gene. Chromosomal context of the viewpoint is shown with chromatin associated RNA, active enhnancers (H3K27ac in the presence of H3K4me1) and distal regulatory elements (H3K4me1) histone modification data in untreated or LPS induced mouse macrophages.

These dynamic changes appeared more evident 2h after the LPS stimulation, when the maximum peak of activation was reached, and were supported by analysis of partial correlations, which showed a loss of correlation at time points of maximal transcription activity (Figure 15a). Furthermore, the quantification of chromatin contacts around the TSS of *Cxcl10* during gene activation, revealed a relative increase in the interactions in a 20kb region surrounding the viewpoint and including LPS-activated enhancers (as indicated by the increase of H3K27ac in H3K4me1 positive regions, figures 14 and 15b).

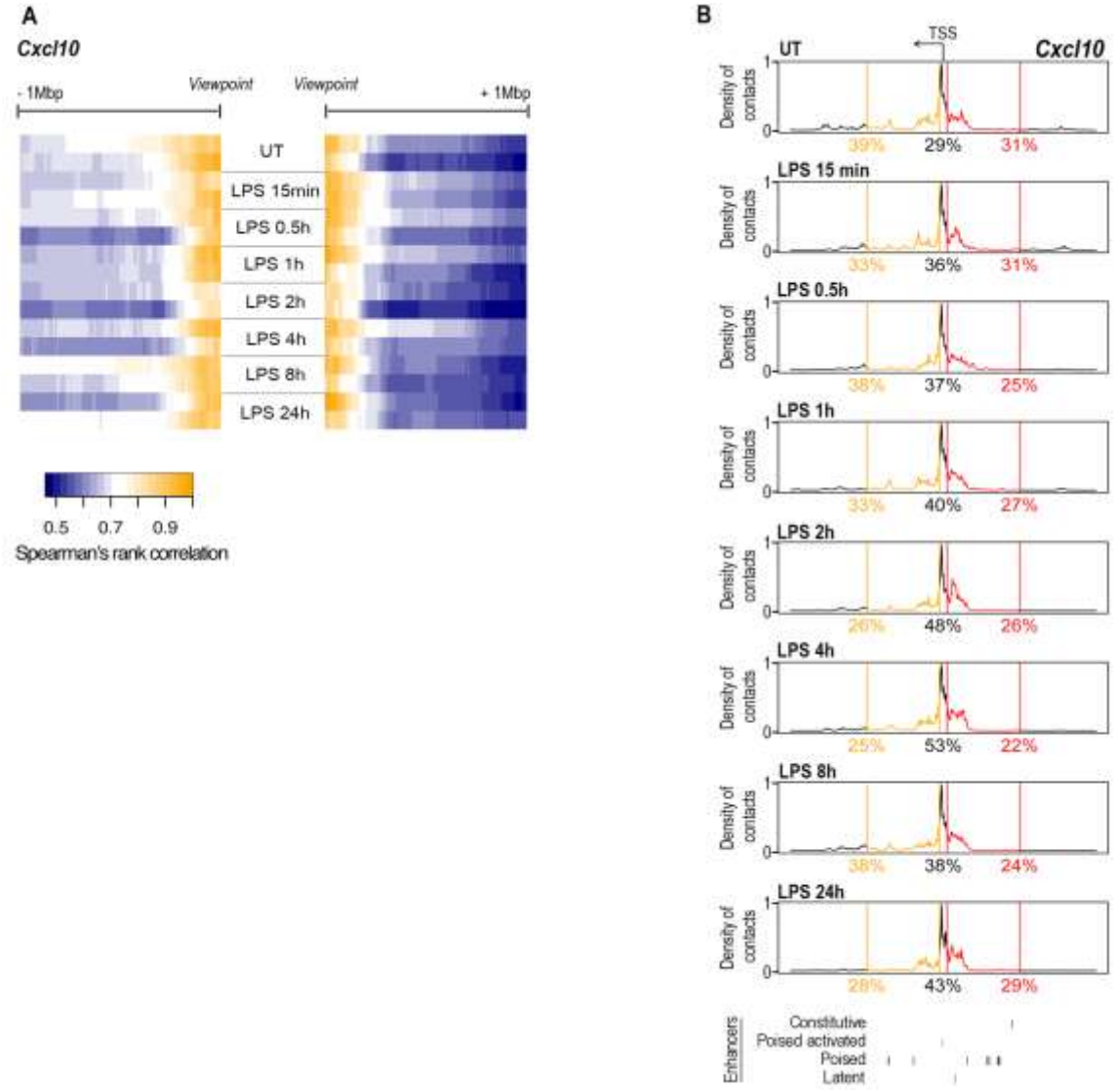

Figure 15 - 4C analysis of chromatin contacts during gene activation for Cxcl10
(a) Partial correlations among time points. Spearman's rank correlation coefficient is calculated for a spatial window of +/- 1Mbp starting from the viewpoint and subdivided into 2Kbp-bins. (b) Percentage of chromatin contacts (smooted counts) around the TSS in a genomic window of 400Kb. The orange window includes a region whose range is [-200Kb, -10Kb] from each anchor, while the red one has a range of [+10Kb, + 200Kb] from each anchor. The central window measures $\pm$ 10 Kb around the TSS. The black arrow indicates transcription direction.

Overall, these observations suggest that dynamic changes in chromatin contacts occur concurrently with the activation of selected LPS-inducible genes.

Data obtained for *Ccl5* and *Cxcl10* suggest that the control of gene expression in response to external stimuli is associated with a partial rearrangement of a set of predetermined chromatin contacts occurring in genomic regions rich of regulatory elements.

However, the same analysis performed on other inducible genes did not detect any dynamic change in response to LPS, thus indicating that chromatin interactions are not only preformed, but also stable during gene activation (Figures S1 and S2). For example, *Jun* encodes a TF whose expression in macrophages is induced by LPS. As indicated by chromatin associated RNA data in Figure 14, *Jun* reached its maximum peak of activation 0.5h post stimulation and was then almost immediately completely turned off. The general interaction profile of this gene was asymmetric with respect to the anchor (Figure 16), with a strong bias for the H3K4me1-decorated gene desert at the 5' of the gene. These regions did not show any detectable increase in H3K27ac in response to stimulation, but they were constitutively associated with acetylated nucleosomes. LPS activation of *Jun* did not cause major changes in the general 4C interaction profile, which appears to be stable and preformed, as confirmed by the quantification of chromatin contacts around the TSS of the gene (Figure 17b). Analysis of partial correlations also indicated a strong conservation of chromatin contacts among different time points for the enhancer-rich region upstream of the gene, suggesting that no re-configuration of chromatin interactions occurs after stimulation (Figure 17a).

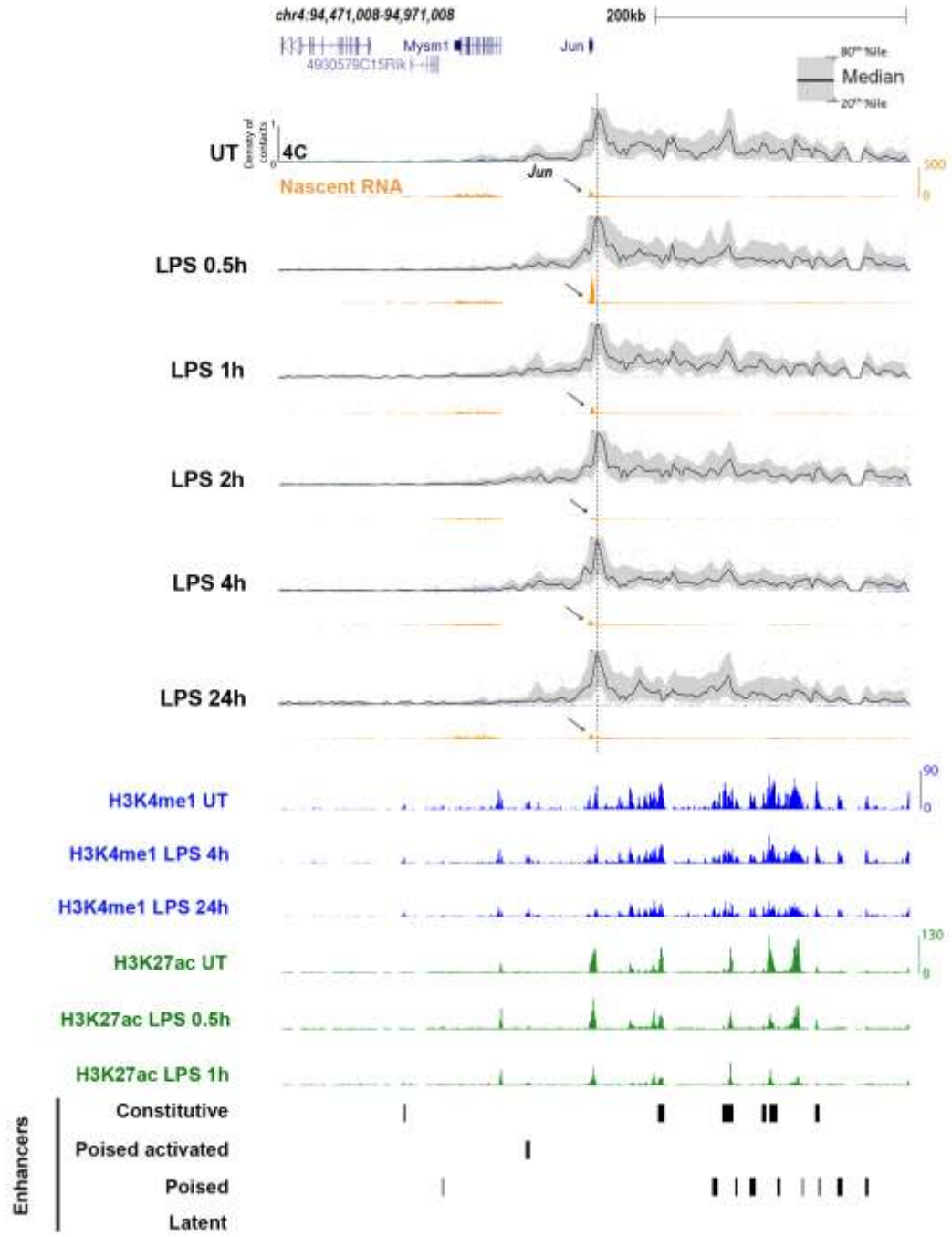

Figure 16 - Representative TSS-genome interactions for Jun
4C contact profile for Jun using a 500kb window size. Dashed gray line indicates viewpoint position representing the TSS of the gene. Chromosomal context of the viewpoint is shown with chromatin associated RNA, active enhnancers (H3K27ac in the presence of H3K4me1) and distal regulatory elements (H3K4me1) histone modification data in untreated or LPS induced mouse macrophages.

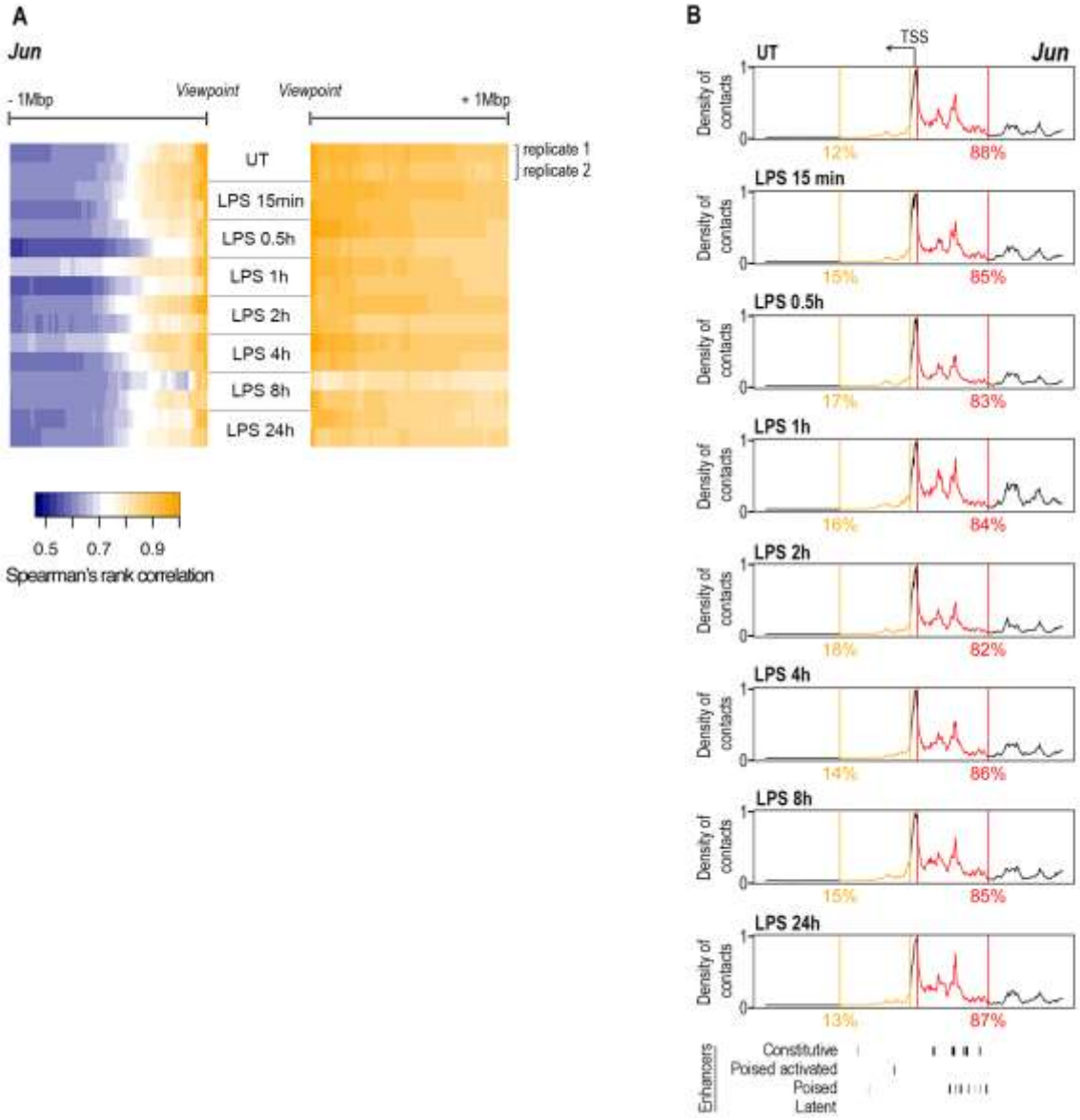

Figure 17 - 4C analysis of chromatin contacts during gene activation for Jun
(a) Partial correlations among time points. Spearman's rank correlation coefficient is calculated for a spatial window of +/- 1Mbp starting from the viewpoint and subdivided into 2Kbp-bins. (b) Percentage of chromatin contacts (smooted counts) around the TSS in a genomic window of 400Kb. The orange window includes a region whose range is [-200Kb, -10Kb] from each anchor, while the red one has a range of [+10Kb, + 200Kb] from each anchor. A $\pm$ 10 Kb region from the TSS was considered as a blind spot and it was excluded from the analysis. The black arrow indicates transcription direction.

3. CELL-TYPE SPECIFICITY OF CHROMATIN CONTACTS AT INFLAMMATORY GENES

We next used 4C-seq to compare the interaction profiles obtained in differentiated BMDM and in mESC.

Figure 18 shows four representative 4C data sets referring to two inflammatory genes preferentially induced in myeloid cells (*Ccl5* and *Nos2*) and two genes broadly expressed

in multiple cell types and/or further induced in response to a large number of stimuli, *Jun* and *Nfkbia*. The 4C-seq profiles showed an overall increase in chromatin interactions in BMDM compared to ESCs for all the considered genes.

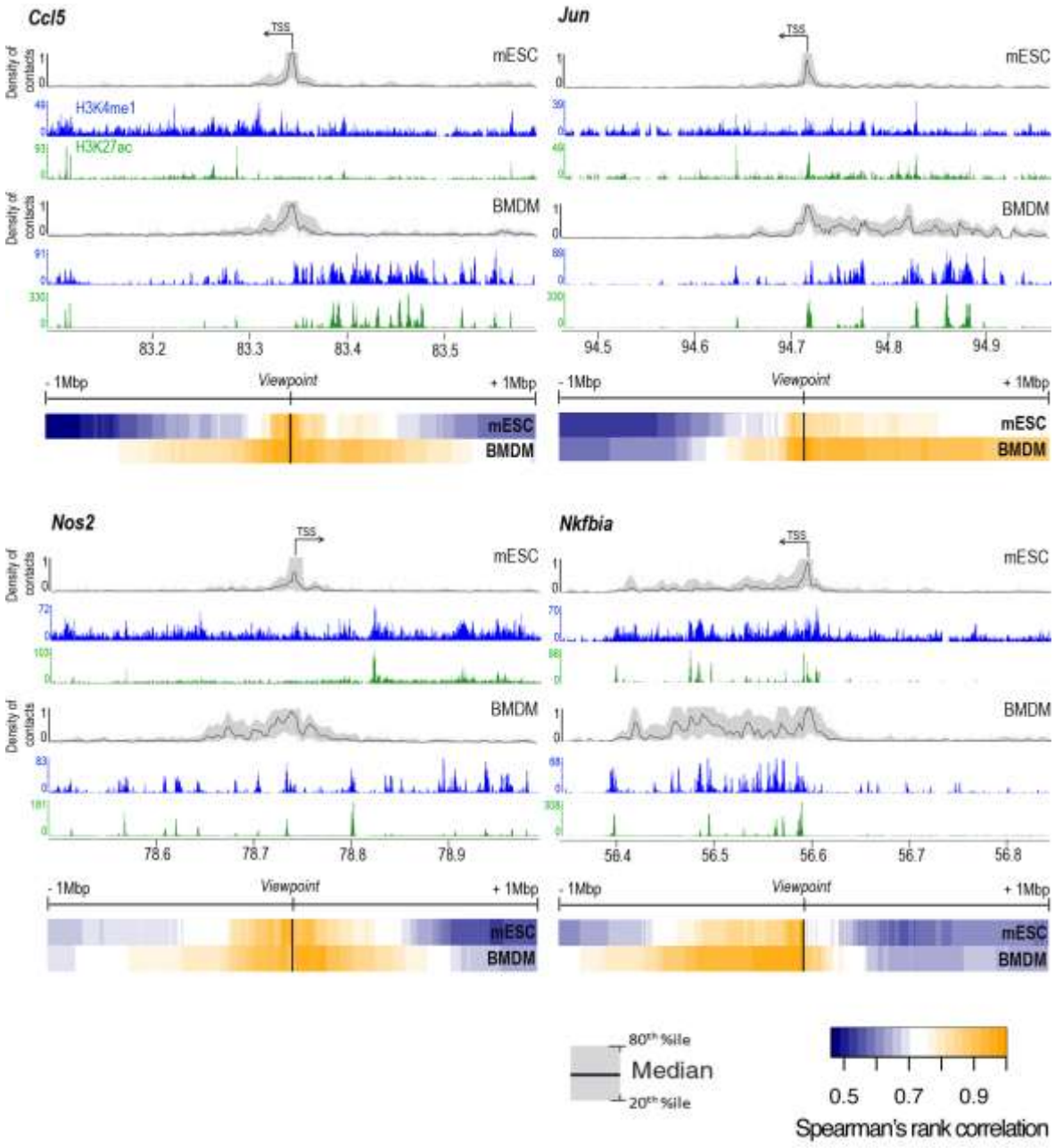

Figure 18 - Cell-type specificity of chromatin contacts
4C contact profile for Ccl5, Jun, Nos2 and Nfkbia using a 500kb window. Chromosomal context of the viewpoint is shown with putative active enhancers (H3K27ac - green) and distal regulatory elements (H3K4me1 - blue) histone modification data in BMDM and mESC. Partial correlation between different cell types is shown above each panel. The black arrow indicates transcription direction.

Different interactions correlated with differences in the distribution of H3K4me1 and H3K27Ac. For *Jun* and *Nfkbia*, some of the H3K4me1 and H3K27Ac peaks were conserved between the two cell types, while others appeared to be specific for BMDM or mESC, suggesting the existence of both cell type-specific and shared enhancers. The enrichment of H3K27Ac in the *Ccl5* and *Nos2* regulatory regions was detected exclusively in BMDMs, indicating that the presence of a specific set of active enhancers in a given cell type correlates with, and may determine a specific set of chromatin contacts relevant for the selective expression of cell type-specific genes.

We then evaluated the partial correlations among 4C-seq profiles of different cell types as described above. Near the viewpoint, the correlations were generally very high but they decreased differently in BMDM and mESC. This loss of correlation indicates that different cell-types are characterized by different configurations of basal *cis*-chromatin interactions.

Taken together data suggest that chromatin contacts between promoters and regulatory sequences vary among different cell-types for both broadly expressed genes and cell type-specific genes.

4. IMPACT OF REDUCED LEVELS OF THE MYELOID TF PU.1 ON 3D CHROMATIN ORGANIZATION

We next tested if the cell type-specific macrophage nuclear organization is supervised by the master regulator PU.1. To this aim we performed 4C-seq in terminally differentiated macrophages in which PU.1 was depleted by RNA interference. We used lentiviral vectors expressing either Pu.1-specific or control shRNAs (pLKO.1shPu.1 and pLKO.1shLuc, respectively). Acute depletion resulted in a ca. 75% reduction in PU.1 protein levels (Figure 19a and b) and was associated with a reduced occupancy of bound genomic regions, as detected by ChIP-qPCR (Figure 19c).

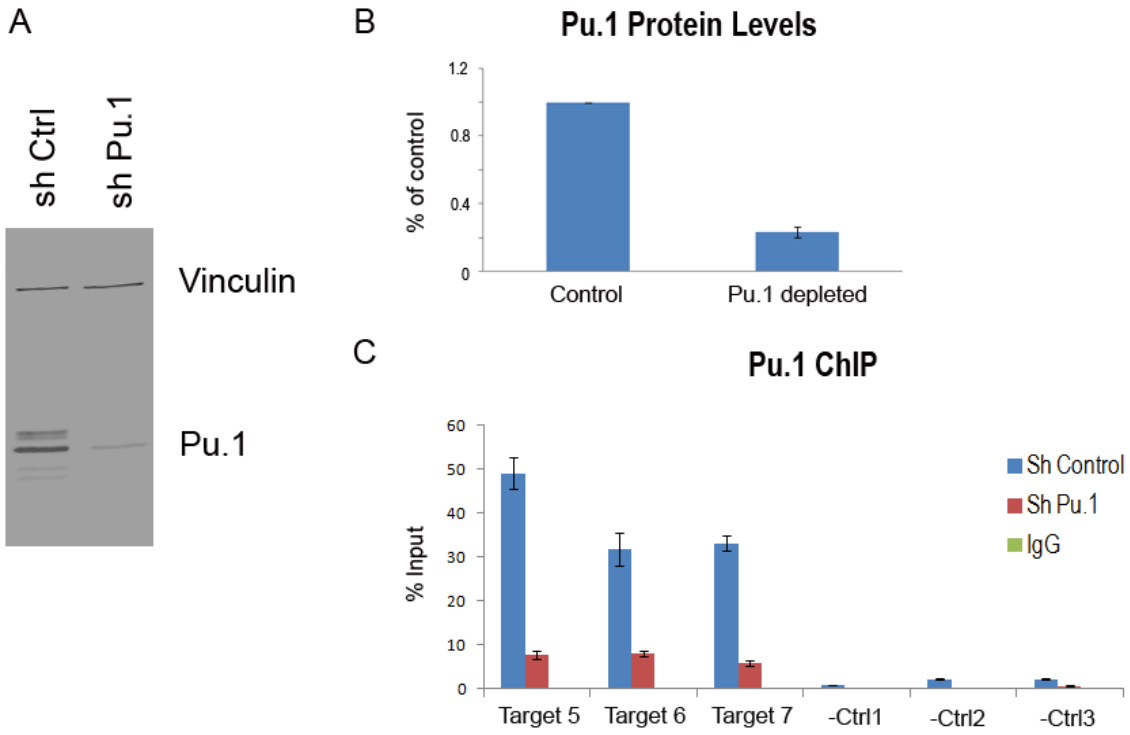

Figure 19 - Lentiviral Pu.1 shRNA efficiency on PU.1 protein levels and on PU.1 genomic occupancy
BMDM were infected with lentiviruses expressing a scrambled sequence (shLuciferase) or an shRNA silencing Pu.1 (shPu.1) and selected with puromycin for two days. Whole cell extracts were prepared from knock down or mock infected cells and analyzed by ChIP-qPCR and western blot analysis for PU.1. (a) Levels of PU.1 protein following shRNA knock down by western blotting. (b) Western blot quantification. Results refer to the enrichment relative to the control signal. Error bars represent the Standard Deviation of the mean among three biological replicates. (c) ChIP for the PU.1 protein followed by qPCR for PU.1 targets. Results are represented as enrichment relative to the input. qPCR were done in triplicate and error bar denotes the standard deviation of the mean.

Figure 20 shows the 4C interactions profile of *Jun*, *Nfkbia*, *Ccl5* and *Nos2* in control and PU.1-depleted macrophages. The comparison of 4C interaction signals from control and PU.1 depleted cells revealed a general conservation of chromatin contacts for all the considered viewpoints. These results suggest either that low residual levels of PU.1 are sufficient to maintain the macrophage-specific genome configuration, or that PU.1 plays a marginal role in assisting reciprocal looping between regulatory elements. It is also possible that PU.1 acts redundantly with additional TFs, such as the highly related Ets family protein FLI1, which has a similar genomic distribution as PU.1 (unpublished data).

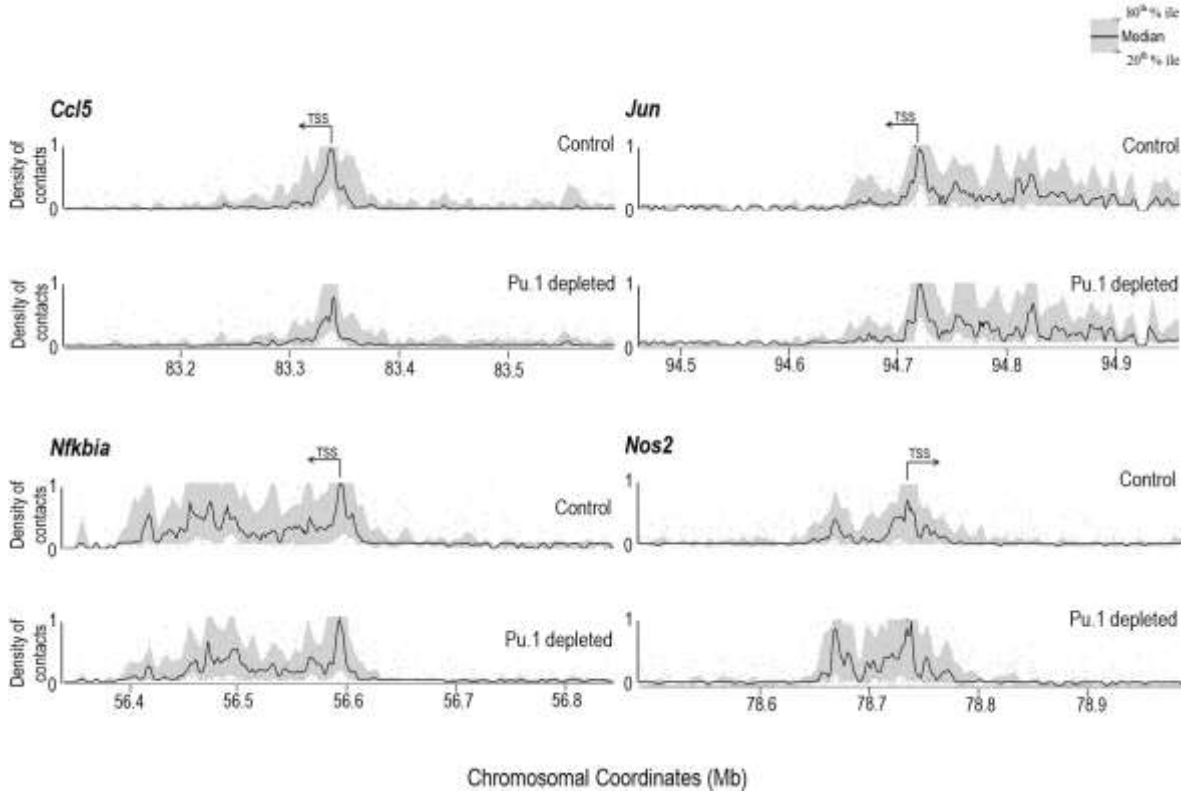

Figure 20 - PU.1 knock down mouse macrophages show general conservation in long range interactions

Chromosomal map of 4C signal for Ccl5, Jun, Nfkbia and Nos2 genes in PU.1 knock down or mock transfected mouse macrophages. Representative 4C data are showed in a 500kb window size in the main trend subpanel. The black arrow indicates transcription direction.

To assess the impact of PU.1 on chromosome configuration, we performed 4C-seq in NIH-3T3 fibroblasts transduced with a PU.1 expression vector. To this aim we used a retroviral vector expressing Pu.1 (pMSCV-Pu.1) or an empty vector (pMSCV-EV) as control. The vector used efficiently expressed Pu.1 mRNA and protein in NIH-3T3 cells (Figure 21a and c) and the occupancy of PU.1 at its bound genomic regions was detected by ChIP-qPCR (Figure 21b). Furthermore, the retroviral expression of Pu.1 induced the activation of some macrophage-specific genes such as Cd68, Emr1 and Lyz2 (Figure 21c).

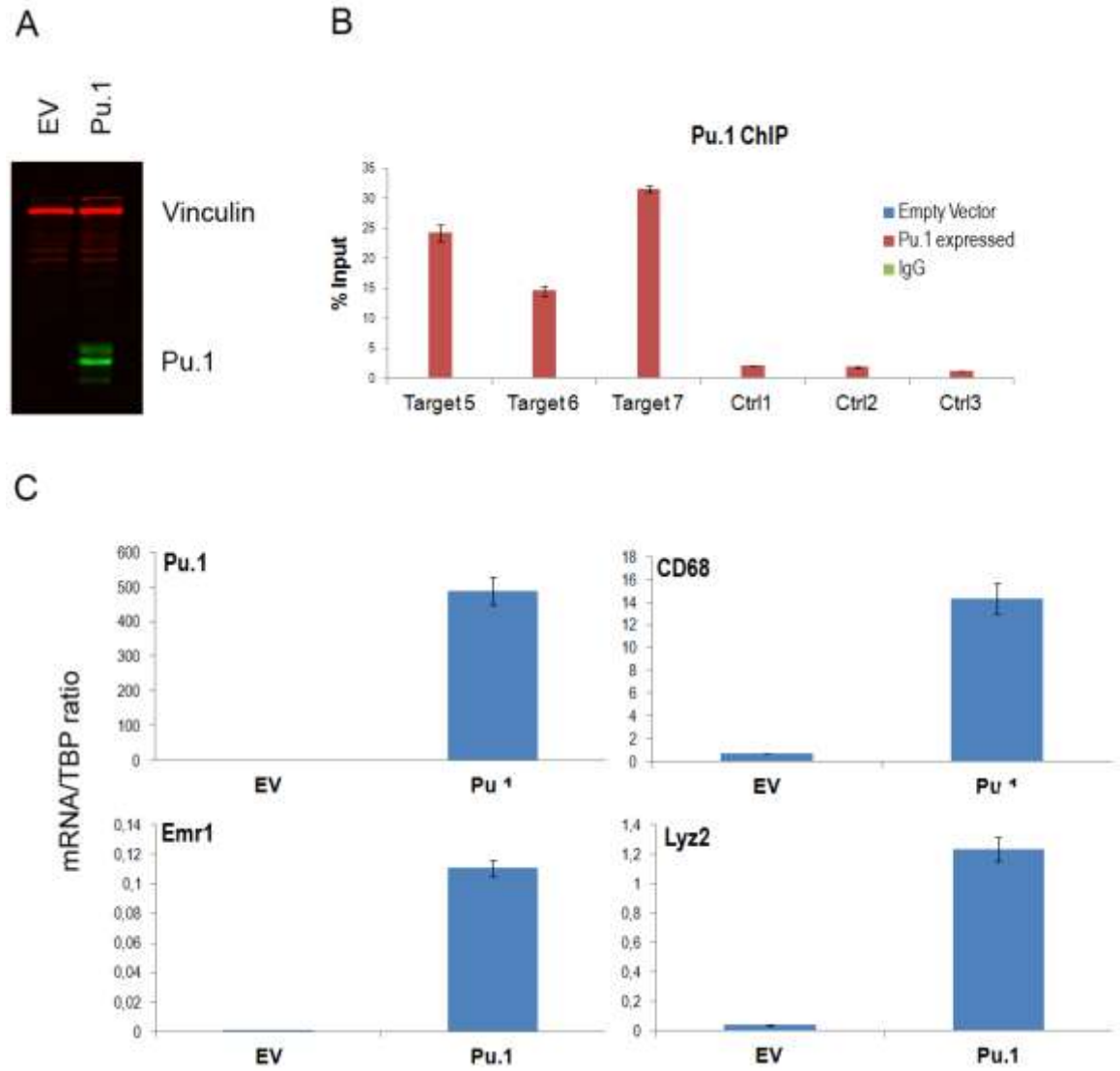

Figure 21 - Pu.1 expression in NIH-3T3 cells
NIH-3T3 cells were infected with retroviruses expressing empty vector (EV) or Pu.1 (Pu.1) and selected with puromycin for two days. RNA and whole cell extracts were prepared from infected cells and analyzed by quantitative RT-PCR, ChIP and western blot analysis for Pu.1. qPCR were done in triplicate and error bar denotes the standard error of the mean. (a) Levels of PU.1 protein following Pu.1 expression by western blotting. (b) ChIP for the PU.1 protein followed by qPCR for PU.1 targets. Results are represented as enrichment relative to the input. (d) qPCR analysis for macrophage markers.

Figure 22 shows the 4C-seq profile of *Jun*, *Nfkbia*, *Ccl5* and *Nos2* genes in normal and Pu.1-expressing NIH-3T3 cells. The figure also shows that the genomic distribution (analyzed by ChIP-seq) of Pu.1 in infected NIH-3T3 and BMDM is similar. The comparison of 4C-seq interaction signals from the two cell types revealed a general conservation of chromatin contacts for all the considered viewpoints (in fact, compared to ESCs, NIH-3T3 were more similar to BMDMs). Furthermore, the comparison between the

4C interaction maps from Pu.1-expressing NIH-3T3 cells and BMDMs showed an overall similarity in the global chromatin interactions landscape. These data confirm that the general chromosome topology in BMDM is not supervised by the master regulator PU.1. Whether this TF contributes to assist reciprocal looping between regulatory elements remains to be assessed.

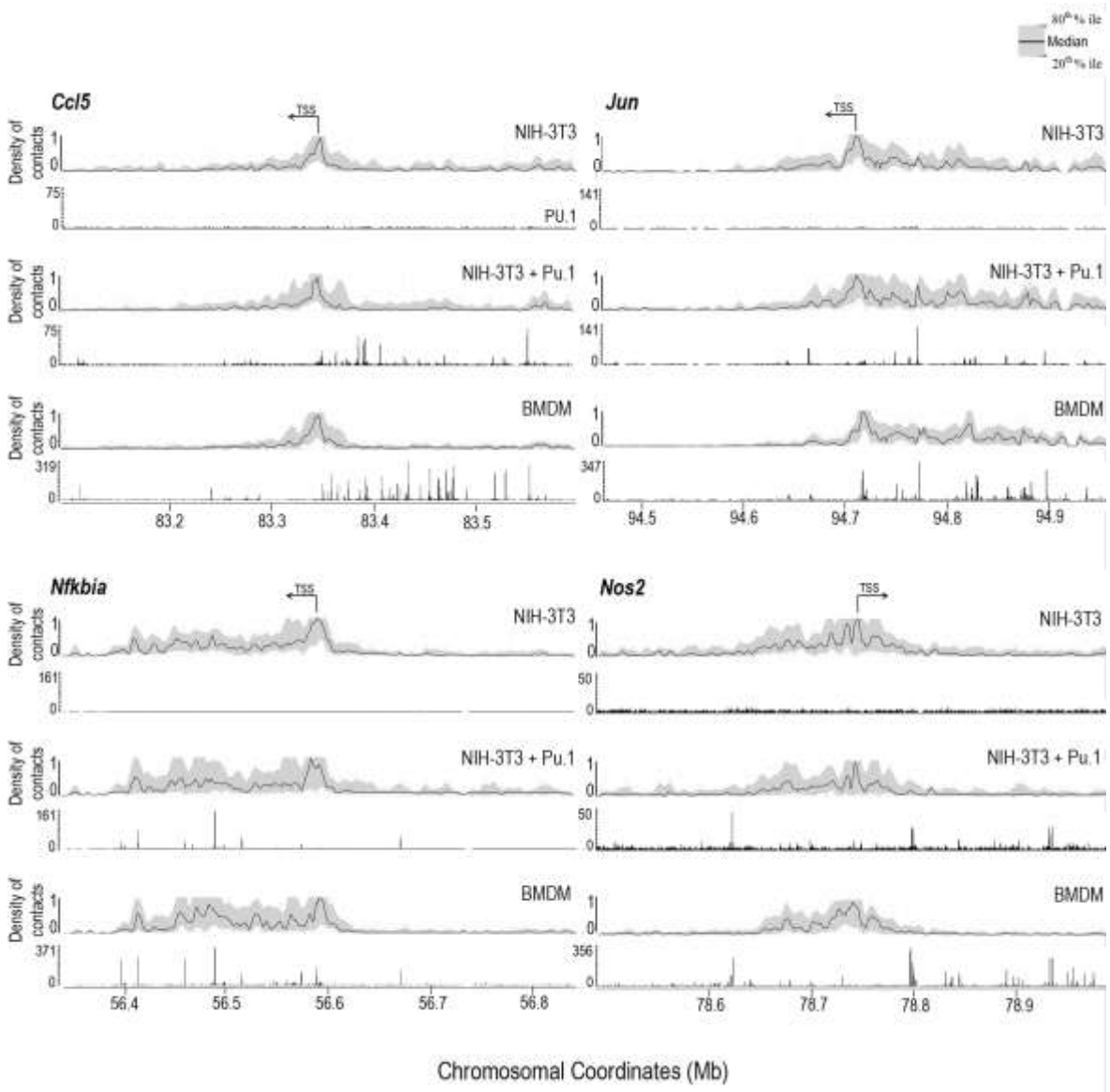

Figure 22 - NIH-3T3 expressing Pu.1 show general conservation in long range interactions
Chromosomal map of 4C signal for Ccl5, Jun, Nfkbia and Nos2 genes in mouse macrophages and normal or Pu.1 expressing NIH-3T3 cells. Representative 4C data are showed in a 500kb window size in the main trend subpanel. Chromosomal context of the viewpoint is shown with the deposition of PU.1. The black arrow indicates transcription direction.

DISCUSSION

Here we used technologies based on the chromosomes conformation capture (3C) to investigate the global 3D organization of the macrophage epigenome, analyzing the specific role of the lineage specifying TF PU.1, in both basal conditions and after perturbation.

We first exploited the 5C technique to characterize one selected locus on chromosome 11 containing several LPS-inducible genes as well as genes unaffected by stimulation. This allowed us to investigate the relationship between dynamic changes in gene expression and interactions between enhancers and cognate TSSs. Our 5C maps showed an inverse relationship between genomic distance and interaction frequency within the region of interest, with an higher number of chromatin contacts for pairs of fragments closer in the linear genome and located within the same TAD. These results are consistent with the known general features of spatial chromatin organization, specifically the prevalence of local chromatin interactions.

We next investigated how TSS–distal fragment interactions relate to gene expression dynamics during the LPS response. At the simplest level of interpretation, our data revealed that chromatin contacts are to a large extent unaffected by stimulation, suggesting that the general interactions landscape is a predetermined feature of this specific cell type and that most gene expression changes occur within such landscape. In other words, most spatial contacts required for gene expression may be present already before stimulation, a result consistent with a recent analysis (Jin et al., 2013). Indirectly, this result implies that the locus must be constitutively in a configuration that is compact enough to enable spatial proximity among regulatory elements relevant for gene regulation. However, a TSS-centered analysis of the 5C map for the Ccl5 gene revealed that gene activation induced by external stimuli leads to local dynamic rearrangements in the interaction pattern of a specific TSS with its regulatory elements, consistent with previous studies showing that enhancer-driven gene activation is characterized by an alteration of chromatin interactions

(Ong and Corces, 2011; Schoenfelder et al., 2010a; Smallwood and Ren, 2013). This suggests that even if the majority of contacts between genes and distal *cis*-regulatory elements are preformed, the control of gene activation may require a fine reorganization of local chromatin contacts which perhaps may differ among different stimuli and cell types. In this regard, it would be of interest to evaluate if the activation of Ccl5 induced by external stimuli different from LPS, such as Poly (I:C), would lead to the same kind of chromatin rearrangements observed in this study. Furthermore, it would be interesting to understand whether the induction of Ccl5 in different cell types, for example in T lymphocytes by interferon gamma (IFNγ), involves similar chromatin contacts around the TSS of this gene. These studies will improve our knowledge on the common basal genomic organization among different cell types and their similar or different rearrangements during gene activation.

To obtain higher resolution data on the interactions between the TSS of individual inducible genes and the rest of the genome, we used multiplexed high-resolution 4C, which has also the advantage over 5C of a lower technical complexity (which may imply a lower tendency to generate experimental artifacts). 4C libraries were prepared from mouse macrophages subjected to external stimuli over a complete time course (8 different time points from 15min to 24h), thus obtaining an high temporal resolution which allowed us to evaluate if possible rearrangements in chromatin contacts during the LPS response are caused by transcriptional activation or precede transcriptional changes. In this study we considered genes with different pattern of induction after LPS treatment (Bhatt et al., 2012).

Consistent with our 5C data and other 4C studies on several loci induced by p53, FOXO3 and glucocorticoid receptor (Eijkelenboom et al., 2013; Hakim et al., 2009; Melo et al., 2013), we observed that LPS gene activation doesn't drastically affect the general 4C interaction profiles, thus confirming that in general, enhancer–promoter interactions are already formed in untreated cells and that these pre-existing chromatin contacts are not

significantly altered after target gene induction. These findings may indicate that pre-existing DNA chromatin interactions between enhancers and promoters could function as a signal for stimulus-dependent transcription factors to indicate the presence of specific target genes even before their activation.

However, we observed that chromatin contacts within a short genomic distance from the TSS of Ccl5 and Cxcl10 undergo relatively few changes during transient transcriptional activation, especially in genomic regions rich of inducible regulatory elements. In particular, we observed a strongest reconfiguration of chromatin interactions at the maximum peak of gene activation, indicating that changes in chromatin contacts during the LPS response are caused by transcriptional activation and don't precede transcriptional changes. We also noticed that the long-range interaction landscape of active genes is asymmetric respect to the TSS and unfolds on either side of the gene promoter, usually within H3K4me1-rich regions which gain H3K27 acetylation, suggesting that enhancers activated by the stimulus probably play a key role in the reconfiguration of chromatin contacts. These data suggest that enhancer-promoter chromatin contacts are only partially predetermined and require a fine rearrangement to accurately control gene expression during inducible responses.

It's interesting to note that the LPS treatment strongly activates Ccl5 and Cxcl10, which become highly expressed in the majority of treated macrophages, while other genes considered in this study, such as Il12a, are less activated and expressed only in a small fraction of the cells because of the heterogeneity of macrophage populations. Given that 4C data are obtained from an average of the whole cell population, it is possible that our analysis failed to detect dynamic changes in chromatin interactions occurring nearby the anchor of genes not equally induced in all the cells.

On the other hand, the 4C profiles of other inducible genes analyzed in this work, such as Jun and Nfkbia, appeared to be conserved among different time points despite being highly transcribed. This might be explained by the fact that these genes are broadly expressed in

different cell types and for this reason they might be characterized by stable chromatin contacts which tend to remain unaffected after stimulation.

Taken together, these results suggest that after gene activation by external stimuli, looping interactions between enhancers and other regulatory elements remain largely unchanged. However, more complex approaches to data analysis may unveil still undetected regulatory patterns. Specifically, we noticed that cumulative tag distribution in 4C experiments tends to change over time more than individual peaks, possibly indicating compaction/decompaction events taking place over the activated locus.

At the submegabase scale, the molecular mechanism determining chromatin folding is still largely unknown. Here, we expand our understanding of macrophage chromosome organization at the submegabase scale focusing on specific looping interactions between TSS and other regulatory sequences. The comparison of 4C profiles obtained from BMDM and mESCs for myeloid specific genes (e.g. Ccl5) and more ubiquitously expressed genes (e.g. Jun) revealed a strong cell type specificity of chromatin contacts which correlate with a different distribution of potentially active enhancer. In particular, we found a general gain of interaction in BMDM, indicating that gene regulation in differentiated cells might engage more stable and/or less random interactions. Collectively, these data suggest that chromatin contacts at promoters are mostly cell-type-specific both for broadly expressed genes and cell type-specific genes and are probably determined by specific set of active enhancers. Furthermore, our data also indicate that chromatin structure may have a possible causative role in controlling transcriptional patterns, eventually determining cell identity.

Cell-type-specific expression programs are coordinated by multiple regulatory signals integrated at promoter regions, which constitute important platforms where TFs can bind. In macrophages the transcription factor PU.1 acts as master regulator, binding the vast majority of macrophage enhancers and transcription start sites (TSS), and regulating the

deposition of functional enhancer-specific chromatin marks, such as H3K4me1 (Ghisletti et al., 2010; Heinz et al., 2010).

In order to determine whether macrophage-specific nuclear organization is supervised by this lineage specifying TF, we performed 4C in terminally differentiated macrophages in which PU.1 was depleted by RNA interference lentiviruses technology. 4C results showed small dynamic differences between Pu.1 depleted or not-depleted cells with largely unaffected overall chromosome topology. These data suggest that PU.1 may only partially affect the looping between specific regulatory elements or that full knockout of the protein is necessary to affect chromosome conformation in mouse macrophages.

To assess this last hypothesis, we performed 4C in NIH-3T3 fibroblasts in which the overexpression of Pu.1 by an expression vector led to their transdifferentiation into macrophage-like cells. The comparison between the 4C interaction profiles revealed a general conservation of chromatin interactions, confirming the idea that PU.1 doesn't play a key role in shaping the 3D chromatin organization in macrophages. However, it is possible that PU.1 acts redundantly with additional TFs (such as the highly related Ets family protein Fli1). In this regard, it would be interesting to evaluate the impact of a PU.1/additional TFs double knock-down on the macrophage genome.

Taken together, these data indicate that the long-range interactions in macrophages are not dependent on PU.1, suggesting that this TF probably acts on pre-formed chromatin conformations, which may allow the rapid induction of its target genes in response to different stimuli.

At the present moment the role of the macrophage master regulator PU.1 on inducible looping interactions is still under evaluation and further experiments are needed to better elucidate to what extent this TF can affect the 3D chromatin conformation in macrophages.

SUPPLEMENTARY FIGURES

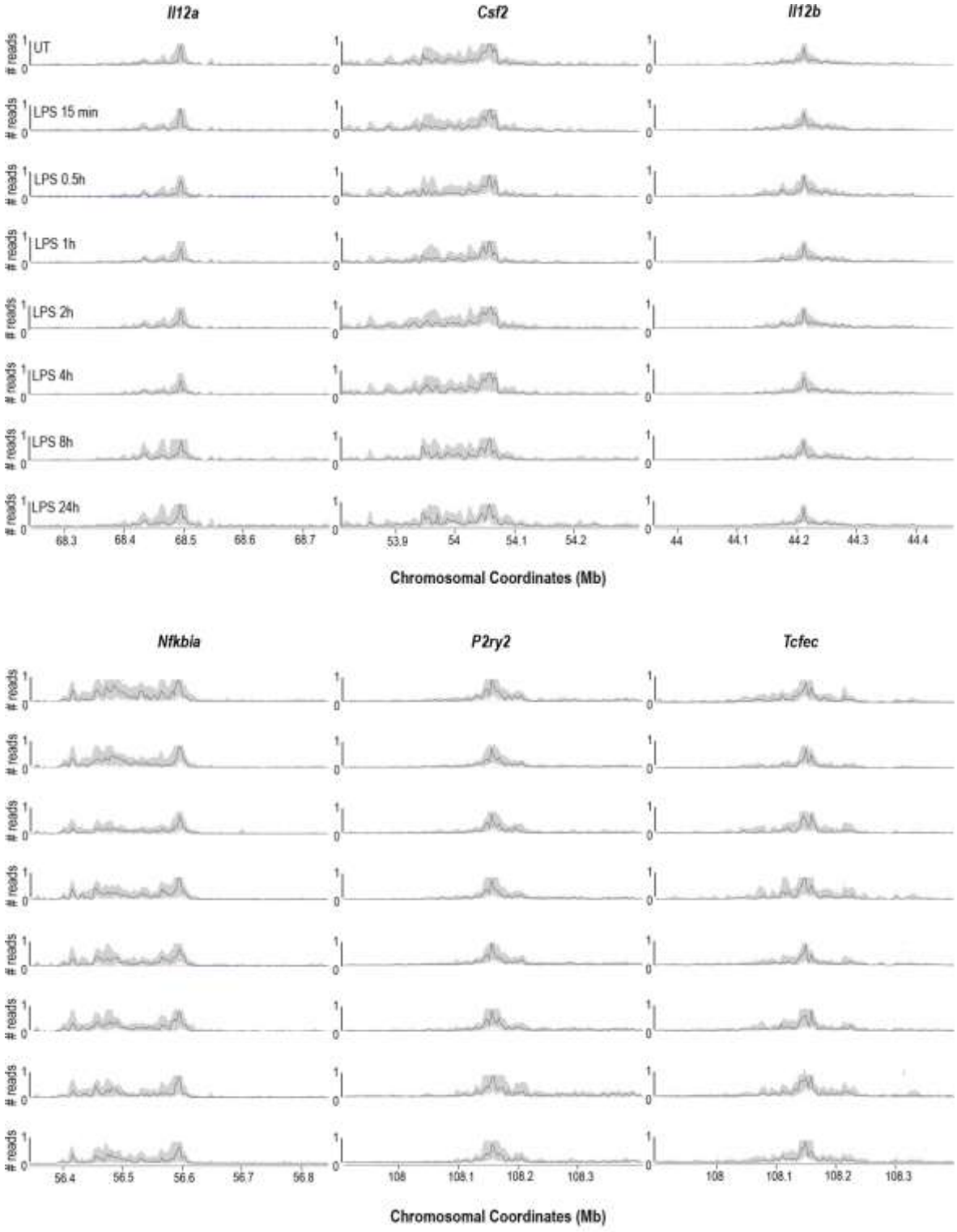

Figure S 1 - Representative TSS-genome interactions for LPS inducible genes
4C contact are shown using a 500kb window size.

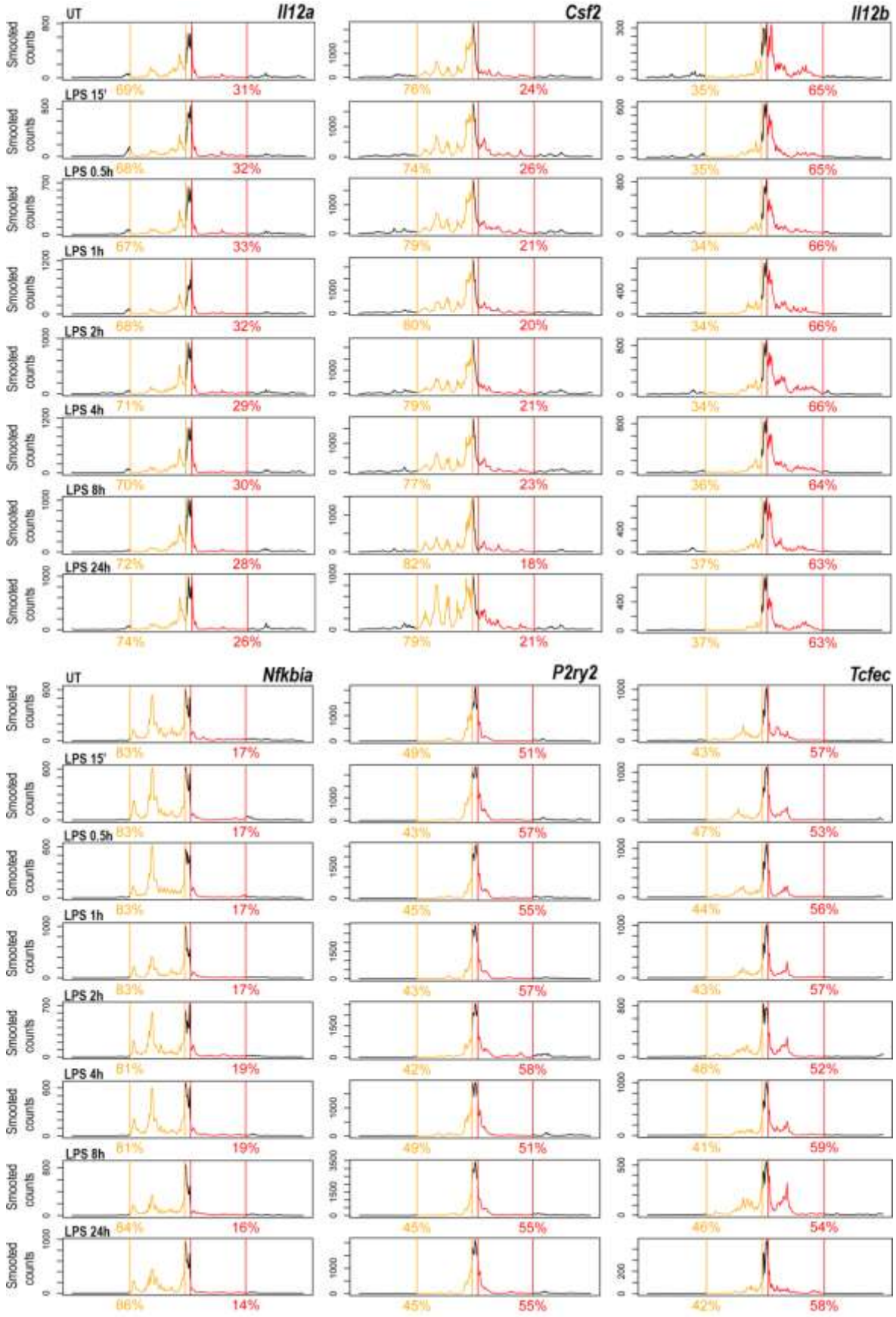

Figure S 2 - 4C analysis of chromatin contacts during gene activation

Percentage of chromatin contacts (smooted counts) around the TSS in a genomic window of approximately 400Kb. The orange window includes a region whose range is [-200Kb, -10Kb] from each anchor, while the red one has a range of [+10Kb, + 200Kb] from each anchor. A $\pm$ 10 Kb region from the TSS was considered as a blind spot and it was excluded from the analysis.

REFERENCES

Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. Science *340*, 1234167.

Apostolou, E., Ferrari, F., Walsh, R.M., Bar-Nur, O., Stadtfeld, M., Cheloufi, S., Stuart, H.T., Polo, J.M., Ohsumi, T.K., Borowsky, M.L., *et al.* (2013). Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. Cell stem cell *12*, 699-712.

Barozzi, I., Simonatto, M., Bonifacio, S., Yang, L., Rohs, R., Ghisletti, S., and Natoli, G. (2014). Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. Molecular cell *54*, 844-857.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Bau, D., Sanyal, A., Lajoie, B.R., Capriotti, E., Byron, M., Lawrence, J.B., Dekker, J., and Marti-Renom, M.A. (2011). The three-dimensional folding of the alpha-globin gene domain reveals formation of chromatin globules. Nature structural & molecular biology *18*, 107-114.

Bellomy, G.R., and Record, M.T., Jr. (1990). Stable DNA loops in vivo and in vitro: roles in gene regulation at a distance and in biophysical characterization of DNA. Progress in nucleic acid research and molecular biology *39*, 81-128.

Beyaert, R. (2011). SHP works a double shift to control TLR signaling. Nature immunology *12*, 725-727.

Bhatt, D.M., Pandya-Jones, A., Tong, A.J., Barozzi, I., Lissner, M.M., Natoli, G., Black, D.L., and Smale, S.T. (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. Cell *150*, 279-290.

Biddie, S.C., John, S., Sabo, P.J., Thurman, R.E., Johnson, T.A., Schiltz, R.L., Miranda, T.B., Sung, M.H., Trump, S., Lightman, S.L., *et al.* (2011). Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding. Molecular cell *43*, 145-155.

Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS biology *4*, e138.

Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell *144*, 327-339.

Burcin, M., Arnold, R., Lutz, M., Kaiser, B., Runge, D., Lottspeich, F., Filippova, G.N., Lobanenkov, V.V., and Renkawitz, R. (1997). Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone receptors, is identical to the multivalent zinc finger repressor CTCF. Molecular and cellular biology *17*, 1281-1288.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Molecular cell *49*, 825-837.

Chen, X., Barozzi, I., Termanini, A., Prosperini, E., Recchiuti, A., Dalli, J., Mietton, F., Matteoli, G., Hiebert, S., and Natoli, G. (2012). Requirement for the histone deacetylase Hdac3 for the inflammatory gene expression program in macrophages. Proceedings of the National Academy of Sciences of the United States of America *109*, E2865-2874.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., *et al.* (2010). Histone H3K27ac separates active from

poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America *107*, 21931-21936.

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. The Journal of cell biology *145*, 1119-1131.

de Laat, W., Klous, P., Kooren, J., Noordermeer, D., Palstra, R.J., Simonis, M., Splinter, E., and Grosveld, F. (2008). Three-dimensional organization of gene expression in erythroid cells. Current topics in developmental biology *82*, 117-139.

De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS biology *8*, e1000384.

De Santa, F., Narang, V., Yap, Z.H., Tusi, B.K., Burgold, T., Austenaa, L., Bucci, G., Caganova, M., Notarbartolo, S., Casola, S., *et al.* (2009). Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. The EMBO journal *28*, 3341-3352.

de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M., *et al.* (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227-231.

de Wit, E., and de Laat, W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes & development *26*, 11-24.

Dekker, J., Marti-Renom, M.A., and Mirny, L.A. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nature reviews Genetics *14*, 390-403.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

DeKoter, R.P., and Singh, H. (2000). Regulation of B lymphocyte and macrophage development by graded expression of PU.1. Science *288*, 1439-1441.

DeMare, L.E., Leng, J., Cotney, J., Reilly, S.K., Yin, J., Sarro, R., and Noonan, J.P. (2013). The genomic landscape of cohesin-associated chromatin interactions. Genome research *23*, 1224-1234.

Deng, W., Rupon, J.W., Krivega, I., Breda, L., Motta, I., Jahn, K.S., Reik, A., Gregory, P.D., Rivella, S., Dean, A., *et al.* (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell *158*, 849-860.

Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., Pellegrini, M., and Plath, K. (2013). Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. Cell stem cell *13*, 602-616.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Dostie, J., and Dekker, J. (2007). Mapping networks of physical interactions between genomic elements using 5C technology. Nature protocols *2*, 988-1002.

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C., *et al.* (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome research *16*, 1299-1309.

Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y.J., Lee, C., Shendure, J., Fields, S., Blau, C.A., and Noble, W.S. (2010). A three-dimensional model of the yeast genome. Nature *465*, 363-367.

Eijkelenboom, A., Mokry, M., de Wit, E., Smits, L.M., Polderman, P.E., van Triest, M.H., van Boxtel, R., Schulze, A., de Laat, W., Cuppen, E., *et al.* (2013). Genome-wide analysis of FOXO3 mediated transcription regulation through RNA polymerase II profiling. Molecular systems biology *9*, 638.

Engel, N., Raval, A.K., Thorvaldsen, J.L., and Bartolomei, S.M. (2008). Three-dimensional conformation at the H19/Igf2 locus supports a model of enhancer tracking. Human molecular genetics *17*, 3021-3029.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., *et al.* (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43-49.

Escoubet-Lozach, L., Benner, C., Kaikkonen, M.U., Lozach, J., Heinz, S., Spann, N.J., Crotti, A., Stender, J., Ghisletti, S., Reichart, D., *et al.* (2011). Mechanisms establishing TLR4-responsive activation states of inflammatory response genes. PLoS genetics *7*, e1002401.

Ferraiuolo, M.A., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X.Q., Nadler, M., Blanchette, M., and Dostie, J. (2010). The three-dimensional architecture of Hox cluster silencing. Nucleic acids research *38*, 7472-7484.

Ferraiuolo, M.A., Sanyal, A., Naumova, N., Dekker, J., and Dostie, J. (2012). From cells to chromatin: capturing snapshots of genome organization with 5C technology. Methods *58*, 255-267.

Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J., and Lobanenkov, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Molecular and cellular biology *16*, 2802-2813.

Fowler, T., Sen, R., and Roy, A.L. (2011). Regulation of primary response genes. Molecular cell *44*, 348-360.

Fraser, J., Rousseau, M., Shenker, S., Ferraiuolo, M.A., Hayashizaki, Y., Blanchette, M., and Dostie, J. (2009). Chromatin conformation signatures of cellular differentiation. Genome biology *10*, R37.

Fraser, P., and Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. Nature *447*, 413-417.

Garber, M., Yosef, N., Goren, A., Raychowdhury, R., Thielke, A., Guttman, M., Robinson, J., Minie, B., Chevrier, N., Itzhaki, Z., *et al.* (2012). A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. Molecular cell *47*, 810-822.

Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.L., *et al.* (2010). Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity *32*, 317-328.

Gordon, S., and Taylor, P.R. (2005). Monocyte and macrophage heterogeneity. Nature reviews Immunology *5*, 953-964.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., *et al.* (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948-951.

Hakim, O., John, S., Ling, J.Q., Biddie, S.C., Hoffman, A.R., and Hager, G.L. (2009). Glucocorticoid receptor activation of the Ciz1-Lcn2 locus by long range interactions. The Journal of biological chemistry *284*, 6048-6052.

Hargreaves, D.C., Horng, T., and Medzhitov, R. (2009). Control of inducible gene expression by signal-dependent transcriptional elongation. Cell *138*, 129-145.

Hatzis, P., and Talianidis, I. (2002). Dynamics of enhancer-promoter communication during differentiation-induced gene activation. Molecular cell *10*, 1467-1477.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., *et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108-112.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., *et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature genetics *39*, 311-318.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Herschman, H.R. (1991). Primary response genes induced by growth factors and tumor promoters. Annual review of biochemistry *60*, 281-319.

Hotamisligil, G.S. (2006). Inflammation and metabolic disorders. Nature *444*, 860-867.

Hu, G., Cui, K., Northrup, D., Liu, C., Wang, C., Tang, Q., Ge, K., Levens, D., Crane-Robinson, C., and Zhao, K. (2013). H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. Cell stem cell *12*, 180-192.

Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290-294.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., *et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231-1245.

Klenova, E.M., Nicolas, R.H., Paterson, H.F., Carne, A.F., Heath, C.M., Goodwin, G.H., Neiman, P.E., and Lobanenkov, V.V. (1993). CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. Molecular and cellular biology *13*, 7612-7624.

Kohne, A.C., Baniahmad, A., and Renkawitz, R. (1993). NeP1. A ubiquitous transcription factor synergizes with v-ERBA in transcriptional silencing. Journal of molecular biology *232*, 747-755.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology *10*, R25.

Lawrence, T., and Natoli, G. (2011). Transcriptional regulation of macrophage polarization: enabling diversity with identity. Nature reviews Immunology *11*, 750-761.

Levasseur, D.N., Wang, J., Dorschner, M.O., Stamatoyannopoulos, J.A., and Orkin, S.H. (2008). Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. Genes & development *22*, 575-580.

Ley, K., Laudanna, C., Cybulsky, M.I., and Nourshargh, S. (2007). Getting to the site of inflammation: the leukocyte adhesion cascade updated. Nature reviews Immunology *7*, 678-689.

Li, Q., Barkess, G., and Qian, H. (2006). Chromatin looping and the probability of transcription. Trends in genetics : TIG *22*, 197-202.

Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D.C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Human genetics *80*, 224-234.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Lobanenkov, V.V., Nicolas, R.H., Adler, V.V., Paterson, H., Klenova, E.M., Polotskaja, A.V., and Goodwin, G.H. (1990). A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. Oncogene *5*, 1743-1753.

Lower, K.M., Hughes, J.R., De Gobbi, M., Henderson, S., Viprakasit, V., Fisher, C., Goriely, A., Ayyub, H., Sloane-Stanley, J., Vernimmen, D., *et al.* (2009). Adventitious changes in long-range gene expression caused by polymorphic structural variation and promoter competition. Proceedings of the National Academy of Sciences of the United States of America *106*, 21771-21776.

Majumder, P., Gomez, J.A., Chadwick, B.P., and Boss, J.M. (2008). The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. The Journal of experimental medicine *205*, 785-798.

Marenduzzo, D., Faro-Trindade, I., and Cook, P.R. (2007). What are the molecular ties that maintain genomic loops? Trends in genetics : TIG *23*, 126-133.

Medzhitov, R. (2008). Origin and physiological roles of inflammation. Nature *454*, 428-435.

Medzhitov, R., and Horng, T. (2009). Transcriptional control of the inflammatory response. Nature reviews Immunology *9*, 692-703.

Melo, C.A., Drost, J., Wijchers, P.J., van de Werken, H., de Wit, E., Oude Vrielink, J.A., Elkon, R., Melo, S.A., Leveille, N., Kalluri, R., *et al.* (2013). eRNAs are required for p53-dependent enhancer activity and gene transcription. Molecular cell *49*, 524-535.

Meyer, L.R., Zweig, A.S., Hinrichs, A.S., Karolchik, D., Kuhn, R.M., Wong, M., Sloan, C.A., Rosenbloom, K.R., Roe, G., Rhead, B., *et al.* (2013). The UCSC Genome Browser database: extensions and updates 2013. Nucleic acids research *41*, D64-69.

Miele, A., and Dekker, J. (2008). Long-range chromosomal interactions and gene regulation. Molecular bioSystems *4*, 1046-1057.

Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. Cell *128*, 787-800.

Moissiard, G., Cokus, S.J., Cary, J., Feng, S., Billi, A.C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B.R., McCord, R.P., *et al.* (2012). MORC family ATPases required for heterochromatin condensation and gene silencing. Science *336*, 1448-1451.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Spitz, F., and Duboule, D. (2011). A regulatory archipelago controls Hox genes transcription in digits. Cell *147*, 1132-1145.

Murray, P.J., Allen, J.E., Biswas, S.K., Fisher, E.A., Gilroy, D.W., Goerdt, S., Gordon, S., Hamilton, J.A., Ivashkiv, L.B., Lawrence, T., *et al.* (2014). Macrophage activation and polarization: nomenclature and experimental guidelines. Immunity *41*, 14-20.

Natoli, G. (2010). Maintaining cell identity through global control of genomic organization. Immunity *33*, 12-24.

Naumova, N., and Dekker, J. (2010). Integrating one-dimensional and three-dimensional maps of genomes. Journal of cell science *123*, 1979-1988.

Naumova, N., Smith, E.M., Zhan, Y., and Dekker, J. (2012). Analysis of long-range chromatin interactions using Chromosome Conformation Capture. Methods *58*, 192-203.

Nolis, I.K., McKay, D.J., Mantouvalou, E., Lomvardas, S., Merika, M., and Thanos, D. (2009). Transcription factors mediate long-range enhancer-promoter interactions. Proceedings of the National Academy of Sciences of the United States of America *106*, 20222-20227.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., *et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Ohneda, K., and Yamamoto, M. (2002). Roles of hematopoietic transcription factors GATA-1 and GATA-2 in the development of red blood cell lineage. Acta haematologica *108*, 237-245.

Ong, C.T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nature reviews Genetics *12*, 283-293.

Orlova, D.Y., Stixova, L., Kozubek, S., Gierman, H.J., Sustackova, G., Chernyshev, A.V., Medvedev, R.N., Legartova, S., Versteeg, R., Matula, P., *et al.* (2012). Arrangement of nuclear structures is not transmitted through mitosis but is identical in sister cells. Journal of cellular biochemistry *113*, 3313-3329.

Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., and Natoli, G. (2013). Latent enhancers activated by stimulation in differentiated cells. Cell *152*, 157-171.

Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nature genetics *35*, 190-194.

Parada, L.A., Roix, J.J., and Misteli, T. (2003). An uncertainty principle in chromosome positioning. Trends in cell biology *13*, 393-396.

Pennacchio, L.A., Loots, G.G., Nobrega, M.A., and Ovcharenko, I. (2007). Predicting tissue-specific enhancers in the human genome. Genome research *17*, 201-211.

Phillips-Cremins, J.E. (2014). Unraveling architecture of the pluripotent genome. Current opinion in cell biology *28*, 96-104.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., *et al.* (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281-1295.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194-1211.

Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., and Gray, J. (1988). Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proceedings of the National Academy of Sciences of the United States of America *85*, 9138-9142.

Qiao, Y., Giannopoulou, E.G., Chan, C.H., Park, S.H., Gong, S., Chen, J., Hu, X., Elemento, O., and Ivashkiv, L.B. (2013). Synergistic activation of inflammatory cytokine genes by interferon-gamma-induced chromatin remodeling and toll-like receptor signaling. Immunity *39*, 454-469.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279-283.

Ragoczy, T., Telling, A., Sawado, T., Groudine, M., and Kosak, S.T. (2003). A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *11*, 513-525.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell *138*, 114-128.

Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., and Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. Development *132*, 797-803.

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature *489*, 109-113.

Schoenfelder, S., Clay, I., and Fraser, P. (2010a). The transcriptional interactome: gene expression in 3D. Current opinion in genetics & development *20*, 127-133.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., *et al.* (2010b). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nature genetics *42*, 53-U71.

Schonheit, J., Kuhl, C., Gebhardt, M.L., Klett, F.F., Riemke, P., Scheller, M., Huang, G., Naumann, R., Leutz, A., Stocking, C., *et al.* (2013). PU.1 level-directed chromatin structure remodeling at the Irf8 gene drives dendritic cell commitment. Cell reports *3*, 1617-1628.

Scott, E.W., Simon, M.C., Anastasi, J., and Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science *265*, 1573-1577.

Servant, N., Lajoie, B.R., Nora, E.P., Giorgetti, L., Chen, C.J., Heard, E., Dekker, J., and Barillot, E. (2012). HiTC: exploration of high-throughput 'C' experiments. Bioinformatics *28*, 2843-2844.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., *et al.* (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.

Simonis, M., Klous, P., Homminga, I., Galjaard, R.J., Rijkers, E.J., Grosveld, F., Meijerink, J.P., and de Laat, W. (2009). High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology. Nature methods *6*, 837-842.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006a). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nature genetics *38*, 1348-1354.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006b). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nature genetics *38*, 1348-1354.

Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. Nature methods *4*, 895-901.

Smale, S.T. (2010). Selective transcription in response to an inflammatory stimulus. Cell *140*, 833-844.

Smallwood, A., and Ren, B. (2013). Genome organization and long-range regulation of gene expression by enhancers. Current opinion in cell biology *25*, 387-394.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994-1004.

Spitz, F., and Furlong, E.E. (2012). Transcription factors: from enhancer binding to developmental control. Nature reviews Genetics *13*, 613-626.

Splinter, E., de Wit, E., Nora, E.P., Klous, P., van de Werken, H.J.G., Zhu, Y., Kaaij, L.J.T., van IJcken, W., Gribnau, J., Heard, E., *et al.* (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes & development *25*, 1371-1383.

Splinter, E., de Wit, E., van de Werken, H.J., Klous, P., and de Laat, W. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. Methods *58*, 221-230.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes & development *20*, 2349-2354.

Staber, P.B., Zhang, P., Ye, M., Welner, R.S., Nombela-Arrieta, C., Bach, C., Kerenyi, M., Bartholdy, B.A., Zhang, H., Alberich-Jorda, M., *et al.* (2013). Sustained PU.1 levels balance cell-cycle regulators to prevent exhaustion of adult hematopoietic stem cells. Molecular cell *49*, 934-946.

Strickfaden, H., Zunhammer, A., van Koningsbruggen, S., Kohler, D., and Cremer, T. (2010). 4D chromatin dynamics in cycling cells: Theodor Boveri's hypotheses revisited. Nucleus *1*, 284-297.

Szutorisz, H., Dillon, N., and Tora, L. (2005). The role of enhancers as centres for general transcription factor recruitment. Trends in biochemical sciences *30*, 593-599.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. Cell *140*, 805-820.

Umbarger, M.A., Toro, E., Wright, M.A., Porreca, G.J., Bau, D., Hong, S.H., Fero, M.J., Zhu, L.J., Marti-Renom, M.A., McAdams, H.H., *et al.* (2011). The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation. Molecular cell *44*, 252-264.

van de Werken, H.J., Landan, G., Holwerda, S.J., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Oz, Y., Bouwman, B.A., *et al.* (2012). Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nature methods *9*, 969-972.

Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. The EMBO journal *26*, 2041-2051.

Vernimmen, D., Marques-Kranc, F., Sharpe, J.A., Sloane-Stanley, J.A., Wood, W.G., Wallace, H.A., Smith, A.J., and Higgs, D.R. (2009). Chromosome looping at the human alpha-globin locus is mediated via the major upstream regulatory element (HS -40). Blood *114*, 4253-4260.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., *et al.* (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854-858.

Vostrov, A.A., and Quitschke, W.W. (1997). The zinc finger protein CTCF binds to the APBbeta domain of the amyloid beta-protein precursor promoter. Evidence for a role in transcriptional activation. The Journal of biological chemistry *272*, 33353-33359.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., *et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, Q., Carroll, J.S., and Brown, M. (2005). Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. Molecular cell *19*, 631-642.

Wei, Z., Gao, F., Kim, S., Yang, H., Lyu, J., An, W., Wang, K., and Lu, W. (2013). Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. Cell stem cell *13*, 36-47.

Weintraub, H., and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. Science *193*, 848-856.

Williams, A., Spilianakis, C.G., and Flavell, R.A. (2010). Interchromosomal association and gene regulation in trans. Trends in genetics : TIG *26*, 188-197.

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nature genetics *30*, 167-174.

Xi, H., Shulha, H.P., Lin, J.M., Vales, T.R., Fu, Y., Bodine, D.M., McKay, R.D., Chenoweth, J.G., Tesar, P.J., Furey, T.S., *et al.* (2007). Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. PLoS genetics *3*, e136.

Yoon, Y.S., Jeong, S., Rong, Q., Park, K.Y., Chung, J.H., and Pfeifer, K. (2007). Analysis of the H19ICR insulator. Molecular and cellular biology *27*, 3499-3510.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes & development *25*, 2227-2241.

Zhang, H., Jiao, W., Sun, L., Fan, J., Chen, M., Wang, H., Xu, X., Shen, A., Li, T., Niu, B., *et al.* (2013a). Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. Cell stem cell *13*, 30-35.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.

Zhang, Y., McCord, R.P., Ho, Y.J., Lajoie, B.R., Hildebrand, D.G., Simon, A.C., Becker, M.S., Alt, F.W., and Dekker, J. (2012). Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell *148*, 908-921.

Zhang, Y., Wong, C.H., Birnbaum, R.Y., Li, G., Favaro, R., Ngan, C.Y., Lim, J., Tai, E., Poh, H.M., Wong, E., *et al.* (2013b). Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. Nature *504*, 306-310.

Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., *et al.* (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nature genetics *38*, 1341-1347.

Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nature reviews Genetics *12*, 7-18.

Zlatanova, J., and Thakar, A. (2008). H2A.Z: view from the top. Structure *16*, 166-179.

ACKNOWLEDGEMENTS

I miei primi ringraziamenti sono rivolti al mio supervisor, Gioacchino Natoli, che con fiducia mi ha accolta nel suo gruppo di ricerca, affidandomi un progetto tanto interessante quanto ambizioso. La sua professionalità e la sua competenza sono stati un'inesauribile fonte di esempio e motivazione durante questo mio percorso professionale.

Un grazie speciale va a tutti i membri del gruppo GN, in particolar modo a Serena Ghisletti per essere stata un'insostituibile punto di riferimento in laboratorio, a Iros Barozzi che ha svolto tutte le analisi bioinformatiche di questo lavoro (avendo anche l'infinita pazienza di farmele comprendere) e a Giulia della Chiara, che più di tutti ha potuto capire le difficoltà tecniche del mio progetto.

Many thanks to my co-supervisors, Bruno Amati at IEO and Matthias Merkenschlager at MRC (London, UK), for critical discussions and support.

Grazie a Luca Giorgetti dell' Institut Curie a Parigi che ha posto le fondamenta di questa ricerca.

I would like to thank Wouter de Laat, Britta Bouwman and Elzo de Wit at Hubrecht Institute (Utrecht, The Netherlands), for teaching me how to perform 4C experiments.

Grazie alla NGS Unit, specialmente a Luca Rotta, per aver assecondato tutte le mie innumerevoli e bizzarre richieste di sequenziamento.

Nulla sarebbe stato possibile senza il calore umano di persone speciali che mi sono state vicine tra i banconi di laboratorio e non solo. Alessia, Valeria e Valentina, non ci sono parole per descrivere quanto siano stati importanti l'affetto e il supporto che mi avete dato in questi anni. Parinaz and Gianluca (Giangi for friends), you are the best "benchmates" ever! Un grazie a Camma e a Max per le utilissime "scientific discussions" al tavolo da ping-pong. Un ringraziamento speciale va inoltre a tutti i miei compagni di dottorato e a quei colleghi che mi hanno vissuta e sopportata quotidianamente sul posto di lavoro.

I miei ringraziamenti sono rivolti anche a coloro che mi hanno sostenuta dalla ridente Appiano City: Alessio, Marchino, Alison e Marta.

Fede e Silvia, il vostro nome sulla pagina dei ringraziamenti delle mie tesi è ormai diventato un solito cliché…

Omar, semplicemente grazie.

Infine sarò sempre a grata alla mia famiglia. Mamma, papà, Marco e Cristiana, il vostro amore rende sempre tutto possibile!