universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Does the transcription factor C/EBPa induce topological change after binding to chromatin?

verfasst von / submitted by

Julia Erber, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirement for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl it Studienblatt / degree code as it appears on the student record sheet:	UA 066 834
Studienrichtung it. Studienblatt / Degree program as it appears on the student record sheet:	Masterstudium Molekulare Biologie UG2002
Betreut von / Supervisor:	ao. Univ.-Prof. Dipl.-Ing. Dr. Johannes Nimpf

Table of content

Abstract

Lineage-instructive transcription factors can short-circuit gene expression by binding DNA sequence-specific to regulatory elements, resulting in the activation of new genes while the old ones are silenced. However, gene expression is also regulated at a higher order by chromatin modification and the complex three-dimensional organization of chromatin. Chromatin contact maps obtained by chromosome conformation capture techniques, such as Hi-C, revealed that chromatin can be separated at the megabase (Mb) level into active (A) and inactive (B) compartments. However, the role of lineage-instructive transcription factor in shaping the high order of genome organization remains elusive. To study these regulatory processes, the Graf laboratory has recently established a system converting B cells into induced macrophages by the overexpression of the transcription factor C/EBPα. During this process called transdifferentiation, changes in gene expression are associated with alterations of A and B compartments, potentially linked to C/EBPα binding. In particular, C/EBPα binds in a region 500 kbp downstream of the *JUN* locus, resulting in the switch of the entire region from B to A compartment and the activation of JUN expression. To investigate the topological role of C/EBPα binding, a CRISPR/Cas9 genome editing technology to delete the binding-site was employed. Next, the transdifferentiation for the wildtype (WT) and C/EBPα knock-out (KO) cells was compared by FACS analysis and qPCR. That revealed a faster cell conversion compared to the wildtype and an impaired yet, not abolished JUN expression. Using Hi-C, it was observed that the B to A compartment switch at the *JUN* locus was not affected due to the deletion. Though, the genome topology displayed alterations, as one out of two loop domains formed by C/EBPα was destroyed, and in general fewer interactions were observed. Overall, this suggests that C/EBPα binding site KO affects JUN expression and accelerates transdifferentiation but does not affect the switch from B to A compartment potentially compensated by additional C/EBPα binding-sites.

1. Introduction

Multicellular organisms are defined by the ability to differentiate and acquire new fates. The gene expression profile ultimately determines the specific cellular state (Davidson, 2010). Beginning with the first division of a fertilized egg, a sophisticated sequence of cell-state transitions leads to the generation of a variety of different cell types. These cell-fate decisions are driven by the microenvironment of a cell consisting of cell-cell interaction, cytokines, and other factors triggering signal transduction into the nucleus (Vaquerizas *et al.*, 2009). The upregulation of genes characteristic for new cell types and the downregulation of respective old ones is crucial in developmental differentiation and hence, in cell fate changes. Here, signaling pathways play a fundamental role in the activation or silencing of DNA sequence-specific regulators with transcription factors being the most distinguished. Transcription factors function by binding to specific DNA motifs at enhancers or promoters to regulate gene expression. While the latter ones are found adjacent to the transcription start site of a gene, enhancers can be localized distant from their gene of interest (Lambert *et al.*, 2018).

1.1. Cell fate and lineage-instructive transcription factors

Usually in a cell's microenvironment located modifiers initiate signaling pathways that eventually lead to changes in gene expression. However, this process can be short-circuited by DNA sequence-specific lineage-instructive transcription factors (liTFs) binding to regulatory elements (Ralph Stadhouders, 2019). They are essential for achieving a new gene expression program as well as inactivating key regulators of the old cell type thereby altering the cell's identity (Graf and Enver, 2009). The first-time lineage conversion (also called transdifferentiation) by transcription factors was reported, by Harold Weintraub's laboratory when they demonstrated the forced expression of MyoD in a fibroblast cells can turn them into myoblasts (Davis, Weintraub and Lassar, 1987). This process transforms one specialized cell type into another without returning first into a pluripotent state (Graf and Enver, 2009). Further proof for the reciprocal regulation of lineage-restricted genes, originated from the hematopoietic system. The expression of the transcription factor GATA1 in monocytic cell lines activates the expression of erythroid-megakaryocyte lineage markers as well as downregulates

monocytic markers (Kulessa, H. Frampton, J. & Graf, 1995). This process functions vice versa as well: By expressing PU.1 in an erythroid-megakaryocyte cell line, the conversion into a monocytic lineage is induced and at the same time the inhibition of GATA1 (Heyworth and Pearson, Stellaay, 2002). This mechanism of switching between GATA1 or PU.1 depicts a fundamental principle of lineage specification and represents one of the first lineage decisions in hematopoietic development and these two transcription factors act as a prime example for cross antagonistic transcription factor interactions (Arinobu *et al.*, 2007). Another example of transdifferentiation has been established in the Graf laboratory when they showed that expression of C/EBPα, a transcription factor essential for the transition of common myeloid progenitors into granulocyte/monocyte progenitors, converts committed B cell progenitors into macrophages (Xie *et al.*, 2004).

Beyond the classical li-TF mechanisms, genome topology has recently emerged as a new player of gene regulation. Chromatin contact maps obtained by chromosome conformation capture techniques, such as Hi-C, revealed a high degree of spatial organization of the genome into the nucleus (Lieberman Aiden 2009).

1.2. An introduction to Hi-C

Hi-C is a chromosomal capture methodology that sheds light on chromatin interactions and the organization of chromatin genome-wide. The number of interactions between specific pairs of loci is quantified resulting in a contact map of Hi-C data. (Lieberman-Aiden *et al.*, 2009). Briefly, nuclei are first crosslinked, and the chromatin is digested with a restriction enzyme. Resulting overhangs are filled and marked with a biotin-labeled residue. Next, a proximity ligation is performed creating chimeric molecules made of fragments that were in close spatial proximity within the nucleus. To generate a Hi-C library, the DNA is sheared, and the biotinylated fragments are pulled down and then analyzed by paired-end sequencing (Rao *et al.*, 2014) (**Figure 1**).

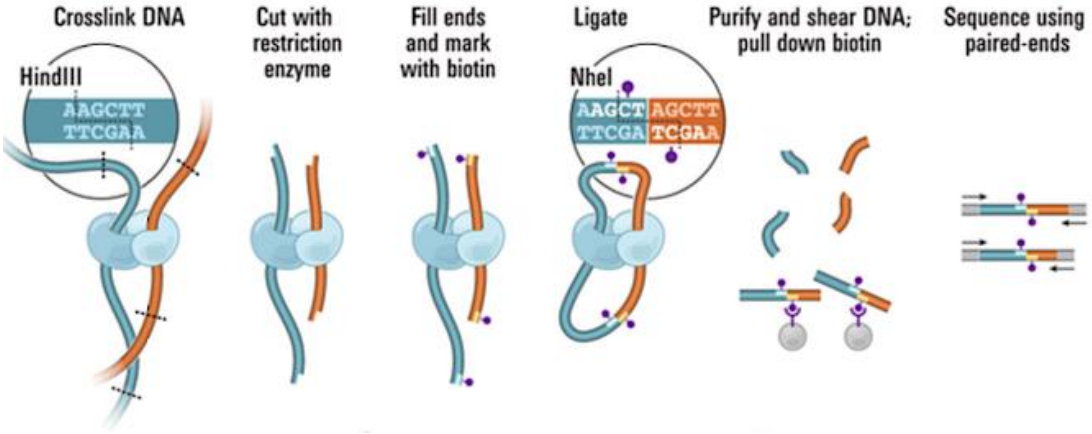


Figure 1 – Hi-C experimental outline. Cells are crosslinked and cut with a restriction enzyme, here for instance by HindIII. In our study MboI, a 4-cutter restriction enzyme was used. Then the sticky ends (digested) are filled with biotinylated residues, resulting in blunt ends. Ligation is performed to produce chimeric fragments. Through the fill-in and ligation, a new site is generated. Following DNA purification and shearing, the biotinylated fragments are pulled down by streptavidin beads. Finally, paired-end sequencing is performed to determine interacting regions. The DNA fragments are indicated in dark blue and orange. Proteins that can facilitate chromatin interactions are marked in light blue. The purple dot illustrates the biotinylated residue (Lieberman-Aiden *et al.*, 2009).

1.2.1. Data resolution

The resolution of Hi-C data is determined by two factors. First, the frequency of the restriction site of the enzymes used during digestion in the experiment determines the fragment size and represents the final resolution limit. Secondly, the other crucial factor is the depth of sequencing. Additionally, the number of cells and the quality of the library define the overall number of unique chimeric fragments that represents the library complexity (Lajoie, Dekker and Kaplan, 2016).

1.2.2. Data processing

Sequencing of Hi-C library is done by paired-end technology producing one read from each 5'end of a molecule. First, paired-end reads are independently mapped to the reference genome and the corresponding fragment is assigned to the same genomic location (Lajoie, Dekker and Kaplan, 2016). Next, all non-informative fragments are excluded including digested unligated fragments ("dangling end") or circular ligated fragment ("self-circle") as well as identical PCR artifacts ("duplicates") (Hansen *et al.*, 2019). Then the data is clustered into genomic bins of a fixed size ranging from 1 kb to 1 Mb increasing the signal to noise ratio. Finally, balancing corrects for biases of the Hi-C data (Lajoie, Dekker and Kaplan, 2016).

1.2.3. Hi-C maps

After a Hi-C experiment, the data generated is visualized in a Hi-C map. It can be described as a list of DNA-DNA contacts. A contact matrix “M” is produced by dividing the genome into regions (or “bins”) of a fixed size where the matrix entry “$M_{i,j}$” is defined by the number of contacts detected between “locus i” and “locus j”. A contact itself represents a read pair that remains after filtering (exclusion of duplicates, dangling ends, duplicates, etc.). Most often these contact maps are depicted as heat maps with the intensity correlating to the contact frequency (Rao *et al.*, 2014) (Lieberman-Aiden *et al.*, 2009).

1.3. Principals of genome folding

Hi-C analysis revealed that DNA inside the nucleus within eukaryotic cells is packaged into hierarchical layers. The correct folding of the three-dimensional conformation of chromatin is complex and crucial for essential biological functions including gene expression and cell fate. It is an extremely organized series of actions and highly dynamic. Correct structuring involves the formation of genome territories (**Figure 2a**) and intrachromosomal hubs (**Figure 2b**), chromosomal compartmentalization (**Figure 2c**), the organization of insulated spatial neighborhoods (**Figure 2d**) and interactions occurring within them (**Figure 2e**) (Ralph Stadhouders, 2019).

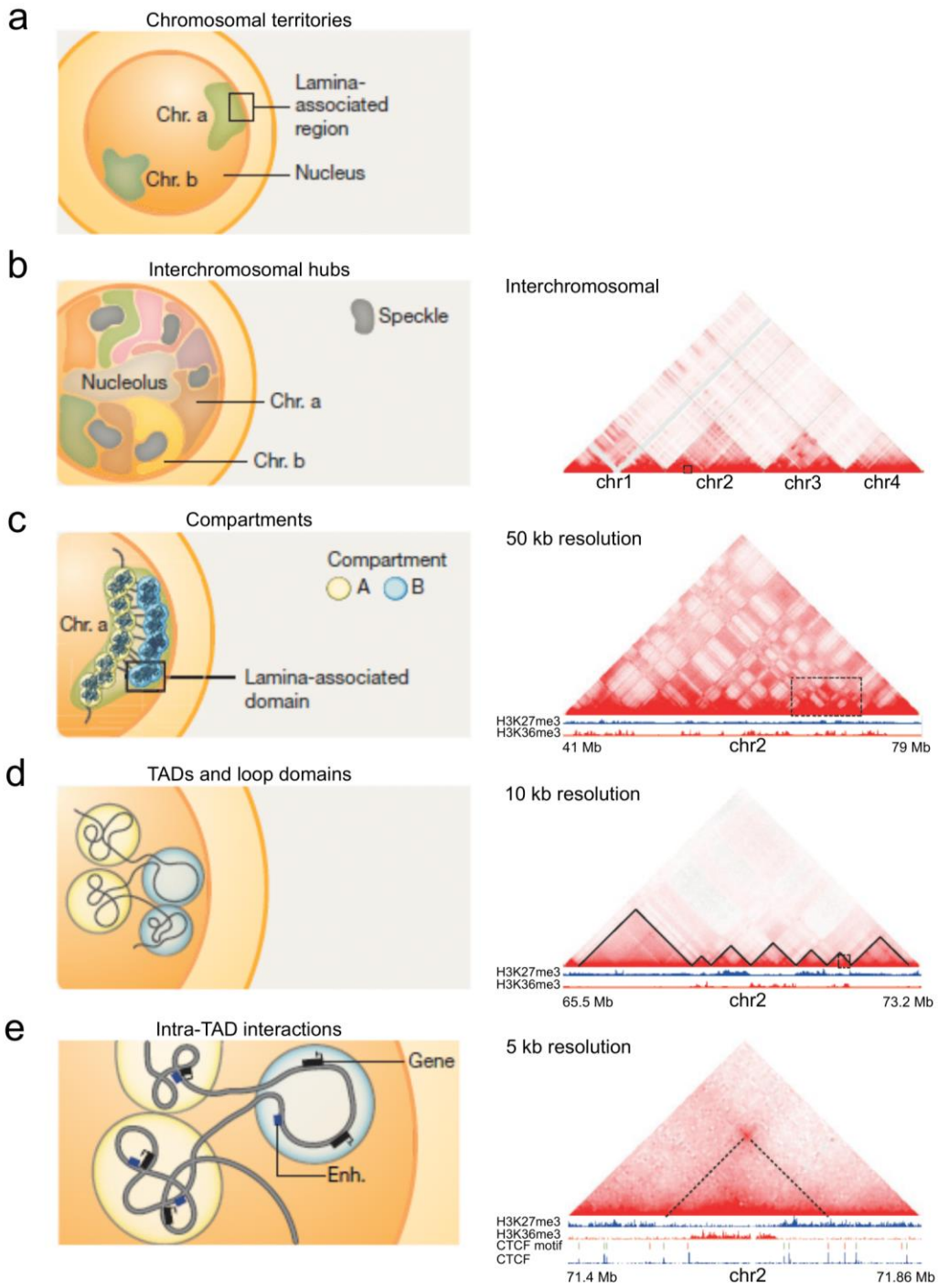


Figure 2 – Overview of 3D genome folding. (**a-e**) Different layers of chromatin compaction with a corresponding exemplary Hi-C heat map are ordered from top to bottom based on the resolution. (**a**) Each chromosome (indicated by chr.a or chr.b) is located within a distinct territory. (**b**) They form interchromosomal hubs with active chromatin situated near nuclear speckles (euchromatin) and inactive chromatin located next to the nucleolus (heterochromatin). On a Hi-C map (right) on the lowest level of resolution different chromosomes can be visualized as red triangles (here Chr 1 to 4). (**c**) Each chromosome within a chromosomal territory is composed of active and inactive chromatin corresponding to the A and B compartment (indicated in yellow and blue). The A compartment is positioned more central while the B compartment forms lamina-associated domains (LAD) at the edge of the nucleus. The compartments on a Hi-C map (right side, compartment indicated as a dotted rectangular) can be identified by the corresponding active or inactive chromatin marks or by principal component analysis. (**d**) Within each compartment, chromatin is organized into loop structures and topological associated neighborhoods (TADs) (altogether termed insulated spatial neighborhoods). Loops are marked by a black line on the Hi-C map (right). (**e**) Inside TADs, regulatory elements such as promoters and enhancers can interact to control gene expression.

Regulatory elements are restrained to the TAD-boundaries. Interactions between two specific loci are visualized as a punctuated interaction signal on a Hi-C map. The specific interacting loci are indicated by a dotted line (Ralph Stadhouders, 2019) (Bonev and Cavalli, 2016).

1.3.1. Chromosome territories and chromosomal compartmentalization

During interphase individual chromosomes are situated in distinct nuclear territories which represents the largest scale of chromatin organization (**Figure 2a**). They are organized into two interchromosomal hubs. Gene-dense and active chromatin (euchromatin) is associated with RNA polymerase II and is located near nuclear speckles (Cremer and Cremer, 2010) (Ralph Stadhouders, 2019). Inactive chromatin (heterochromatin) can be found close to the nucleolus and is concentrated with ribosomal RNA genes and centromeric chromatin (**Figure 2b**) (Ralph Stadhouders, 2019).

Furthermore, at the megabase scale, each chromosome can be divided into an active A compartment and an inactive B compartment. Each compartment favors interactions with its kind (**Figure 2c**) (Lieberman-Aiden et al., 2009). Chromatin belonging to the A compartment consists of primarily transcribed genes and active histone modifications such as H3K27 acetylation. Correspondingly, the B compartment displays mainly inactive genes with repressive histone modifications including H3K27me3 and H3K9me3 (Ralph Stadhouders, 2019).

Phase separation, a process where proteins quickly concentrate in the cell and form condensates which are liquid-like droplets (**Figure 3**) has been recently proposed to act in chromatin compartmentalization through proteins associating with active or inactive chromatin, respectively (Shin et al., 2018) (Lesne et al., 2019).

1.3.2. Intrachromosomal topological domains

Each compartment is further segregated into topological associated domains (TADs) and loop domains, approximately 0.1 – 1 Mb (Dixon et al., 2012) (Nora et al., 2012) (Sexton et al., 2012). Therefore, TADs often present homogenous chromatin marks corresponding to an active or a repressive state which dictates their location in the A or B compartment (**Figure 2d**) (Le Dily and Beato, 2015). TADs preferentially exhibit intra-domain interactions and have boundaries (regions with low interaction signals (Andrey and Mundlos, 2017)) enriched for CTCF-binding sites, as well as cohesin and highly transcribed genes (Rowley and Corces, 2018). It is believed that TADs are formed through a so-called loop extrusion mechanism (**Figure 3**). A loop extrusion factor (thought to be cohesion) builds a loop by extruding chromatin through its ring-shaped

structure. Once an extrusion barrier is reached this process stops. CTCF represents this barrier which must be bound to DNA in a specific orientation to interact with cohesin (Rao *et al.*, 2014) (Fudenberg *et al.*, 2016). Although the deletion of CTCF disrupts the TAD structure, it only slightly impairs the gene expression (Despang *et al.*, 2019) (Nora *et al.*, 2017) (Stik *et al.*, 2020).

Still, there is no universal definition for TADs as they present in different sizes and often with so-called sub-TADs and loops. Yet, they are all related to and dependent on cohesin. (Phillips-Cremins *et al.*, 2013) (Bonev and Cavalli, 2016). The current nomenclature can be misleading as insulated neighborhoods such as TADs, loop domains, and CTCF contact domains all describe loops with CTCF boundaries (Schoenfelder and Fraser, 2019). Recently, however, a new concept of compartmental domains was introduced. They exist besides CTCF loop domains and originate directly from compartmentalization into A and B compartments. This new type of TADs is defined by the chromatin state instead of proteins like CTCF (Rowley *et al.*, 2017).

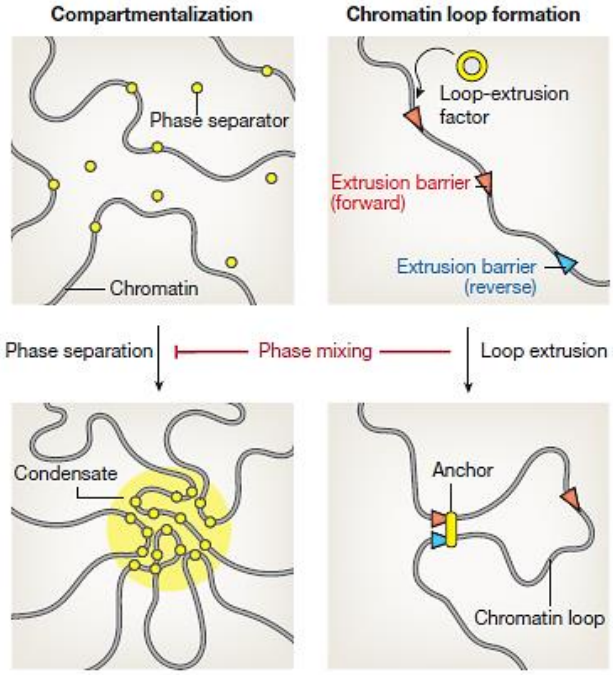


Figure 3 – Mechanisms of genome folding. (Left) Compartmentalization of the genome by phase separation. Phase separators (yellow) (for instance, transcription factors or chromatin-associated proteins) bound to chromatin, form condensates much like oil drops in water. 3D hubs are generated by condensates facilitating gene regulation and displaying a high concentration of important associated proteins. (Right) Formation of chromatin loops by extrusion. A loop-extrusion factor such as cohesin extrudes chromatin through its ring until the process is blocked at an extrusion barrier. A barrier could be CTCF that is bound in forward (red) and reverse (blue) orientation as interaction with cohesin only occurs when in convergent orientation. This creates a stable loop. Additionally, loop extrusion can antagonize phase separation (Ralph Stadhouders, 2019).

1.3.3. Gene regulatory elements and their role in chromatin architecture

Topologically associated domains have been considered as units of gene regulation by bringing promoters and there distantly located enhancer into close spatial proximity (Ralph Stadhouders, 2019).

However, many other proteins are implicated in forming interactions between promoters and enhancers among them TF and chromatin-associated proteins. For instance, Mediator, a coordinator of transcription, the TF KLF4, and Pax5, are proofed to form physical contacts of promoters and enhancers (Ralph Stadhouders, 2019). Additionally, another ubiquitously expressed transcription factor Ying Yang 1 (YY1) assists in creating promotor-enhancer interaction through dimerization and tethers them together, similarly to CTCF in TAD formation. In summary, this provides proof for the importance of transcription factors in such a process (Weintraub *et al.*, 2017).

However, recent studies interrogate the connection between TADs and gene expression. They showed that TADs and particularly CTCF sites are unnecessary for proper gene expression during development but instead provide precision and robustness (Ghavi-Helm *et al.*, 2019) (Despang *et al.*, 2019) (Stik *et al.*, 2020). Thereby, one could conclude that TADs are only important for a limited number of genes or to further adjust gene transcription (Ralph Stadhouders, 2019). This supports the significance of regulatory elements and (lineage-instructive) transcription factors in shaping the topological genome for correct gene expression (Stik *et al.*, 2020).

1.4. Genome conformation and its role in cell fate

Cell-fate decisions go hand in hand with changes in the three-dimensional conformation of the genome (Ralph Stadhouders, 2019). For TF to bind DNA, chromatin modifications (epigenome) and the complex three-dimensional structure of the chromatin must be navigated through. Altogether, a permissive/non-permissive or active/inactive chromatin structure is arranged (Azagra *et al.*, 2020). Therefore, cell identity can be regarded as an emergent property that results from the interaction of gene expression, the epigenome, and the genome topology triggered by a lineage-instructive transcription factor (Ralph Stadhouders, 2019).

Overall, the arrangement of A and B compartment do not change much throughout different cell types. However, experiments of B cell reprogramming into induced pluripotent stem cells (iPSC) show compartment switching, meaning a shift in the arrangement of A-B compartments, during differentiation, indicating a role of transcription factors in A-B compartmentalization and consequently gene positioning

(Stadhouders *et al.*, 2018). Also, a substantial part of TAD boundaries seems to be cell-type specific. All this supports the notion of a cell-type-specific aspect of the genome conformation (Bonev *et al.*, 2017) (Stadhouders *et al.*, 2018), with in general changes in gene expression preceded by changes in the genome topology (Stadhouders *et al.*, 2018).

1.5. Cell conversion model

The Graf laboratory as recently established a human leukemic B cell line (BLaER) that was derived from the RCH-ACV cell line. C/EBPα fused to an estrogen receptor hormone-binding domain (C/EBPαER) was introduced into the cell line. Induction with β-estradiol (E2), relocates the transcription factor C/EBPα to the nucleus and thus, the conversion of this B cell into an induced macrophage (**Figure 4a**). During this process, B cell-specific genes become downregulated and macrophage-specific genes become upregulated (Rapino et al., 2013).

1.5.1. Changes of genome topology during transdifferentiation

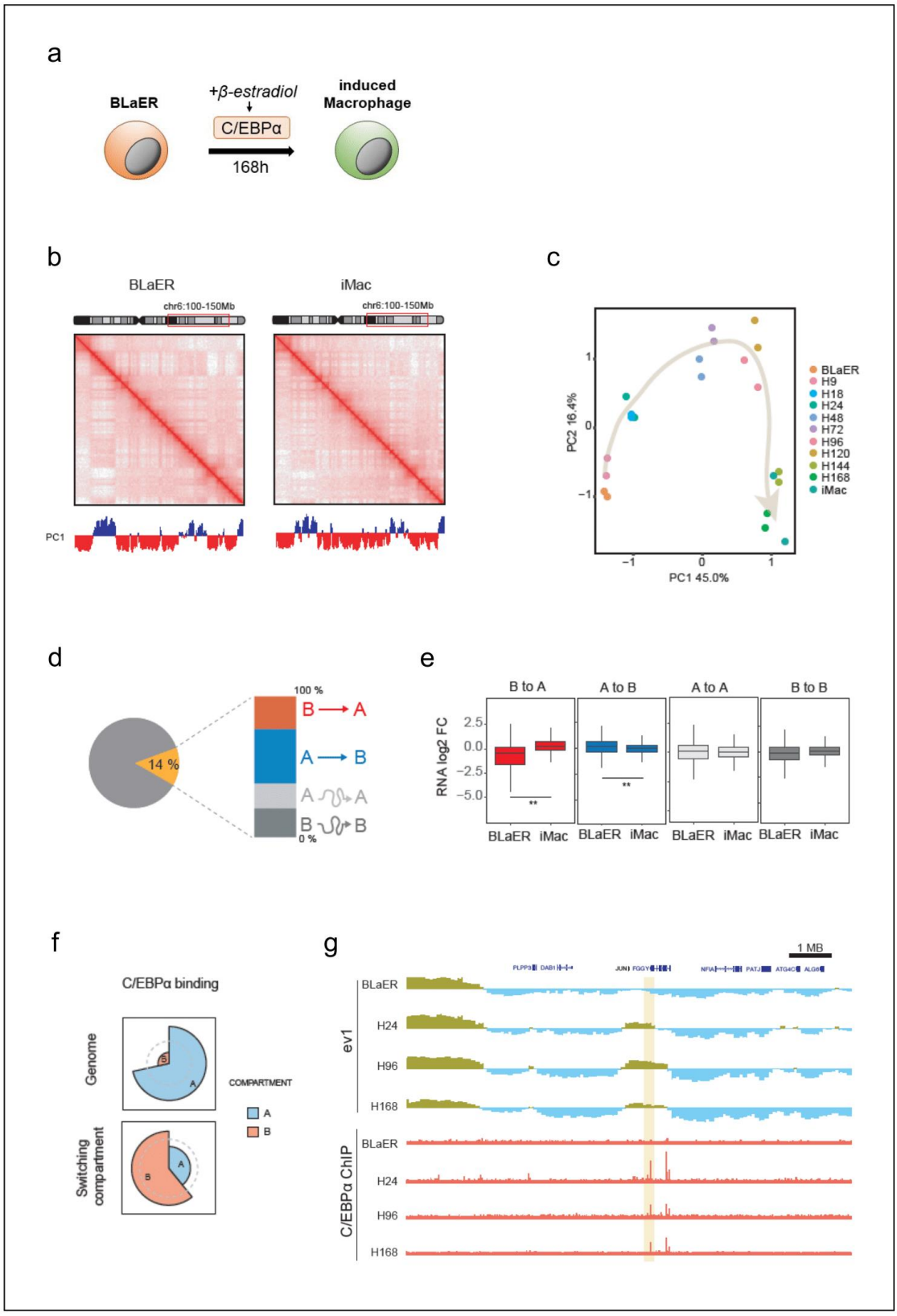


Figure 4 – Compartment switching during transdifferentiation. (**a**) BLaER transdifferentiation model. BLaER cell converts into an induced macrophage upon induction with β-estradiol, resulting in the expression of C/EBPα. (**b**) Hi-C heat map of a region on chromosome 6 before (left) and after (right) transdifferentiation from a BLaER into an induced macrophage (iMac). Below principal component 1 (PC1) indicates the two compartments within this region. (**c**) Principal component analysis (PCA) of the whole genome during transdifferentiation. PC1 explained 45 % and PC2 explained 16.4 % of the variability within the data set. The grey arrow indicates the trajectory transdifferentiation follows. (**d**) Percentage of the genome undergoing compartment switches. 14 % (yellow) of the whole genome (grey) switched compartment between A and B. (**e**) mRNA expression levels of BLaER and iMac during compartment switching. mRNA levels were upregulated once a B to A switch (red panel) occurred and vice versa correspondingly downregulated (blue panel). (**f**) C/EBPα binding preferences. (top) Throughout the genome C/EBPα primarily binds to the A compartment (blue). (Bottom) In switching compartments, C/EBPα preferentially binds B compartments (red) converting to A compartments. (**g**) Compartmentalization of *JUN* locus correlated with C/EBPα binding. The *JUN* locus, a B compartment (dark yellow) in BLaER cells, switched into an A compartment (blue) during transdifferentiation into iMacs. C/EBPα ChIP-seq data (red) revealed the protein is bound to JUN 24h post induction throughout cell conversion. (Stik *et al.*, 2020) *(Stik et al., personal communication)*.

This model (**Figure 4a**) was used to investigate topological changes of the genome during transdifferentiation (Stik *et al.*, 2020). It was reported while most of the compartments remained stable (**Figure 4b**), approximately 14 % of the genome was undergoing a compartment switch (**Figure 4d**), following a transdifferentiation trajectory when analyzed by principal component analysis (PCA) (**Figure 4c**). The transcriptional changes correlated with the altered compartmentalization (**Figure 4e**) (Stik *et al.*, 2020). Also, C/EBPα binding was explored. In general, throughout the genome C/EBPα preferentially binds to the A compartment. However, focusing on that 14 % of the genome undergoing a compartment switch, C/EBPα favored regions that turn from B into an A compartment during transdifferentiation (**Figure 4f**) (Stik *et al.*, personal communication). Interestingly, one of the regions that switched from B to A compartment is located on chromosome 1 between the *JUN* and *FGGY* locus and coincides with three C/EBPα binding-sites. During transdifferentiation, already after 24h, this region at the *JUN* locus converted into an A compartment. Intriguingly, ChIP-seq analysis revealed the TF C/EBPα also binds after 24h following induction of transdifferentiation (**Figure 4g**) (Stik *et al.*, personal communication). Thus, implying that C/EBPα might be responsible for the switch from inactive B to active A compartment (**Figure 5**). Consequently, lineage instructive transcription factor could have an as of yet undescribed function.

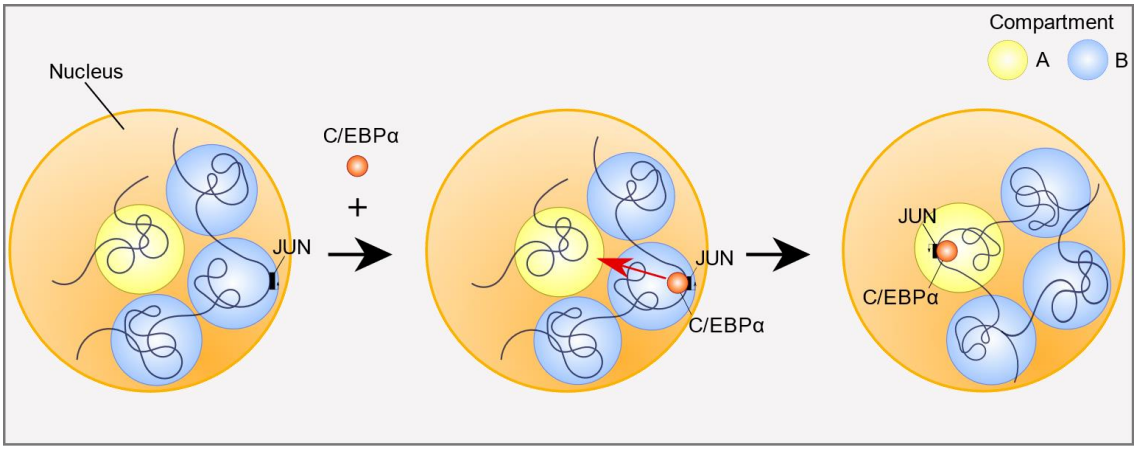


Figure 5 – Possible compartment switch by C/EBPα. Potentially, once C/EBPα (orange) is expressed and enters the nucleus, it binds to its motif adjacent to JUN in the inactive B compartment (blue) and triggers the switch into a more centrally located active A compartment (yellow).

To analyze this possible connection between the transcription factor and an activating compartment switch (**Figure 5**), CRISPR technology was employed to destroy specific C/EBPα binding-sites in this region of interest. Binding-sites were identified based on epigenetic analysis determining the accessibility of chromatin (ATAC-seq), active enhancer chromatin marks (H3K27ac), and the binding-motif of C/EBPα (ChIP-seq). Subsequently, the resulting knockout cell line will be investigated throughout transdifferentiation from BLaER to iMac to determine the effects of the deletion. Additionally, the gene expression of JUN and FGGY as well as B cell and myeloid genes will be monitored to examine possible alterations. Finally, the genome topology at the *JUN* locus will be studied with Hi-C to test whether in the absence of C/EBPα binding the B to A compartment switch no longer occurs.

2. Material & Methods

2.1. Cell Culture

BLaER cell line was previously established in the laboratory of Thomas Graf. This system was derived from a RCH-ACV cell line in which C/EBPα fused to a β-estradiol receptor (C/EBPαER) as well as a GFP was introduced giving rise to the human leukemic B cell line, BLaER (Rapino *et al.*, 2013). Transdifferentiation can be induced within 7 days by introducing 100 nM β-estradiol, human colony-stimulating factor 1 (hCsf-1) (10 ng/ul) and human interleukin-3 (hIL-3) (10 ng/ul). For cell culture RPMI 25 mM HEPES medium (containing 10% inactivated FBS, 1% Pen/strep (10000 units/mL penicillin and 10000 ug/mL streptomycin), 1% L-glutamine 200 mM and 0.1% 2-Mercaptoethanol 50 mM) was used.

2.2. CRISPR Cas9-guided knockout of C/EBPα motifs

To generate new CRISPR-edited cell lines based on the BLaER cell line, C/EBPα binding motifs was identified based on previously generated ATAC-seq as well as C/EBPα and H3K27ac ChIP-seq data on the human chromosome 1 between the locus of *JUN* and *FGGY* at approximately 500 kbp of JUN.

2.2.1. CRISPR plasmid digestion

gRNAs were designed and synthetized by IDT ® and cloned into px330_mCherry or px459_PuroR (Addgene, plasmid #62988) that expressed the Cas9 protein as well. Digestion of the plasmids was performed using 4 µg of plasmid DNA, 4 µl of BbsI High Fidelity NEB 20 U/µl (R3539L), 2.5 µl of CutSmart® Buffer NEB, 2 µl of Alkaline Phosphatase Calf Intestinal (CIP) NEB 10 U/µl (M0290S) and ddH2O up to 25 µl during 72 hours at 37 °C. Additional 2 µl BbsI High Fidelity NEB 20 U/µl (R3539L) was added about 1.5h after the start of incubation. The digested vectors were visualized with a 0.6 % agarose gel, followed by a purification using QIAquick-Gel Extraction Kit (Cat. Nos 28704 and 28706)

2.2.2. Guide RNAs annealing and phosphorylation

Annealing reaction was performed using 100 µM of forward and reverse gRNA sequences (Integrated DNA Technologies), 1 ul T4 Polynucleotide Kinase (New England

BioLabs (NEB), #M0201S), 1 ul T4 Polynucleotide Kinase Buffer with 10nM ATP (10x) (New England BioLabs, #B0202S) and 6 µl of ddH2O, ending up with a total volume of 10 µl. Incubation was carried out at 37 °C for 30 minutes, 95 °C for 5 minutes, and finally ramping down to 25 °C decreasing the temperature at 5 °C/min. The annealed gRNAs were diluted to 1:200 in ddH2O for subsequent ligation.

2.2.3. Ligation of gRNAs and plasmids

All ligation reactions were performed using 1 µl of T4 DNA Ligase Thermo Fisher Scientific 5 U/µl (EL0011), 2 µl of 10x T4 DNA Ligase Buffer (Thermo Fisher Scientific, B69), 50 ng of vector and a quantity of insert DNA to obtain a ratio 1:7 in a final volume of 20 µl. Incubation was done at 22°C for 1 hour.

2.2.4. Transformation

50 µl of Library Efficiency™ DH5α Competent Cells (Thermo Fisher Scientific, Cat. No. 18263012) were thawed on ice for 10 minutes, mixed with 5 µl of the respective ligation product and left on ice for 15 minutes. Heat-shock was done at 42 °C for 40 seconds followed by the addition of 450 µl of SOC 1x, and incubation at 37 °C shaking for 30 - 60 min. After that the bacteria were seeded in LB agar plates containing ampicillin and left at 37 °C overnight.

2.2.5. PCR assay of transformed bacterial colonies

The colony PCR was performed using 0.4 µl of 5 µM forward primer (Integrated DNA Technologies), 0.4 µl of 5 µM reverse primer (Integrated DNA Technologies) 0.4 µl of dNTPs mix 10 mM, 0.2 µl of Phire Hot Start II DNA Polymerase (Thermo Fisher Scientific, F122S), 4 µl of 5x Phire Green Reaction Buffer (Thermo Fisher Scientific, F527L), and ddH2O up to 20 µl. Bacterial colonies were added to the PCR mix. PCR was performed with an initial denaturation at 98 °C for 3 minutes, 25 cycles of denaturation at 98 °C for 10 seconds, annealing at 55 °C for 20 seconds, extension at 72 °C for 30 seconds, and a final extension at 72 °C for 5 minutes. Electrophoresis was prepared to run the PCR products using a 1 % agarose gel.

2.2.6. Amplification of transformed Bacteria and purification of plasmid DNA

The transformed bacteria were amplified using 200 ml of LB medium containing ampicillin with a dilution of 1:1000 and incubated on a shaker at 180 rpm, at 37 °C overnight. The DNA was purified using PureLink HiPure Plasmid Filter DNA purification

Kit (Invitrogen, MAN0003720, Cat. No. K210015). To confirm successful cloning the DNA was sequenced by GATC Biotech (Ebersberg, Germany).

2.2.7. Nucleofection and Cell Sorting

For each reaction $3*10^6$ BlaER cells in 100 µl transfection solution consisting of 82 µl Solution C and 18 µl of Supplement (AmaxaTM Cell Line NucelofectorTM Kit C, Cat. No. VCA-1004) were nucleofected using the Nucleofector® Device (Amaxa) with 2 µg of plasmid DNA in total, 1 µg containing upstream and 1 µg downstream binding gRNA, respectively. On the following day, cells transfected with the px459_PuroR plasmid were treated with puromycin (10 mg/ml) 1: 10,000. Cell sorting was performed on a BD Influx™ Cell Sorter by selecting for alive single mCherry positive cells on day 3 after nucleofection into p96-well plates.

2.2.8. Cell lysis and PCR screening

After amplification and colony growth of the single cells, a mirror plate was generated and genomic DNA was isolated according to Alt-R Genome Edition Detection Kit (Integrated DNA Technologies (IDT), Cat. No. 1075932) and amplified by PCR with an initial denaturation at 98 °C for 30 sec, 25 cycles of denaturation at 98 °C for 10 sec, annealing at a primer specific temperature around 60 °C and extension at 72 °C for 20 sec, with a final extension at 72 °C for 2 min. Next, the PCR product was visualized by gel electrophoresis to check for homozygote and heterozygote knockouts. Potential samples with a successful knockout were validated by sequencing (GATC ®).

2.3. Transdifferentiation and FACS analysis

Transdifferentiation of 300.000 cells/ml of BLaER cells into macrophages was induced by introducing human colony-stimulating factor 1 (hCsf-1) (10 ng/ul) and human interleukin-3 (hIL-3) (10 ng/ul) as well as 100nM β-estradiol to the cells. Specific time points were set at t = 0 h, t = 24 h, t = 48 h, t = 96 h and at t = 168 h to monitor the transdifferentiation into macrophages. Human FcBlock (Human BD Fc Block™, BD Pharmingen, Cat.564219, dilution 1/20), a blocking antibody, was applied followed by antibodies specific for macrophages, Mac-1 labeled with APC (APC Mouse Anti-Human CD11b/Mac-1 BD Pharmingen, Cat 550019, dilution 1/10), and for B cells, CD19 labeled with PE (PE Mouse Anti-Human CD19, BD Pharmingen, Cat 555413, dilution 1/10), to observe the conversion into macrophages. For FACS a BD LSR Fortessa Analyzer was used. DAPI (1 ug/ml) served as a viability marker. Cells were first gated by size and

single cells that were DAPI positive and GFP tagged were selected. Finally, they were discriminated against by APC and PE label corresponding to antibody Mac-1 specific for macrophages and CD19 for B cells. Data were analyzed with FlowJo software (Tree Star, Ashland, San Diego) and GraphPad Prism7 (GraphPad Software, San Diego) to perform statistical analysis using a student t-test.

2.4. Targeted Hi-C Library Preparation

2.4.1. Cross-linking

For targeted Hi-C Library Preparation $2*10^6$ BLaER KO cells were used for Hi-C with a concentration of $1*10^6$ cells/ml in a fresh culture medium. Cells were cross-linked with 37 % stock formaldehyde to a 1 – 2 % final concentration and incubated for 10 minutes at room temperature (RT). Glycine was added to a final concentration of 0.125 M for quenching followed by a 5 minutes incubation step under rotation and centrifugation at 300 xg for 5 min at 4 °C. After washing with ice-cold PBS, the pellet can be snap-freezed in liquid nitrogen and stored at -80 °C.

2.4.2. Lysis and restriction enzyme digestion

The pellet (2 – 5 million cells) was resuspended in 0.25 ml freshly prepared ice-cold Hi-C lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2 % Igepal CA630 and 1X Roche complete protease inhibitor), incubated for 15 minutes on ice and centrifuge at 1000 xg for 5 minutes at 4 °C. After washing with 500 µl Hi-C lysis buffer, the pellet was resuspended in 50 µl 0.5 % SDS in 1X NEBuffer 2 (NEB2) (NEB, B7002) and incubated at 62 °C for 10 minutes. For quenching of SDS 170 µl of 1X NEB2 buffer containing 100 % Triton X-100 was added, followed by incubation at 37 °C for 15 minutes. Then, 25 µl of 1X NEB2 buffer mixed to the samples by inverting, an aliquot as a control (8 µl) was taken and 100U MboI restriction enzyme (NEB, R0147M) was added to the remaining nuclei and digested for 2 hours at 37 °C under rotation. An additional shot of 100U MboI restriction enzyme (NEB, R0147M) was added and left incubating for 2 hours. The samples were left incubating overnight with a final dose of 100U MboI restriction enzyme (NEB, R0147M).

On the next day, 100U MboI restriction enzyme (NEB, R0147M) was mixed to the samples and incubated for 3 more hours. After 1 hour a control aliquot (8 µl) was taken. The de-crosslinking of all aliquots before and after digestion was done by adding 80 µl of TE buffer (10 mM Tris pH 8, 1 mM EDTA) and 10 µl Proteinase K (10 mg/ml) and

incubated at 65 °C for 1 hour. The control aliquots were visualized on a 0.6 % agarose gel, with the fragment majority ranging from 3 to 0.5 kb.

2.4.3. Biotin fill-in, proximity ligation and crosslink reversal

To inactivate MboI, the samples were heated to 65 °C for 20 minutes, cooled to RT, and centrifuged at 1000 xg at RT for 5 minutes. The pellet was dissolved in 250 μl fresh 1X NEB2 buffer and the restriction fragment overhangs were filled-in using 50 μl of a mix containing 37.5 µl of 0.4mM biotin-14-dATP (Life Technologies, 19524-016), 1.5 µl of 10mM dCTP (Thermo Fisher Scientific, Ref R0151), 1.5 µl of 10mM dGTP (Thermo Fisher Scientific, Ref R0161), 1.5 µl of 10mM dTTP (Thermo Fisher Scientific, R0171) and 8 µl of 5U/µl DNA Polymerase I, Large (Klenow) Fragment (NEB, M0210) and incubate at 37 °C for 1.5 hours under rotation. Next, add ligation mix consisting of 10X NEB T4 DNA ligase buffer (500 mM Tris-HCl pH 7.5, 100 mM MgCl2, 10 mM ATP, 100 mM DTT), 20 mg/ml Bovine Serum Albumin (100X BSA molecular biology grade, NEB, B9000S), 2000 U/µl T4 DNA Ligase (NEB, M0202M) = 10000 CEU (= ± 30 Weiss Units) and water adding up to 900 ul. Incubation at RT overnight with slow rotation followed. RNAs and proteins were degraded through 50 µl proteinase K (20 mg/ml) and 10 µl RNAse A (10 mg/ml) (Thermo Fisher Scientific, #EN0531), incubation at 55°C for 45 – 60 minutes and then at 65 °C for 4 hours.

2.4.4. DNA Shearing and Size Selection

First, the tubes were cooled to RT, then each sample was split into three 400 μl aliquots followed by adding glycogen (20 mg/ml) (Thermo Fisher Scientific, Ref R0561), 0.1X volumes of 3M sodium acetate (pH 5.2), and 2.5X volumes of 100 % ethanol. The tubes were mixed by inverting and incubated at -80 °C for 45 – 60 minutes. After the tubes thawed, they were centrifuged at maximum speed at 4 °C for 25 minutes and kept on ice afterward. The pellets of the three aliquots were washed in 70 % ethanol and pooled together. Following an additional washing step with 70 % ethanol and air-drying, the pellet was dissolved in 130 μl 1X Tris buffer (10 mM Tris-HCl, pH 8) and incubated at 37 °C for 15 minutes. The yield was measured using Qubit dsDNA HS Assay Kit (Invitrogen by Thermo Fisher Scientific, Ref. Q32851) and the ligation was controlled by running about 200 ng of Hi-C product on a 0.6 % agarose gel with an expected fragment size of more than 3 kb.

Next, the samples were diluted in a 0.65 ml Diagenode tube to 10 ng/ul in 1X Tris buffer with 1 μg per tube. The biotinylated DNA was sheared to a size of 150 – 700 bp to make it suitable or high-throughput sequencing using Illumina sequencers. Shearing was

performed on the Bioruptor Pico with 6 – 8 cycles (7 cycles) of 20'' ON and 60'' OFF. The sheared DNA was transferred to normal safe-lock tubes and 1.8X of AMPure XP beads (Beckmann Coulter, Ref A63881) were added and resuspended gently. After a 5-minute incubation period at RT, the beads were collected using a magnet and washed two times with freshly prepared 80 % ethanol while keeping them on the magnet. The beads were briefly air-dried, and the DNA was eluted using 300 µl 1X Tris buffer (10 mM Tris-HCl, pH 8), incubating for 2 – 5 minutes at RT and collecting the supernatant by gathering the beads with a magnet. Finally, the yield was measured by Qubit and the samples were analyzed on a 1.5 % agarose gel.

2.4.5. Biotin Pull-Down and Preparation for Illumina Sequencing

First, 2X Binding & Washing Buffer (BWB) containing 10 mM Tris-HCl, pH 7.4; 1 mM EDTA and 2 M NaCl was prepared and the 1X BWB by diluting 1:2. The samples were prepared for biotin pull-down by washing per sample 150 µl of 10mg/ml Dynabeads MyOne Streptavidin C1 beads (Life Technologies, # 65001) twice with 800 µl of 1X BWB and gathering the beads with a magnet. After resuspending the beads in 300 µl of 2X BWB, they were transferred to the Hi-C sample, incubated for 20 minutes at RT under rotation, and then again separated on the magnet to discard the solution to bind the biotinylated DNA to the streptavidin beads. The beads were washed again in 600 µl of 1X BWB as before, then resuspended in 200 µl 1X NEBuffer 2 (New England BioLabs (NEB), B7002S), transferred into a new tube and reclaimed again by discarding the solution. A 100 µl End Repair master mix (NEBNext DNA Library prep kit (NEB, E6040)) containing 10X NEBNext End Repair Reaction Buffer, 5 µl NEBNext End Repair Enzyme Mix and water was prepared to resuspend the beads in order to repair the sheared ends of the DNA and remove biotin from unligated ends. Following incubation at 37 °C for 30 minutes, the beads were washed again in 600 µl 1X BWB and reclaimed, resuspended in 200 µl 1X NEBuffer 2 (New England BioLabs (NEB), B7002S), transferred to a new tube and again reclaimed.

Then, the beads were resuspended in 50 µl of dA-Tailing master mix (NEBNext DNA Library prep kit (NEB, E6040)) containing 10X NEBNext dA-Tailing Reaction Buffer, 3 µl DNA Polymerase I Ig (Klenow) Fragment (NEB, M0212) and water. Following incubation at 37 °C for 30 minutes, the beads were washed again in 600 µl 1X BWB and reclaimed, resuspended in 200 µl 1X NEB T4 DNA Ligation buffer, transferred to a new tube and again reclaimed.

In the next step, the samples were resuspended in 50 µl of Adaptor Ligation master mix (NEBNext DNA Library prep kit (NEB # E6040)) with reduced Adaptor/Ligase

concentrations containing 5X Quick Ligation Reaction Buffer, 3 μl NEBNext Adaptor (NEB, #E7337A), 2 μl Quick T4 DNA Ligase and water. Following incubation at RT for 15 minutes, 3 μl of USER enzyme (NEB, #E7338A) were added with a continued incubation period of 15 minutes at 37 °C.

The beads were separated on a magnet and the solution discarded. The beads were resuspended in 100 μl 1X Tris buffer (10 mM Tris-HCl, pH 8), transferred into a new tube, and reclaimed again. Finally, the beads were resuspended in 50 μl 1X Tris buffer (10 mM Tris-HCl, pH 8).

2.4.6. Whole-genome chromatin interaction library amplification and purification

The Hi-C library samples were amplified by PCR in duplicates and optimization of PCR cycle number was done previously by testing for 6, 8, 10 and 12 cycles (initial denaturation at 98 °C for 2 minutes, 8 cycles of denaturation at 98 °C for 20 seconds - annealing at 65 °C for 30 seconds – extension at 72 °C for 45 seconds, and a final extension at 72 °C for 3 minutes). The PCR mix contained 10 μl of Adapter ligated library-on beads, 3.75 μl of 25 μM NEB Universal primer (NEB Next Multiplex Oligos, (NEB, #E73352)), 3.75 μl of Index primer (NEB Next Multiplex Oligos, (NEB, #E73352)), 10 μl of 5x Herculase II Reaction Buffer (Agilent, Cat. No. 600675-52), 1.25 μl of 10 mM dNTPs (NEB, #N0447L), 2 μl of Herculase II Fusion DNA Polymerase (Agilent, Cat. No. 600675-51) and PCR graded water adding up to 50 μl. For each sample, a different index primer was used.

The PCR products for each library were pooled and purified with 1X AMPure XP bead (Beckmann Coulter, Ref A6388). The yield was measured on the Qubit (Qubit dsDNA HS Assay Kit (Ref. Q32851, Invitrogen by Thermo Fisher Scientific)) and fragment size was estimated based on a 1.5 % agarose gel for the calculation of the library DNA molarity. Finally, the final Hi-C products were sequenced on the NextSeq500 for library validation.

2.4.7. In situ Hi-C processing and normalization

Hi-C data were processed using an in-house pipeline based on TADbit (Serra *et al.*, 2017).

The quality of reads was checked by FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to discard problematic samples and systemic artifacts. Trimmomatic (Bolger, Lohse and Usadel, 2014) with the

recommended parameters for paired-end reads was used to remove adapter sequences and poor quality reads (ILLUMINACLIP:TruSeq3-PE.fa:2:30:12:1:true; LEADING:3; TRAILING:3; MAXINFO:targetLength:0.999; and MINLEN:36). For mapping, a fragment-based strategy as implemented in TADbit was used. Briefly, each side of the sequenced read was mapped in full length to the reference genome (hg38, Dec 2017 GRCh38). In following steps, not uniquely mapped reads, were assumed chimeric due to ligation of several DNA fragments. Next ligation sites were searched for, to discard those reads in which no ligation site was found. Remaining reads were split as often as ligation sites were detected. Individual split read fragments were then mapped independently. These steps were repeated for each read in the input FASTQ files. Multiple fragments from a single uniquely mapped read will result in as many contacts as possible pairs can be made between the fragments. For instance, if a single read was mapped through three fragments, a total of three contacts (all-versus-all) was represented in the final contact matrix. The TADbit filtering module was used to remove non-informative contacts and to create contact matrices. The different categories of filtered reads applied are:

- self-circle: reads coming from a single restriction enzyme (RE) fragment and point to the outside.
- dangling-end: reads coming from a single RE fragment and point to the inside.
- error: reads coming from a single RE fragment and point in the same direction.
- extra dangling-end: reads coming from different RE fragments but are close enough and point to the inside. The distance threshold used was left to 500 bp (default), which is between percentile 95 and 99 of average fragment lengths.
- duplicated: the combination of the start positions and directions of the reads was repeated, pointing at a PCR artifact. This filter only removed extra copies of the original pair.
- random breaks: the start position of one of the reads was too far from RE cutting site, possibly due to non-canonical enzymatic activity or random physical breaks. The threshold was set to 750 bp (default), > percentile 99.9. From the resulting contact matrices, low-quality bins (those presenting low contacts numbers) were removed as implemented in TADbit's "filter_columns" routine. The matrices obtained were normalized for sequencing depth and genomic biases using OneD (Vidal *et al.*, 2018). Next, they were further normalized for local coverage within the region (expressed as normalized counts per thousand within the region) without any correction for the diagonal decay. Finally, the resulting normalized matrices were directly subtracted from each other for differential analysis.

2.4.8. Identification of subnuclear compartments and topologically associated domains (TADs)

To segment the genome into A/B compartments, normalized Hi-C matrices at 100kb resolution were corrected for decay as previously published. Diagonals were grouped in case the signal-to-noise ratio was below 0.05. Corrected matrices were then split into chromosomal matrices and transformed into correlation matrices using the Pearson product-moment correlation. The first component of a principal component analysis (PCA) (PC1) on each of these matrices was used as a quantitative measure of compartmentalization. H3K4Me2 ChIP-seq data allowed to assign negative and positive PC1 categories to the correct compartments. If necessary, the sign of PC1 (randomly assigned) was inverted for the positive PC1 values to correspond to A compartment regions and vice versa for the B compartments.

2.5. RNA isolation, RT-PCR and qPCR

RNA was extracted from 10^7 * 10 knockout cells in two replicates 24h post induction of transdifferentiation and uninduced with a RNeasy Mini Kit (Qiagen, Cat. No. 7401 and 74106) and the yield was measured with a Nanodrop spectrophotometer.
cDNA was first generated using a High Capacity RNA to cDNA Kit (Applied Biosystems, Cat. No. 4387406). 800 ng of RNA were taken, and a reaction mix was prepared with and without 20X RT Enzyme Mix per sample. Reverse transcription was performed at 37 °C for 60 min followed by 95 °C for 5 min and a cooldown to 4 °C.
1 µl of 1:5 diluted cDNA was used for qPCR with 5 µl Power SYBR Green PCR Master Mix and 1 µl of 5 µM forward and reverse primer, respectively, and ddH2O to reach a total volume of 10 µl (Applied Biosystems, Ref. 4367659). Oligonucleotide sequences are indicated in Table 1. qPCR was performed in triplicate reactions on a ViiA7 by life technologies (Applied Biosystems) and analyzed with GraphPad Prism7 (GraphPad Software, San Diego).

Table 1 – Sequences of oligonucleotides used for q-RT-PCR.

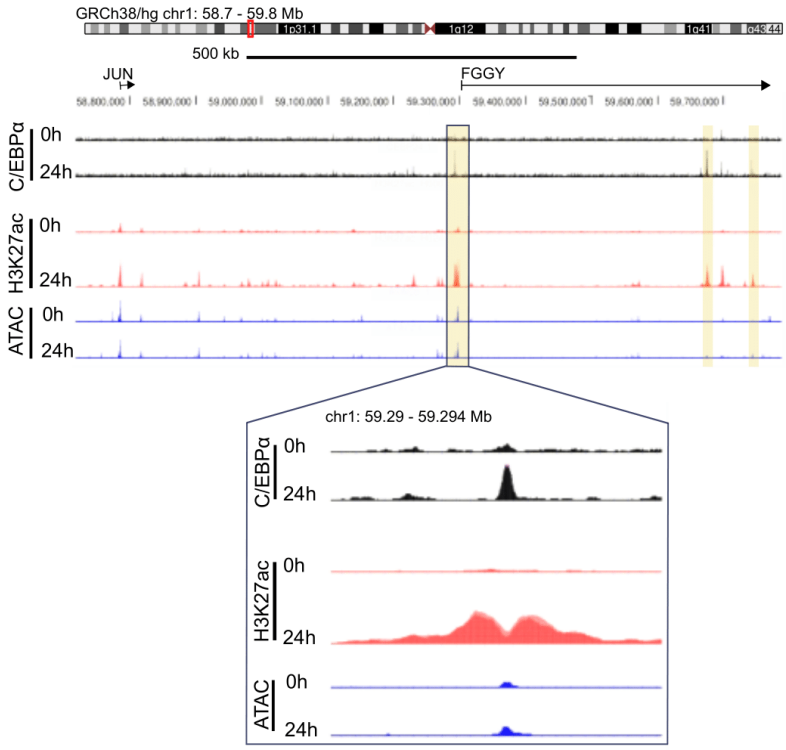
Oligonucleotide	Orientation	Sequence
JUN	forward	AGATGGAAACGACCTTCTATGAC
	reverse	GTCATGCTCTGTTTCAGGATCT
FGGY	forward	CCTGGAGCTTCTCTTGGAAAT
	reverse	TCCTGCATGGGCATCAAT
CSF1R	forward	TCCAAAACACGGGGACCTATC
	reverse	CGGGCAGGGTCTTTGACATA
hPU.1	forward	GAAGACCTGGTGCCCTATGA
	reverse	GGGGTGGAAGTCCCAGTAAT
hEBF1	forward	TGCTACTCCCTGTATGAAAG
	reverse	ATGGTACCGAATATGACCTG
hPGK1	forward	CTGGGCAaGGATGTTCTGTT
	reverse	CACATGAAAGCGGAGGTTCT
hGAPDH	forward	CAGCCTCAAGATCATCAGCA
	reverse	TGTGGTCATGAGTCCTTCCA

3. Results

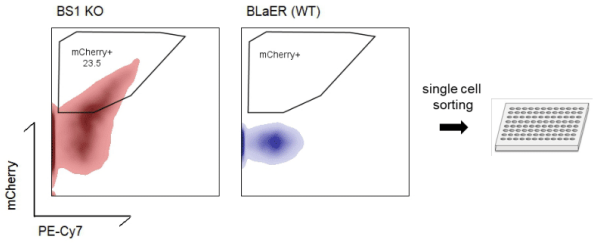
3.1. Generation of knockout cell lines

To investigate the ability of C/EBPα to induce a B to A compartment switch during transdifferentiation of human B cells (BLaER cell line) into macrophages (**Figure 5**), this study was focused on a well-characterized genomic region at the *JUN* locus. This region, bound by C/EBPα about 500 kb downstream of the *JUN* locus is concomitantly switched from B to A compartment during transdifferentiation, resulting in the activation of JUN expression. The binding-site undergoes epigenomic alterations after C/EBPα binding with an increase of chromatin accessibility (ATAC-seq) and acetylation of histones (H3K27ac) correlating with active chromatin marks (**Figure 6a**). First, to employ CRISPR technology to remove the C/EBPα binding-site, gRNA flanking the binding-site were designed and cloned into plasmids containing either a mCherry fluorescence marker or a puromycin resistance. In order to increase editing-efficiency, three upstream and downstream binding gRNAs were designed, cloned, and consequently tested to select the best performing one. The two most efficient gRNA-Cas9 plasmids were nucleofected into BLaER cells, shortly cultured in medium containing puromycin and single-cell sorted to select for cell colonies carrying a homozygote deletion (**Figure 6b**). Following the PCR screen to identify successful knockout cells 96 % of the colonies were still wildtype cells, 3 % were heterozygotes and only 1 % was a homozygote knockout, resulting in two bands at 600 bp when analyzed by gel electrophoresis (**Figure 6c-d**).

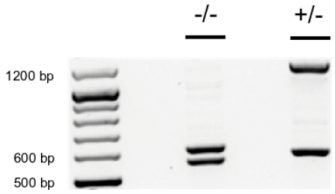
a



b



d



c

region	type	count
binding-site 1	WT [+/+]	90 (96%)
	heterozygote KO [+/-]	3 (3%)
	homozygote KO [-/-]	1 (1%)

Figure 6 – Establishing knockout cell lines. (**a**) Identifying C/EBPα binding-motifs. Three C/EBPα binding-sites were detected on chromosome 1 between JUN and FGGY. This was achieved through previously generated C/EBPα and H3K27ac ChIP-seq as well as ATAC-seq data during transdifferentiation at 0h and 24h after induction. Open chromatin regions (ATAC-seq) with active enhancers (H3K27ac) marks and C/EBPα binding motifs were pinpointed. Their location was 500 bp, 890 bp, and 950 bp upstream of JUN. Binding-site 1 (500 bp upstream of JUN) was targeted with gRNAs that cut upstream and downstream of binding-site 1. Puromycin or a mCherry marker served for selection. (**b**) Single-cell sorting of nucleofected BLaER cells. (Left panel) Nucleofected BLaER cells with a high-intensity mCherry fluorescence signal were sorted into 96-well plates. (Right panel) Wildtype BLaER cells were single-cell sorted as control displaying no red fluorescence signal. (**c**) Cell count based on PCR screen. Following a PCR screen of the single-cell sorted cells 96 % were still wildtype, while 3 % heterozygote and only 1 % homozygote knockouts were identified. This reflects a low CRISPR editing efficiency. (**d**) Confirmation of knockouts. Heterozygotes (+/-) and homozygotes (-/-) were validated by gel electrophoresis resulting in one band at 1200 bp and 600 bp for the heterozygote or two bands at 600 bp for the homozygote.

3.2. Transdifferentiation of knockout cells

Once the knockout cell line (BS KO) was obtained, the impact of the deletion of C/EBPα binding-site on transdifferentiation was assessed. Knockout and wildtype cell lines were then treated with β-estradiol to induce transdifferentiation and cell surface marker changes were monitored at 0h, 24h, 48, 96h, and 168h (fully transdifferentiated) (**Figure 7a**). In general, wildtype and likewise knockout cells exhibit a reciprocal regulation of CD19, a B cell marker, and Mac-1, a macrophage marker during cell conversion. CD19 is downregulated while Mac-1 is upregulated during transdifferentiation. However, the FACS analysis revealed differences between knockout and wildtype, with the transdifferentiation of knockout cells slightly accelerated already at 48h after induction compared to the wildtype cells. This continued throughout the whole process (**Figure 7a**). Indeed, the kinetics of CD19 during transdifferentiation revealed significantly reduced CD19 positive cells at about 40 % in the knockout compared to almost 80 % in the wildtype at 48h (**Figure 7b**). After 96h, the CD19 percentage approximated 0 in the knockout whereas in the wildtype with about 8 % exhibiting a significant difference. Likewise, at 168h once fully transdifferentiated the concentrations of CD19 continued to be at the same levels showing significantly decreased levels in the knockout compared to the wildtype (**Figure 7b**). Correspondingly, the percentage of Mac-1 positive cells was significantly increased in the knockout in comparison to the wildtype at 48h, 96h, and 168h (**Figure 7b**). Altogether, this indicates a faster speed of transdifferentiation in the knockout cell line evaluated against the wildtype. Here, CD19 is earlier downregulated and Mac-1 more rapidly upregulated.

3.3. Gene expression changes in the knockout cell line

The following step was to study the impact of the binding-site knockout on B cell and myeloid gene expression in addition to the expression of the genes JUN and FGGY during transdifferentiation. RNA expression of JUN, FGGY, PU1, CSF1R, EBF1 was then analyzed by q-RT-PCR at 24 hours post-induction (**Figure 7c**). The gene expression of the knockout cell line was examined in contrast to the wildtype in the previously mentioned genes of interest after 24h following induction (**Figure 7c**). During cell conversion, the JUN expression in the knockout and wildtype increased. Yet after 24h, the wildtype displayed by about two-times more elevated JUN mRNA levels than the knockout. Though, the control genes (CSF1R, PU.1, and EBF1) determining proper transdifferentiation were more pronounced up- and downregulated, respectively, in the knockout contrary to the wildtype (**Figure 7c**). In conclusion, the altered gene expression agreed with the results obtained by FACS analysis emphasizing a more rapid transdifferentiation in the knockout over the wildtype.

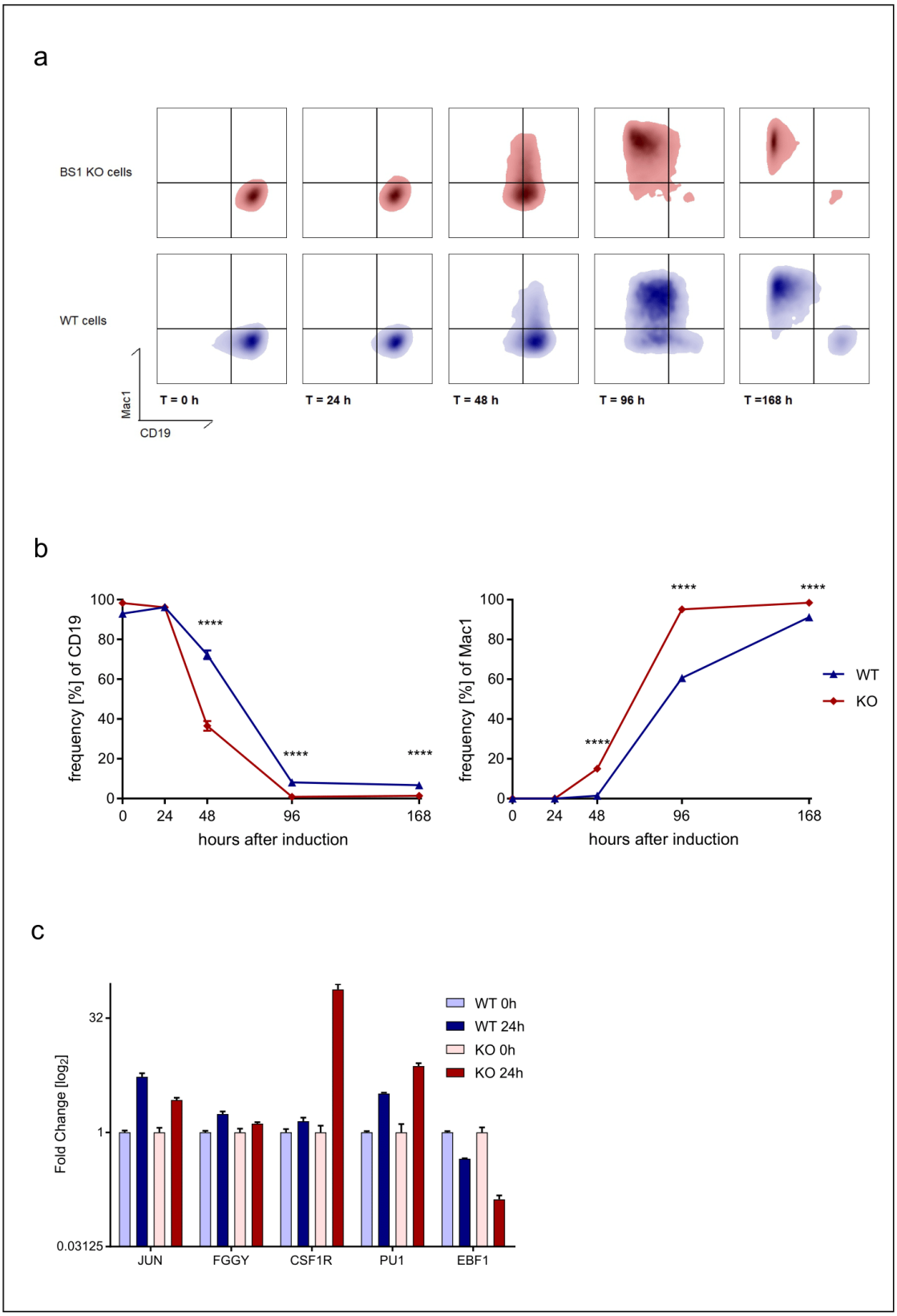

a
BS1 KO cells
WT cells
Mac1
CD19
T = 0 h
T = 24 h
T = 48 h
T = 96 h
T =168 h
b
frequency [%] of CD19
hours after induction
****
frequency [%] of Mac1
WT
KO
c
Fold Change [log2]
32
1
0.03125
WT 0h
WT 24h
KO 0h
KO 24h
JUN
FGGY
CSF1R
PU1
EBF1

Figure 7 – Transdifferentiation and gene expression of the knockout. (**a**) FACS plot of transdifferentiating BS1 KO. The knockout (KO) (red) was analyzed during transdifferentiation with FACS compared to the wildtype (WT) (blue). Antibodies specific for CD19, a B cell marker, and Mac1, a macrophage marker were used to track the process of cell conversion. (**b**) Kinetics of CD19 and Mac1. (Left Panel) The concentration of CD19 during transdifferentiation in the KO (red) compared to the wildtype (blue) were analyzed at the indicated timepoints. (Right Panel) Mac1 concentration of KO and WT during cell conversion. Statistical analysis by student t-test, ****p<0.0001. (**c**) Comparison of gene expression in WT and KO. Gene expression of WT (light blue, blue) and KO (light red, red) were compared at 0h and 24h post induction. The WT was previously analyzed in the laboratory by RNA-seq.

3.4. Hi-C and change in genome topology induced by C/EBPα BS KO

As a next step, Hi-C was performed to allow mapping of chromatin interactions genome-wide (Lieberman-Aiden, 2009) and to determine the effects of the deletion of the binding-site on the genome architecture. Specifically, compartmentalization and genome topology of the region around the deletion was of interest (**Figure 8a**). First, the Hi-C metrics and quality were evaluated. Just below 70% of normal paired reads were obtained (**Figure 8b**). Normal paired reads correspond to reads that were linked to a chimeric fragment consisting of two fused chromatin regions that are in close spatial proximity to each other in the nucleus. Conversely, aberrant reads represent low percentages, with about 4% and 27% of duplicates and dangling ends, respectively. The normal paired reads can be further distinguished into trans and cis, the latter corresponding to cross-ligated fragments of the same chromosome and trans describes cross-ligations of fragments of different chromosomes. (Hansen et al., 2019). Thus, cis intrachromosomal interactions represent 70% of the interactions (**Figure 8b**). Altogether, this indicates a good Hi-C quality and library as a high yield of normal paired reads was obtained and a high cis/trans ratio was reached.

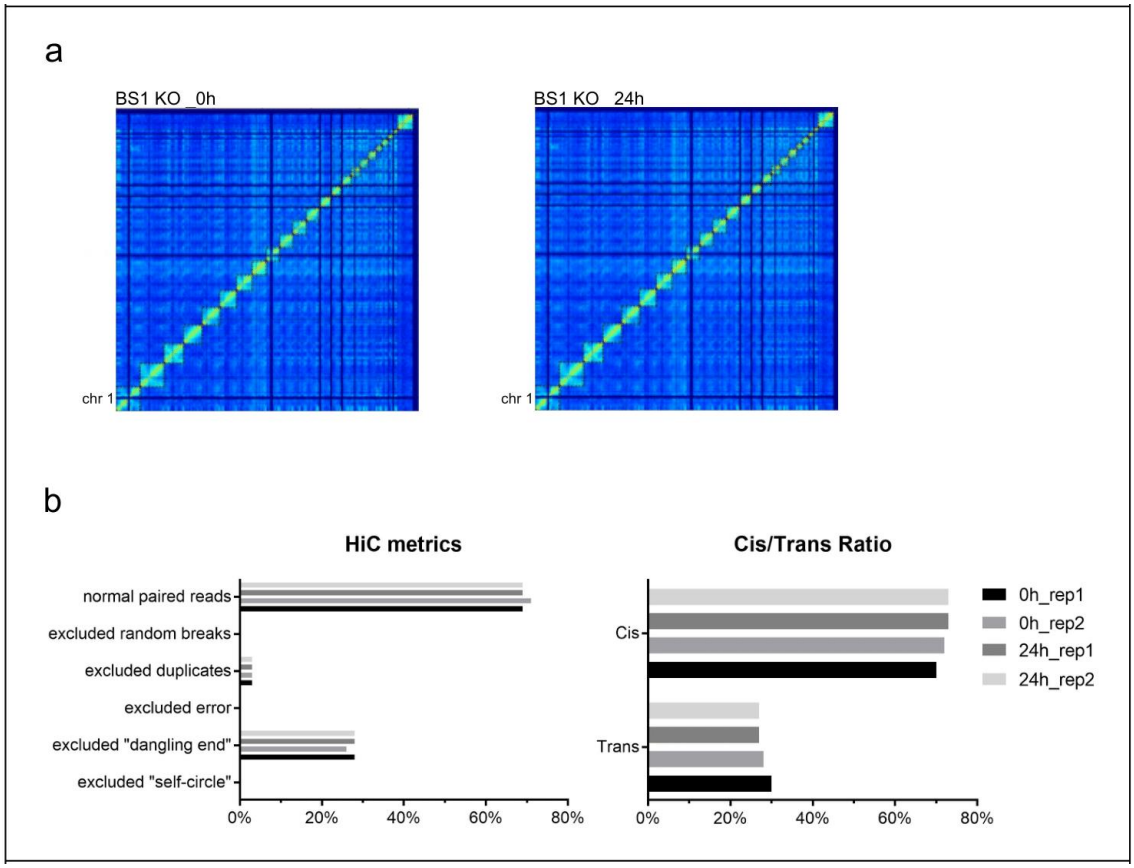


Figure 8 – Hi-C knockout experiment metrics and quality. (**a**) Whole-genome Hi-C heat map of BS1 KO at 0h and 24h following induction of transdifferentiation. Each square domain corresponds to one chromosome. The location of chromosome 1 was indicated. (**b**) Hi-C metrics (left panel). Indicated are normal paired reads for the two 0h-replicates and 24h-replicates of BS1 KO. Random breaks, duplicates, errors, dangling ends, and self-circles were excluded as they are uninformative. Self-circle, reads originating from a circular ligated fragment. dangling-end, a digested but unligated fragment. error, reads from a single restriction fragment pointing in the same direction. duplicates, identical PCR artifacts. random breaks, start position is located too far from restriction enzyme cutting site. (Right panel) Cis/Trans ratio. The percentage of cis and trans reads for each 0h and 24h replicate (rep1/rep2) was determined. Reads in cis map to the same chromosome and in trans to a different chromosome. A low cis/trans ratio points towards abundant random cross-ligation (Hansen *et al.*, 2019).

3.5. Compartmentalization and chromatin topology in knockout cells

Hi-C data were then analyzed to determine genome segmentation into A and B compartments. Quantitative changes in the A–B-compartment association (based on the PC1 values of a PCA on the Hi-C correlation matrix) during the first 24h of transdifferentiation were measured. First, a principal component analysis (PCA) of the compartment data was conducted. Here, principal component 1 (PC1) explained 39.2% of the variability and principal component 2 (PC2) 28.6% (**Figure 9a**). PC1 appeared to reflect transdifferentiation of a B cell into an induced macrophage from uninduced at 0h to 24h after induction. Here, these differences between uninduced (0h) and induced (24h) appeared to be more pronounced within the knockout compared to the wildtype. Though, PC2 could discriminate between wildtype and knockout. This observation can be explained by the fact that the Hi-C experiment for the wildtype was conducted previously in the laboratory with a deeper sequencing depth than the Hi-C experiment of the knockout.

Next, to uncover whether the initial hypothesis of stopping the compartment switch from B to A through the knockout of C/EBPα binding-site was achieved, the compartmentalization during transdifferentiation from 0h to 24h after induction was analyzed and visualized in the region between the *JUN* and *FGGY* locus in the knockout compared to the wildtype (**Figure 9b**). The JUN locus in the wildtype, as previously discussed, switched from an inactive B compartment (blue) to an active A compartment (yellow). Likewise, in the knockout, the *JUN* locus also turned into an A compartment after 24h. Thus, posing the question of whether the binding of another protein could be responsible. A region further upstream located within the *FGGY* locus appeared to be switching into an A compartment as well. Interestingly, it contains an additional binding-site of C/EBPα (about 890 kb upstream of JUN) (**Figure 9b**) that could compensate for the knockout.

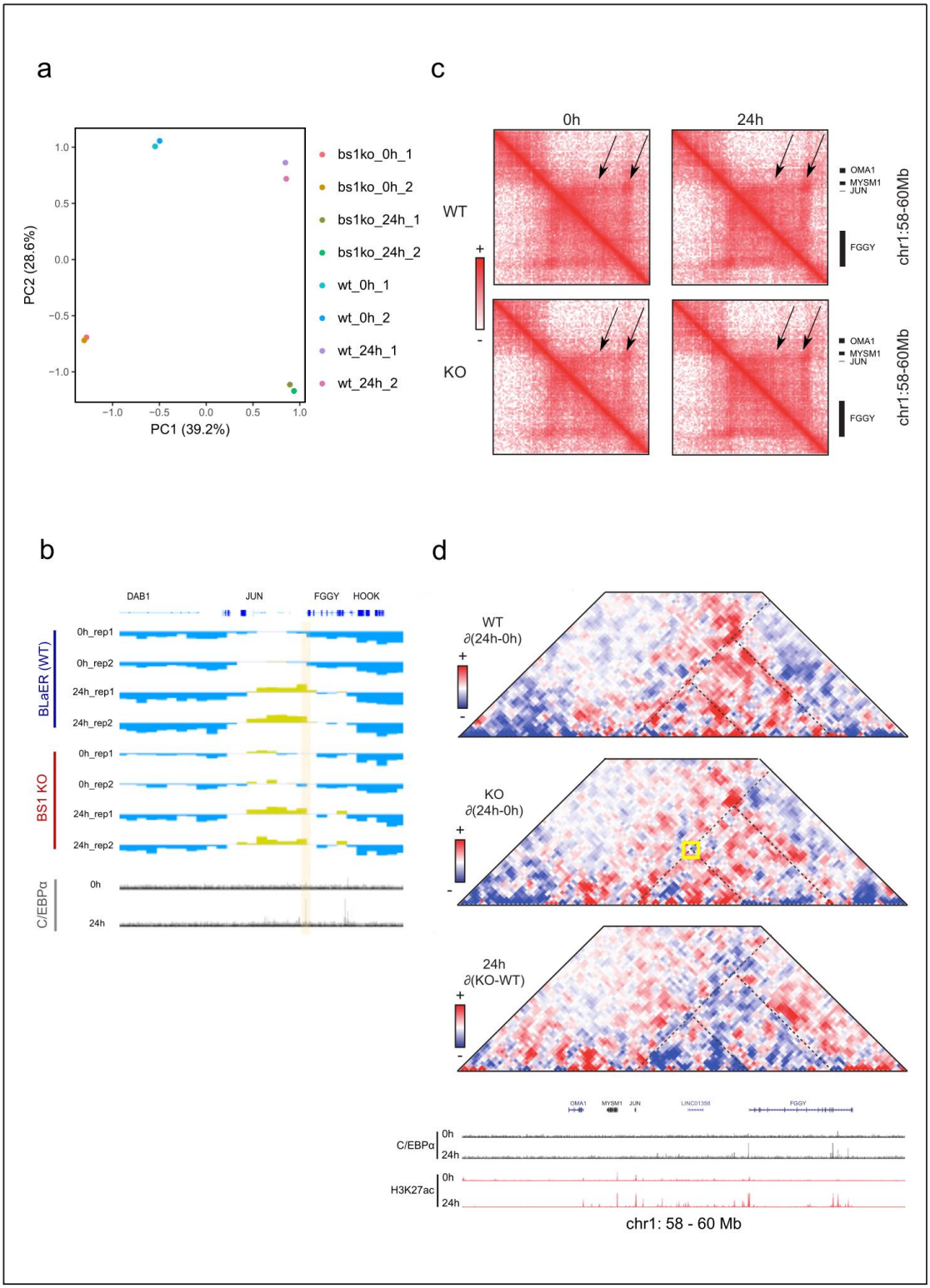

a
PC2 (28.6%)
PC1 (39.2%)
bs1ko_0h_1
bs1ko_0h_2
bs1ko_24h_1
bs1ko_24h_2
wt_0h_1
wt_0h_2
wt_24h_1
wt_24h_2
c
0h
24h
WT
KO
OMA1
MYSM1
JUN
FGGY
chr1:58-60Mb
b
DAB1
JUN
FGGY
HOOK
BLaER (WT)
BS1 KO
0h_rep1
0h_rep2
24h_rep1
24h_rep2
C/EBPα
0h
24h
d
WT
∂(24h-0h)
KO
∂(24h-0h)
24h
∂(KO-WT)
OMA1
MYSM1
JUN
LINC01358
FGGY
C/EBPα
H3K27ac
chr1: 58 - 60 Mb

Figure 9 – Identification of compartments and loop domains in the knockout. (**a**) PCA analysis of compartment data. A PCA analysis of wildtype (WT) and knockout (KO) at 0h and 24h after induction (in two replicates) was performed. PC1 (39.2 %) differentiated between uninduced and induced state within the WT and KO. Additionally, the WT Hi-C experiment was conducted previously in the laboratory with a different sequencing depth than for the KO Hi-C experiment. Thus, PC2 (28.6 %) distinguished between these two experiments. (**b**) Compartmentalization of WT and KO. The compartmentalization of WT (blue) and KO (red) at 0h and 24h after induction at JUN was visualized. The C/EBPα ChIP-seq data (grey) was layered below indicating the deleted binding-site 1 (yellow bar) in the knockout. The A compartment was displayed in yellow and the B compartment in blue. (**c**) Hi-C heat maps of targeted regions. Hi-C maps were prepared for the WT and KO at 0h and 24h for the region located on chromosome 1 between JUN and FGGY. Interactions were indicated in red. (**d**) Difference map. Difference maps were created for WT (top) and KO (middle) (δ(24h-0h)) to visualize changes that occurred during transdifferentiation. Yellow square indicates the interaction, lost due to the deletion. (Bottom) Difference map to view alterations between KO and WT. The C/EBPα (blue) and H3K27ac (red) ChIP-seq data and gene locations were indicated below. For better visualization, the dotted line specifies the interacting loci.

The chromatin organization at the JUN locus was then examined in further detail in the knockout and wildtype. When comparing the Hi-C maps of the wildtype uninduced (0h) and at 24h after induction of transdifferentiation, slight changes can be observed with a characteristic punctuated interaction signal between JUN locus and C/EBPα binding-site (**Figure 9c**). Additionally, a more pronounced punctuate interaction signal appeared at 24h between the JUN locus and the additional C/EBPα binding site at the 3’ end of FGGY. In contrast, the knockout Hi-C map showed no gain in interactions between JUN and the region of the deleted CEBP/α binding-site. However, contact between JUN and the additional C/EBPα binding site at the 3’ end of FGGY still occurred (**Figure 9c**).

These observations can be visualized in more detail through Hi-C differential maps to pinpoint any topological alterations. Here, the 0h was subtracted from 24h map for wildtype, and knockout and additionally, the 24h wildtype was subtracted from the 24h knockout map (**Figure 9d**). The dotted line indicates interacting genomic loci resulting in a punctuate signal. The wildtype difference map clearly demonstrated the interaction of JUN with the C/EBPα binding site at the 5’ end as well as the 3’ end of FGGY resulting in two loop domains. Accordingly, the knockout presented only one loop structure stemming from JUN and the C/EBPα binding-site at the 3’ end of FGGY. Overall, the 24h difference map displayed fewer interactions in the region of interest (**Figure 9d**), indicating that the topology was slightly impacted by the knockout. This appeared to be not enough to abolish the compartment switch from B to A.

4. Discussion

C/EBPα is a lineage-instructive transcription factor inducing the transdifferentiation of B cells into induced macrophages, through alteration of the chromatin architecture specifically the compartmentalization. In particular, C/EBPα appeared to play a significant role by potentially binding to inactive B compartments and converting them into active A compartments (Stik *et al.*, 2020). This suggests that lineage-instructive transcription factors have an additional up until now unidentified function.

To examine whether C/EBPα is responsible for the switch from B to A compartment during transdifferentiation (**Figure 6**) a C/EBPα binding-motif was destroyed. The generation of the cell line containing the deletion of C/EBPα binding-site (500 kb upstream of JUN) was done using two plasmids expressing gRNA but also a mCherry or puromycin marker for selection. However, the low efficiency (1% of the cell population were homozygote knockouts and 3 % heterozygotes (**Figure 6c**) indicates that the puromycin selection was not optimal. The plasmid containing the puromycin resistance gene could have been replaced by another fluorescent protein (e.g BFP since GFP was already expressed in the BLaER cells) to increase the editing efficiency.
Analysis of transdifferentiation kinetics by FACS suggested a faster speed of transdifferentiation for the knockout than for the wildtype cells. The analysis of the RNA expressed by the knockout and wildtype cells during transdifferentiation confirm also the acceleration of the process for the knockout with an increase of myeloid genes (CSF1R, PU1) and a decrease of the B cell gene EBF (**Figure 7c**). Though, the quantification of the wildtype mRNA was done previously in the laboratory by RNA-seq while the quantification of the knockout cell line was done by qPCR, which does not allow for optimal comparison. Although studies support a high correlation between qPCR and RNA-seq results (Everaert *et al.*, 2017) (Li *et al.*, 2019), this observation should be confirmed employing a similar technique for the 2 conditions. Anyway, JUN expression was impaired after 24h in the knockout cells in comparison to the wildtype. Interestingly, a study by Rangatia *et al*. showed that an elevated c-Jun concentration causes the inhibition of C/EBPα (Rangatia *et al.*, 2003). Thus, the downregulation of JUN in the knockout cells could have increased the activity of C/EBPα and then accelerated the expression of myeloid genes. Nonetheless, further studies need to be conducted to identify the exact regulation of these transcription factors.

Finally, to uncover the compartmentalization and topology of the knockout, Hi-C was performed. The Hi-C quality metrics were good (**Figure 8**) and will allow future deeper sequencing of the libraries. During the first 24h of transdifferentiation, very few quantitative alterations linked to the association with A or B compartment were assessed, as shown in the principal component analysis of the chromatin compartment data (**Figure 9a**). However, extensive analysis of the qualitative and quantitative change of genome compartmentalization should be realized genome-wide to measure the impact of the knockout on the entire 3D organization of the nucleus.

Our analysis focused on the compartmentalization of the JUN locus as further analyzed in the wildtype and the knockout. The results clearly demonstrated the B to A compartment switch in the knockout was not abolished through the destruction of the C/EBPα binding-site. Interestingly, a smaller region located within FGGY towards the 3' end turned into an A compartment and displayed two additional C/EBPα binding-sites (**Figure 9b**). When further analyzed at a higher resolution, chromatin contacts were detected between this additional C/EBPα binding-site, the binding-site of C/EBPα that was targeted, and the *JUN* locus. While the contacts between the *JUN* locus and the region around the C/EBPα targeted binding-site disappeared in the knockout, contacts between JUN and additional C/EBPα (close to FGGY) remain. Thus, the loss of contact between the targeted C/EBPα binding-site and the *JUN* locus may explain the decrease of JUN expression, but it cannot be excluded that these additional binding sites may participate in the compartmentalization of the JUN locus and somehow compensate the knockout. In line with this, it was discovered that the entire region between JUN and the 3' end of FGGY also switches in a later stage of transdifferentiation (Stik *et al.*, 2020). The current hypothesis is that the two additional C/EBPα motifs rescued the compartmentalization of JUN (**Figure 10**). This was indicated by the compartment analysis displaying an extra small A compartment domain at the 3' end of FGGY with its two C/EBPα motifs (**Figure 9b**). Also, the Hi-C map and difference map revealed a loop domain that was unaffected by the knockout (**Figure 9c-d**). It will be interesting to destroy the two additional binding-sites separately as well as all three of them at once to identify whether C/EBPα is responsible for the compartment switch.

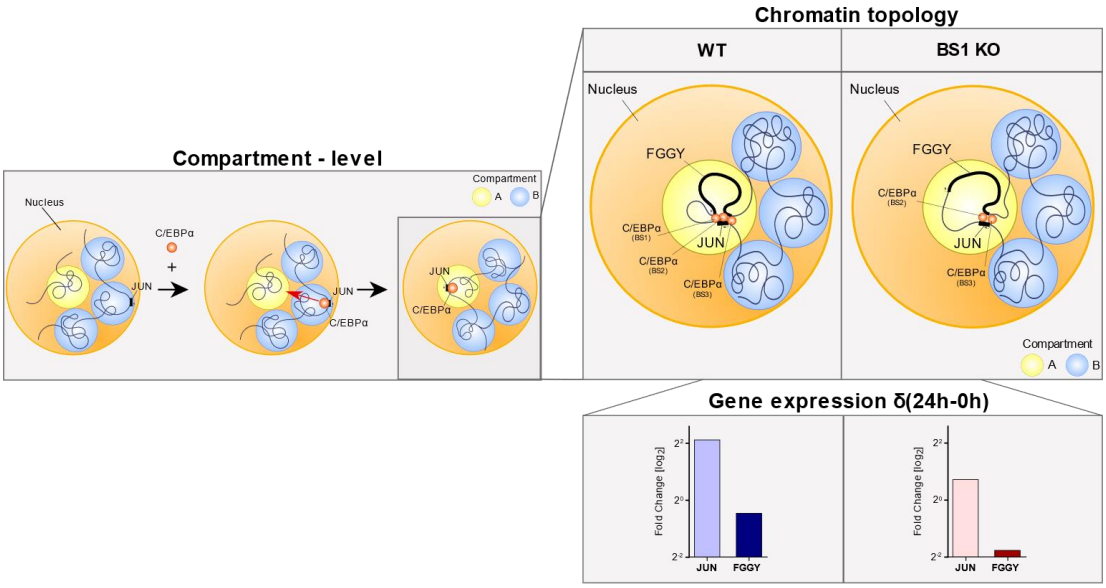


Figure 10 – Genome architecture in BS1 knockouts. Based on the results a draft was prepared to visualize genomic alterations occurring in the binding-site 1 knockout. (**Left**) Compartmentalization in BS1 KO. The deletion of the one C/EBPα binding motif did not affect the switch in compartmentalization. C/EBPα still binds to two additional binding-sites further upstream at the 3' end of FGGY which appears to rescue the effect. The transcription factor, consequently, converts this region (JUN-FGGY) from a B (blue) into an A compartment (yellow). (**Top right**) On the level of chromatin topology, only one loop domain between JUN and the 3'end of FGGY was observed in the BS1 KO mediated by C/EBPα. In contrast, the wildtype displayed two loop domains between JUN and 5' as well as 3' end of FGGY. This alteration was attributed to the binding-site deletion. (**Bottom right**) Changes were additionally observed in gene expression. The wildtype showed severally elevated JUN levels (light blue) while the KO (light red) displayed an a less sever increase 24h post induction of transdifferentiation. Likewise, FGGY expression in the KO (dark red) was downregulated compared to the wildtype (dark blue).

5. Acknowledgments

I would first like to thank my supervisor Prof. Johannes Nimpf of the Max Perutz Labs for helping me whenever I had a question and making part of this endeavor possible.

I would also like to thank Prof. Thomas Graf and all team members of the Thomas Graf laboratory at the Centre for Genomic Regulation in Barcelona, Spain where I conducted my research project as part of my master's thesis for giving me this extraordinary opportunity and support. Specifically, I would like to thank Gregoire Stik, PhD from the Graf laboratory who guided me through my research project and was always open to questions and discussions to support my research project and always gave very valuable input.

Finally, I want to express my deepest gratitude to all people who assisted me in making this scientific journey possible.

6. References

Andrey, G. and Mundlos, S. (2017) 'The three-dimensional genome: Regulating gene expression during pluripotency and development', *Development (Cambridge)*, 144(20), pp. 3646–3658. doi: 10.1242/dev.148304.

Arinobu, Y. *et al.* (2007) 'Reciprocal Activation of GATA-1 and PU.1 Marks Initial Specification of Hematopoietic Stem Cells into Myeloerythroid and Myelolymphoid Lineages', *Cell Stem Cell*. Elsevier, 1(4), pp. 416–427. doi: 10.1016/j.stem.2007.07.004.

Azagra, A. *et al.* (2020) 'From Loops to Looks: Transcription Factors and Chromatin Organization Shaping Terminal B Cell Differentiation', *Trends in Immunology*, 41(1), pp. 46–60. doi: 10.1016/j.it.2019.11.006.

Bolger, A. M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: A flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), pp. 2114–2120. doi: 10.1093/bioinformatics/btu170.

Bonev, B. *et al.* (2017) 'Multiscale 3D Genome Rewiring during Mouse Neural Development', *Cell*. Elsevier Inc., 171(3), pp. 557-572.e24. doi: 10.1016/j.cell.2017.09.043.

Bonev, B. and Cavalli, G. (2016) 'Organization and function of the 3D genome', *Nature Reviews Genetics*, pp. 661–678. doi: 10.1038/nrg.2016.112.

Bussmann, L. H. *et al.* (2009) 'A Robust and Highly Efficient Immune Cell Reprogramming System', *Cell Stem Cell*. Elsevier Ltd, 5(5), pp. 554–566. doi: 10.1016/j.stem.2009.10.004.

Cai, D. H. *et al.* (2008) 'C/EBPα:AP-1 leucine zipper heterodimers bind novel DNA elements, activate the PU.1 promoter and direct monocyte lineage commitment more potently than C/EBPα homodimers or AP-1', *Oncogene*, 27(19), pp. 2772–2779. doi: 10.1038/sj.onc.1210940.

Cremer, T. and Cremer, M. (2010) 'Chromosome territories.', *Cold Spring Harbor perspectives in biology*, 2(3), pp. 1–23. doi: 10.1101/cshperspect.a003889.

Davidson, E. H. (2010) 'Emerging properties of animal gene regulatory networks', *Nature*, 468(7326), pp. 911–920. doi: 10.1038/nature09645.

Davis, R. L., Weintraub, H. and Lassar, A. B. (1987) 'Expression of a single transfected cDNA converts fibroblasts to myoblasts', *Cell*. Elsevier, 51(6), pp. 987–1000. doi: 10.1016/0092-8674(87)90585-X.

Despang, A. *et al.* (2019) 'Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture', *Nature Genetics*. Springer US, 51(8), pp. 1263–1271. doi: 10.1038/s41588-019-0466-z.

Le Dily, F. and Beato, M. (2015) 'TADs as modular and dynamic units for gene regulation by hormones', *FEBS Letters*, 589(20), pp. 2885–2892. doi: 10.1016/j.febslet.2015.05.026.

Dixon, J. R. *et al.* (2012) 'Topological domains in mammalian genomes identified by analysis of chromatin interactions', *Nature*. Nature Publishing Group, 485(7398), pp. 376–380. doi: 10.1038/nature11082.

Everaert, C. *et al.* (2017) 'Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data', *Scientific Reports*. Springer US, 7(1), pp. 1–11. doi: 10.1038/s41598-017-01617-3.

Fudenberg, G. *et al.* (2016) 'Formation of Chromosomal Domains by Loop Extrusion', *Cell Reports*. The Author(s), 15(9), pp. 2038–2049. doi: 10.1016/j.celrep.2016.04.085.

Ghavi-Helm, Y. *et al.* (2019) 'Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression', *Nature Genetics*. Springer US, 51(8), pp. 1272–1282. doi: 10.1038/s41588-019-0462-3.

Graf, T. and Enver, T. (2009) 'Forcing cells to change lineages', *Nature*. Nature Publishing Group, 462(7273), pp. 587–594. doi: 10.1038/nature08533.

Hansen, P. *et al.* (2019) 'Computational processing and quality control of hi-c, capture hi-c and capture-c data', *Genes*, 10(7). doi: 10.3390/genes10070548.

Heyworth, C. and Pearson, Stellaay, G. E. (2002) 'Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells', *The EMBO Journal*, 21(14).

Hong, S. *et al.* (2011) 'AP-1 protein induction during monopoiesis favors C/EBP: AP-1 heterodimers over C/EBP homodimerization and stimulates FosB transcription', *Journal of Leukocyte Biology*, 90(4), pp. 643–651. doi: 10.1189/jlb.0111043.

Kim, S. and Shendure, J. (2019) 'Mechanisms of Interplay between Transcription Factors and the 3D Genome', *Molecular Cell*. Elsevier Inc., 76(2), pp. 306–319. doi: 10.1016/j.molcel.2019.08.010.

Kulessa, H. Frampton, J. & Graf, T. (1995) 'GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, ' and erythroblasts', *Genes & Development*, pp. 1250–1262.

Lajoie, B. R., Dekker, J. and Kaplan, N. (2016) 'The Hitchhiker ' s Guide to Hi-C Analysis : Practical guidelines', pp. 65–75. doi: 10.1016/j.ymeth.2014.10.031.The.

Lambert, S. A. *et al.* (2018) 'The Human Transcription Factors', *Cell*. Elsevier Inc., 172(4), pp. 650–665. doi: 10.1016/j.cell.2018.01.029.

Lesne, A. *et al.* (2019) 'Exploring mammalian genome within phase-separated nuclear bodies: Experimental methods and implications for gene expression', *Genes*. MDPI AG. doi: 10.3390/genes10121049.

Li, Yajuan *et al.* (2019) 'Systematic identification and validation of the reference genes from 60 RNA-Seq libraries in the scallop Mizuhopecten yessoensis', *BMC Genomics*. BMC Genomics, 20(1), pp. 1–12. doi: 10.1186/s12864-019-5661-x.

Lieberman-Aiden, E. *et al.* (2009) 'Comprehensive mapping of long-range interactions reveals folding principles of the human genome', *Science*. American Association for the Advancement of Science, 326(5950), pp. 289–293. doi: 10.1126/science.1181369.

Lieberman-Aiden, V. B. (2009) 'Comprehensive mapping of long-range interactions reveals folding principles of the human genome . Supplemental Online Materials Contents : Materials and Methods II . Computational Analysis Overview III . Supplemental Derivation IV . Supplemental Monte Car'.

Nora, E. P. *et al.* (2012) 'Spatial partitioning of the regulatory landscape of the X-inactivation centre', *Nature*. Nature Publishing Group, 485(7398), pp. 381–385. doi: 10.1038/nature11049.

Nora, E. P. *et al.* (2017) 'Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization', *Cell*, 169(5), pp. 930-944.e22. doi: 10.1016/j.cell.2017.05.004.

Phillips-Cremins, J. E. *et al.* (2013) 'Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment', *Cell*, 153, pp. 1281–1295. doi: 10.1016/j.cell.2013.04.053.

Ralph Stadhouders, G. J. F. & T. G. (2019) 'Transcription factors and 3D genome conformation in cell-fate decisions'. Nature. doi: https://doi.org/10.1038/s41586-019-1182-7.

Rangatia, J. *et al.* (2002) 'Downregulation of c-Jun Expression by Transcription Factor C/EBPα Is Critical for Granulocytic Lineage Commitment', *Molecular and Cellular Biology*, 22(24), pp. 8681–8694. doi: 10.1128/mcb.22.24.8681-8694.2002.

Rangatia, J. *et al.* (2003) 'Elevated c-Jun expression in acute myeloid leukemias inhibits C/EBPα DNA binding via leucine zipper domain interaction', *Oncogene*, 22(30), pp. 4760–4764. doi: 10.1038/sj.onc.1206664.

Rao, S. S. P. *et al.* (2014) 'A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping', *Cell*. Elsevier Inc., 159(7), pp. 1665–1680. doi: 10.1016/j.cell.2014.11.021.

Rapino, F. *et al.* (2013) 'C/EBPα Induces Highly Efficient Macrophage Transdifferentiation of B Lymphoma and Leukemia Cell Lines and Impairs Their Tumorigenicity', *Cell Reports*. The Authors, 3(4), pp. 1153–1163. doi: 10.1016/j.celrep.2013.03.003.

Reddy, V. A. *et al.* (2002) 'Granulocyte inducer C/EBPα inactivates the myeloid master regulator PU.1: Possible role in lineage commitment decisions', *Blood*, 100(2), pp. 483–

490. doi: 10.1182/blood.V100.2.483.

Rowley, M. J. *et al.* (2017) 'Evolutionarily Conserved Principles Predict 3D Chromatin Organization', *Molecular Cell*. Cell Press, 67(5), pp. 837-852.e7. doi: 10.1016/j.molcel.2017.07.022.

Rowley, M. J. and Corces, V. G. (2018) 'Organizational principles of 3D genome architecture', *Nature Reviews Genetics*. Springer US, 19(12), pp. 789–800. doi: 10.1038/s41576-018-0060-8.

Schoenfelder, S. and Fraser, P. (2019) 'Long-range enhancer–promoter contacts in gene expression control', *Nature Reviews Genetics*. Springer US. doi: 10.1038/s41576-019-0128-0.

Serra, F. *et al.* (2017) 'Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors', *PLoS Computational Biology*, 13(7), pp. 1–17. doi: 10.1371/journal.pcbi.1005665.

Sexton, T. *et al.* (2012) 'Three-dimensional folding and functional organization principles of the Drosophila genome', *Cell*. Cell Press, 148(3), pp. 458–472. doi: 10.1016/j.cell.2012.01.010.

Shin, Y. *et al.* (2018) 'Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome', *Cell*. Elsevier Inc., 175(6), pp. 1481-1491.e13. doi: 10.1016/j.cell.2018.10.057.

Stadhouders, R. *et al.* (2018) 'Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming', *Nature Genetics*, 50(2), pp. 238–249. doi: 10.1038/s41588-017-0030-7.

Stik, G. *et al.* (2020) 'CTCF is dispensable for immune cell transdifferentiation but facilitates an acute inflammatory response', *Nature Genetics*. Nature Publishing Group, pp. 1–7. doi: 10.1038/s41588-020-0643-0.

Vaquerizas, J. M. *et al.* (2009) 'A census of human transcription factors: Function, expression and evolution', *Nature Reviews Genetics*, 10(4), pp. 252–263. doi: 10.1038/nrg2538.

Vidal, E. *et al.* (2018) 'OneD: increasing reproducibility of Hi-C samples with abnormal karyotypes', *Nucleic acids research*, 46(8), p. e49. doi: 10.1093/nar/gky064.

Weintraub, A. S. *et al.* (2017) 'YY1 Is a Structural Regulator of Enhancer-Promoter Loops', *Cell*. Elsevier Inc., 171(7), pp. 1573-1588.e28. doi: 10.1016/j.cell.2017.11.008.

Xie, H. *et al.* (2004) 'Stepwise reprogramming of B cells into macrophages', *Cell*, 117(5), pp. 663–676. doi: 10.1016/S0092-8674(04)00419-2.

7. Table of figures

8. List of tables

9. Zusammenfassung

„Lineage-instructive“ (abstammungsinstruktive) Transkriptionsfaktoren können die Genexpression kurzschließen, indem sie DNA sequenzspezifisch an regulatorische Elemente binden. Dies führt zur Aktivierung neuer und Inaktivierung alter Gene. Die Genexpression wird auch durch Chromatinmodifikation und die komplexe dreidimensionale Organisation von Chromatin in einer höheren Ordnung reguliert. “Chromatin contact maps“, die durch „chromosome confirmational capture“-Methoden, wie Hi-C erhalten werden, zeigten, dass Chromatin auf Megabasenpaar (Mb) -Ebene in aktive (A) und inaktive (B) Kompartimente unterteilt werden kann. Die Rolle der „lineage-instructive“ Transkriptionsfaktoren bei der Gestaltung der hohen Ordnung der Genomorganisation bleibt jedoch unklar. Um diese regulatorischen Prozesse zu untersuchen, hat das Graf-Labor kürzlich ein System etabliert, das B-Zellen durch Überexpression des Transkriptionsfaktors C/EBPα in induzierte Makrophagen umwandelt. Während dieser Transdifferenzierung sind Änderungen der Genexpression mit Änderungen der A- und B-Kompartimente assoziiert, die möglicherweise durch die C/EBPα-Bindung verursacht werden. Insbesondere bindet C/EBPα in einer Region 500 kbp downstream des JUN-Locus, was zur Umwandlung der gesamten Region vom B- zum A-Kompartiment und zur Aktivierung der JUN-Expression führt. Um die topologische Rolle der C/EBPα-Bindung zu untersuchen, wurde eine CRISPR/Cas9-Genomeditierungstechnologie zum Entfernen (Knockout) der Bindungsstelle verwendet. Als nächstes wurde die Transdifferenzierung für die Wildtyp- (WT) und C/EBPα-Knockout-Zellen (KO) durch FACS-Analyse und qPCR verglichen. Eine schnellere Zellumwandlung im Vergleich zum Wildtyp und eine beeinträchtigte, jedoch nicht aufgehobene JUN-Expression wurde festgestellt. Hi-C zeigte, dass die B-zu-A-Umwandlung am JUN-Locus aufgrund des Knockouts nicht betroffen war. Die Genomtopologie wies jedoch Veränderungen auf. Eine von zwei durch C/EBPα gebildeten Schleifendomänen wurde zerstört und weniger Wechselwirkungen wurden beobachtet. Dies deutet darauf hin, dass der C/EBPα-Knockout die JUN-Expression beeinflusst und die Transdifferenzierung beschleunigt. Der Wechsel von B- zu A-Kompartiment bleibt jedoch unbeeinflusst und wird möglicherweise durch zusätzliche C/EBPα-Bindungsstellen kompensiert.