Functional Analysis of DNA Demethylation During IL4/GM-CSF-driven Differentiation of Human Monocytes

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von

Ana Karina da Silva Mendes

aus Madeira, Portugal

im Jahr 2020

The present work was carried out from January 2016 to July 2020 at the Clinic and Polyclinic of Internal Medicine III at the University Hospital Regensburg.

Die vorliegende Arbeit entstand im Zeitraum von Januar 2016 bis Juli 2020 an der Klinik und Poliklinik für Innere Medizin III des Universitätsklinikums Regensburg.

Das Promotionsgesuch wurde eingereicht am: 10.07.2020

Die Arbeit wurde angeleitet von: Prof. Dr. Michael Rehli

Unterschrift:

_____

Aos meus pais

Table of Contents

List of Figures

List of Tables

List of Abbreviations

°C	degree celsius
µg	microgram
µl	microliter
µM	micromolar
2HG	2-hydroxyglutarate
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5hmU	5-hydroxymethyluracil
5mC	5-methylcytosine
α-KG	α-ketoglutarate
AID	activation induced deaminase
AML	acute myeloid leukemia
AP	apyrimidinic
APC	antigen-presenting cell
APOBEC	apolipoprotein B mRNA editing enzyme
APS	ammonium persulfate
ATAC	assay for transposase-accessible chromatin
ATP	adenosine triphosphate
BER	base excision repair
BioID	Proximity-Dependent Biotin Identification
bp	base pair
BS	bisulfite sequencing
BSA	bovine serum albumin
C5	carbon at the fifth position
CD	cluster of differentiation
cDC	conventional dendritic cell
cDNA	complementary DNA
CEBP	CCAAT/enhancer Binding Protein
CGI	CpG island
ChIP	chromatin immunoprecipitation
CLL	chronic lymphocytic leukemia
CLP	common lymphoid progenitor
CMML	chronic myelomonocytic leukemia
CMP	common myeloid progenitor
CO_2	carbon dioxide
CoIP	co-immunoprecipitation
CpG	cytosine and guanine separated by a phosphate
CTRL	control
CXXC	cysteine-X-X-cysteine zinc-finger domain
d	day

Da	dalton
DC	dendritic cell
DMR	differentially methylated region
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DNMT	DNA methyltransferase
dNTP	deoxyribonucleotide triphosphate
dsDNA	double strand deoxyribonucleic acid
DSBH	double-stranded β-helix
DSG	disuccinimidyl glutarate
EBF1	early B-cell factor 1
E.coli	*escherichia coli*
EDTA	ethylenediaminetetraacetic acid
E.g.	For example
EGR	early growth response
ESC	embryonic stem cell
EtBr	ethidium bromide
EtOH	ethanol
FCS	fetal calf serum
Fe (II)	iron (II)
FH	fumarate hydratase
g	gram
GC	guanine/cytosine
GM-CSF	granulocyte macrophage colony stimulating factor
GMP	granulocyte-macrophage progenitor
G:T	guanine:tymine
G:U	guanine:uracil
h	hour
H2A/B	histone 2A/B
H3	histone 3
H4	histone 4
HLA	human leukocyte antigen
HSCs	hematopoietic stem cells
IAD1	interferon-regulatory factor (IRF)-association domain 1
iBAQ	intensity based absolute quantification
IDAX	inhibition of the Dvl and axin complex
iDC	immature dendritic cell
IDH	isocitrate dehydrogenase
IDT	integrated DNA technologies
IgG	immunoglobulin G
IGV	integrative genomics viewer
IL-4	interleukin 4

IP	immunoprecipitation
IRF	interferon-regulatory factor
IVT	*in vitro* transcription
JBP	J-binding protein
kb	kilobase
KD	knock-down
kDa	kilodalton
Klf4	Kruppel like factor 4
L	liter
LB	lysogeny broth (Luria-Bertani)
LPS	lipopolysaccharide
M	molar
MAC	macrophage
MBD4	methyl-CpG binding domain protein 4
M-CSF	Macrophage colony-stimulating factor
MDS	myelodysplastic syndrome
mESCs	mouse embryonic stem cells
mg	milligram
MHC	major histocompatibility complex
min	minute
ml	millilitre
MLL	mixed-lineage leukemia
MO	monocyte
Mo-DC	monocyte-derived dendritic cell
Mo-MAC	monocyte-derived macrophage
mM	millimolar
MMLV	Moloney Murine Leukemia Virus
MNCs	mononuclear cells
MPP	multipotent progenitor
MPS	mononuclear phagocyte system
mRNA	messenger ribonucleic acid
NAB	NGFI-A-binding protein
ng	nanogram
NGS	next generation sequencing
NK	natural killer
NLS	nuclear localization sequence
O_2	molecular oxygen
Oct4	octamer-binding transcription factor 4
O-GlcNAc	O-linked β-D-N-acetylglucosamine
OGT	*O*-linked β-D-*N*-acetylglucosamine transferase
OH	hydroxyl group
PBMCs	Peripheral blood mononuclear cells
PBS	phosphate buffered saline

PCR	polymerase chain reaction
pDC	plasmatoid dendritic cell
PGC	primordial germ cell
PMSF	phenylmethanesulfonyl fluoride
Pol	polymerase
PU.1	PU box binding protein 1
PVDF	polyvinylidene fluoride
riBAQ	relative intensity based absolute quantification
RNA	ribonucleic acid
RNase	ribonuclease
RPKM	Reads Per Kilobase Million
rpm	revolutions per minute
RPMI-1640	Roswell Park Memorial Institute-1640
RT	room temperature
RUNX1	Runt-related transcription factor 1
SAM	S-adenosyl-L-methionine
SAP	shrimp alkaline phosphatase
SDH	succinate dehydrogenase
SDS	sodium dodecyl sulfate
sec	second
seq	sequencing
siRNA	small interfering RNA
SMUG	single-strand-selective monofunctional uracil-DNA glycosylase
SNIP1	SMAD nuclear interacting protein 1
Sox2	SRY-Box Transcription Factor 2
SWI/SNF	switch/sucrose non-fermentable
TAE	tris-acetate-EDTA
TBS	tris-buffered saline
TCA	tricarboxylic acid
TDG	thymine DNA glycosylase
TEMED	N,N,N',N'-Tetramethylethylenediamine
TET	ten-eleven translocation
TF	transcription factor
TNFα	Tumor necrosis factor alpha
TSS	transcription start site
TTS	transcription termination site
U	unit
UCSC	University of California, Santa Cruz
UHRF1	ubiquitin-like, PHD and ring finger–containing 1
UV	ultraviolet
V	volt
WGBS	whole genome bisulfite sequencing
WT1	Wilms Tumor 1

1 Introduction

1.1 Epigenetics & DNA Methylation

The term “epigenetics” is used to describe heritable alterations in gene expression that are not encoded in the DNA sequence[1,2]. Accordingly, epigenetics regulates gene expression by switching genes on or off and determining which proteins are transcribed. To accomplish this task, important epigenetic mechanisms are involved, including DNA methylation, histone modifications, nucleosome and higher-order chromatin structure and the processes mediated by non-coding RNAs[1].

In a eukaryotic nucleus, the nucleosome core (the basic packaging unit of DNA) consists of a histone octamer of four dimers of each core histone (H2A, H2B, H3, and H4) wrapped around 147 bp DNA[3,4]. The nucleosomes can be found every 200 ± 40 bp[5] and their positioning throughout a genome has an important regulatory function. In cooperation with the binding of transcription factors (TFs), RNA polymerases or architectural proteins, nucleosomes facilitate different chromatin dynamics[3] (Figure 1.1) thus influencing processes such as transcription, DNA repair, replication and recombination[4].

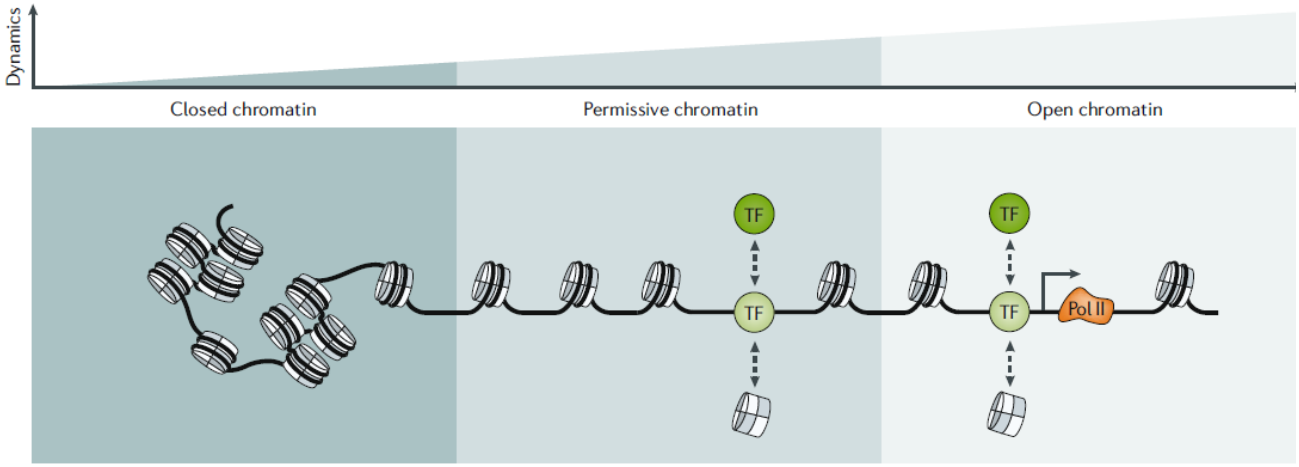

Figure 1.1 Chromatin dynamics across the genome through a continuum of accessibility states

In contrast to closed chromatin, permissive chromatin is more dynamic and allows for TFs to initiate sequence-specific accessibility remodeling towards an open chromatin conformation. Pol II, RNA polymerase II; TF, transcription factor (adapted from Klemm *et al.* [3]).

Considering that, in this thesis the analyses were primarily focused on mechanisms involved in loss of DNA methylation, this chapter provides an overview about the DNA methylation process, one of the most widely investigated epigenetic mechanisms of genome regulation.

DNA methylation exerts a fundamental role in cellular differentiation, development and disease[6–10]. Its role is particularly well established in hematopoeisis. For instance, previous work performed by Bröske and colleagues[8] in mice, demonstrated that DNA methylation-associated mechanisms can control the ability of hematopoietic stem cells (HSCs) to differentiate into myeloerythroid versus lymphoid lineages. In addition, Lessard and colleagues[11] verified that DNA methylation at erythroid enhancers is essential to determine transcriptional and development differences between human fetal and adult erythropoiesis. Also associated with enhancer methylation studies, Agirre and colleagues[12] work revealed that DNA hypermethylation of B-cell specific enhancers is linked with the pathogenesis of multiple myeloma.

Overall, DNA methylation is a key epigenetic process in self-renewal and differentiation of HSCs into different lineages (Figure 1.2) by promoting alternative lineage-specific gene expression[6,8,13]. Besides its function in regulating transcription, DNA methylation is also involved in genomic imprinting, X-chromosome inactivation and suppression of mobile genetic elements[14–17].

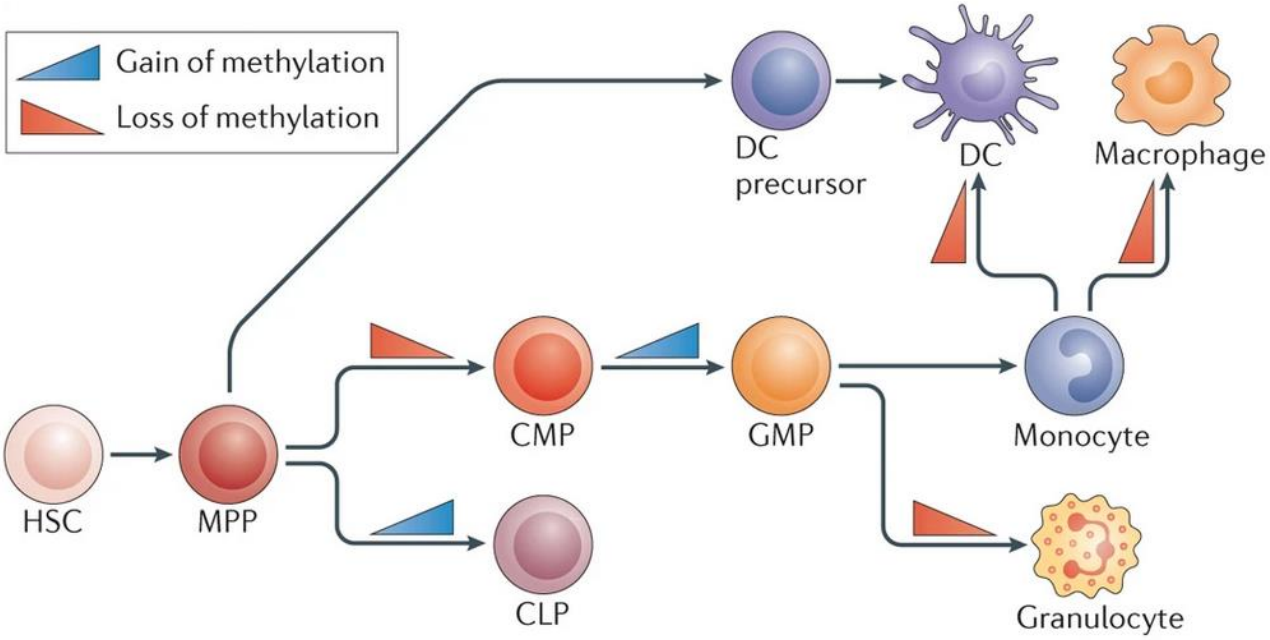

Figure 1.2 DNA methylation patterns in the early stages of hematopoeisis

The figure depicts the main events of methylation that occur during adult hematopoiesis, including data obtained from studies using human and mice primary cells as well as cell lines. The blue triangles indicate an overall increase in methylation, while the red triangles represent a loss of methylation at specific stages of hematopoiesis. HSC, hematopoietic stem cell; MPP, multipotent progenitor; CMP, common myeloid progenitor; CLP, common lymphoid progenitor; GMP, granulocyte-macrophage progenitor; DC, dendritic cell (adapted from Álvarez-Errico *et al.*[13]).

DNA methylation is mediated by DNA methyltransferase (DNMTs) enzymes, which in mammals comprise a family of five members, DNMT1, DNMT2, DNMT3A, DNMT3B and DNMT3L. However, only DNMT1, DNMT3A and DNMT3B possess confirmed methyltransferase activity[14,18,19]. These enzymes catalyse the transfer of a methyl group from S-adenosyl-L-methionine (SAM) to the carbon at the fifth (C5) position of cytosines that are adjacent to guanines in DNA (known as CpG

dinucleotides). As a result, regulatory elements become hypermethylated, which in turn leads to transcriptional repression[2,18–20].

The methylation activity of DNMT1 is tailored to hemimethylated cytosines in CpG dinucleotides sequences, thus maintaining the methylation pattern during DNA replication. On the other hand, DNMT3A and 3B are involved in *de novo* DNA methylation[17].

In the mammalian genome, about 60-80% of CpGs are methylated[10,17,19,21], which actually corresponds only to 5% of all cytosines residues[19]. In addition, most of these CpGs sites are located in short CpG-rich DNA regions called CpGs islands (CGIs) that are preferentially located near the transcription start sites (TSS) of more than 50% of gene promoters[2].

Analyses of differentially methylated regions have shown the co-localization of hypomethylated regions with both promoters and enhancers of actively transcribed, tissue specific and developmental stage-specific genes[6,19,22,23].

Due to the importance of DNA methylation turnover for mammalian development and differentiation (as exemplified for the hematopoeitic system), the next chapters provide more insight into the DNA demethylation process with a major focus on the active DNA demethylation mechanism.

1.2 DNA Demethylation

Over decades DNA methylation was considered a stable, persistent and heritable mark[2,24]. Hence, loss of methylation was believed to take place in a replication-dependent manner in the absence or inhibition of the maintenance methylation machinery[24]. However, a major revolution in this field occurred in 2009, when Kriaucionis and Heintz[25] identified higher levels of 5-hydroxymethylcytosine (5hmC) in mouse Purkinje neurons and granule cells. Another group reported independently that a family of enzymes called Ten-eleven Translocation (TET) oxidize 5mC to 5hmC, both *in vitro* and in mouse embryonic stem cells (mESCs)[26]. These two studies uncovered an enzymatic pathway to erase methylation in a process called active DNA demethylation.

Since then, many reports have demonstrated the occurrence of this process in different biological settings, such as during embryonic and primordial germ cell development (PGC)[10,27,28], HSC differentiation[13] and in post-mitotic cells in the adult brain[29,30]. Similarly, our group also reported active DNA demethylation in post-mitotic human monocytes (MO)[31,32].

1.2.1 TET enzymes

TET proteins are large (approximately 180-230 kDa) multidomain enzymes[33] originally identified in the context of hematological malignancies[34]. TET1 was initially cloned and characterized as an acute myeloid leukemia (AML)-associated protein, which is fused to MLL (mixed-lineage leukemia) in

t(10;11)(q22;q23) translocations. Further sequence homology analysis identified two additional paralogues in human and mouse, namely TET2 and TET3[19,34,35].

The catalytic activity of TET enzymes in the oxidation of 5mC, was predicted based on sequence homology with the trypanosome proteins, JBP1 and JBP2 that are known to oxidize the methyl group of thymine to 5-hydroxymethyluracil (5hmU)[19,24,26].

1.2.1.1 Function & Structure

There is ample evidence supporting the role of TET enzymes in the iterative, stepwise oxidation of 5mC to 5hmC, 5-formylcytosine (5fC) and then 5-carboxycytosine (5caC)[19,36–39]. Besides their function as demethylation intermediates, the oxidized forms of 5mC may also serve as relatively stable epigenetic modifications[40,41] with essential regulatory functions[19]. Supporting this hypothesis, Raiber and colleagues[42] reported that 5fC might affect transcription regulation in terms of chromatin remodeling through its effect on DNA conformation. Nevertheless, further work is necessary to elucidate the potential regulatory roles of these 5mC oxidized derivatives in biological processes.

In addition to the participation of TET enzymes in the oxidation of cytosine-derived nucleobases, these enzymes were shown to catalyze the oxidation of thymine to 5hmU in mESCs, although in very low amounts[43]. Interestingly, TET enzymes were also reported to be involved in chromatin modifications and other cellular processes through the interaction with the O-linked b-N-acetylglucosamine (O-GlcNAc) transferase (OGT)[44]. This enzyme catalyzes the addition of O-GlcNAc onto serine and threonine residues (O-GlcNAcylation) *in vivo*[45–47]. Accordingly, it was proposed that the interaction of TET2 with OGT facilitates OGT-dependent histone O-GlcNAcylation, emphasizing the role of the TET enzymes in gene transcription regulation by chromatin modifications[44,45,48].

TET proteins are iron (II)/α-ketoglutarate (Fe (II)/α-KG)-dependent dioxygenases. As illustrated in Figure 1.3, all TET isoforms contain a C-terminal core catalytic domain, comprising a double-stranded β-helix (DSBH) domain and a cysteine rich domain[10,17,27,33]. The DSBH domain contains the binding sites for Fe (II) and α-KG (essential for the oxidation of 5mC), while the cysteine-rich domain wraps around the DSBH core to stabilize the general structure and the TET–DNA interaction[17]. Moreover, TET1 and TET3 contain N-terminal CXXC zinc finger domains, allowing these enzymes to bind to unmethylated, methylated and hydroxymethylated DNA[2]. In contrast, the CXXC domain in TET2 was lost during evolution as a result of a genomic inversion and is now coded as a separated protein called CXXC finger protein 4 (CXXC4)[24].

Myeloid cells, including MO and MO-derived cells, primarily express the TET2 isoform[32]. Therefore, the present work focused on the TET2-mediated active DNA demethylation.

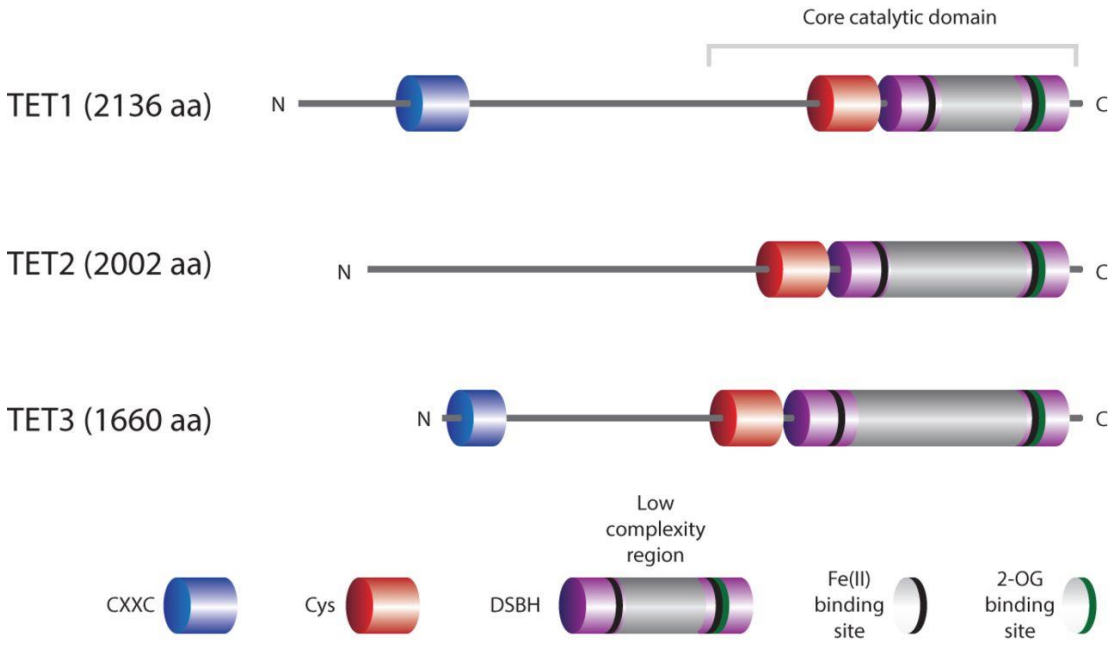

Figure 1.3 Domain structure of TET proteins

All TET isoforms share an identical C-terminal catalytic domain, which includes the DSBH domain, a cysteine-rich (Csy) domain and binding sites for the Fe (II) and 2-OG (2-oxoglutarate, also known as α-KG) co-factors. The DSBH domain possesses a low complexity region of unknown function. In contrast to TET2 isoform, TET1 and TET3 have an N-terminal CXXC domain that can bind directly to DNA (adapted from Rasmussen and Helin[33]).

1.2.1.2 Substrates & Co-factors

TET enzymes use molecular oxygen (O_2) and α-KG as substrates and Fe (II) as a co-factor to successively oxidize DNA, generating oxidized DNA, succinate and carbon dioxide (CO_2) as co-products (Figure 1.4)[17,34,49]. Consequently, the reaction kinetic is directly affected by the availability of substrates and co-factors.

α-KG is generated by isocitrate dehydrogenase enzymes (IDH1-3), which convert isocitrate into α-KG as part of the tricarboxylic acid (TCA) cycle [17,34,50].

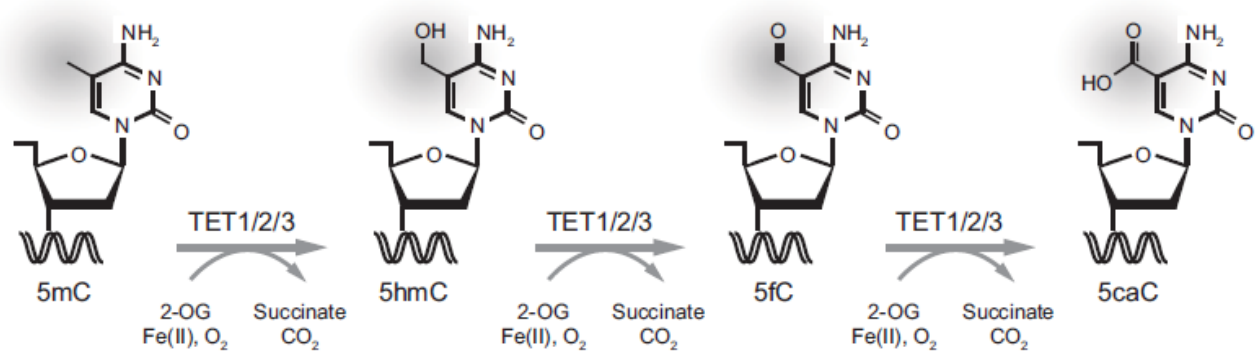

Figure 1.4 Stepwise oxidation of 5mC to 5hmC, 5fC and 5caC by TET enzymes

TET enzymes use O_2 and 2-OG as substrates and Fe (II) as a co-factor to successively oxidize DNA, generating 5hmC, 5fC and 5caC as well as the co-products succinate and CO_2 (adapted from Rasmussen and Helin[33]).

Similarly, vitamin C has been repeatedly reported to act as a co-factor of TET enzymes, by maintaining the essential iron in the catalytic site in a reduced state (Fe^{2+}) and enhancing the enzymatic activity. Consequently, there is an increase in the 5hmC content, which facilitates the DNA demethylation[17,51–56]. Additionally, interaction of vitamin C with the TET enzymes catalytic domain may promote its folding and/or recycling of the co-factor Fe^{2+} [56].

1.2.1.3 Dysregulation

Myeloid cells express high levels of DNMT3A and TET2, epigenetic enzymes involved in cytosine modifications[57]. This fact, associated with the high frequency of recurrent mutations in TET2 and DNMT3A found in a wide spectrum of myeloid malignancies[57–60], suggests a fundamental role of these proteins in the normal hematopoiesis[13,44], as well as in the specification of myeloid identity[57].

Besides the genetic alterations in genes of these epigenetic regulators, changes in genes involved in cellular metabolism are also implicated in the pathogenesis of myeloid malignancies[33,34,61]. This includes perturbations of the TCA cycle that lead to accumulation of aberrant metabolites, which inhibit TET enzymes (Figure 1.5) and other α-KG-dependent dioxygenases[34,62,63].

Some key enzymes of the TCA cycle like IDH1/IDH2 that catalyze the conversion of isocitrate into α-KG (critical co-factor for the activity of TET2[44]) are recurrently mutated in hematological malignancies. Their gain-of-function mutations lead to the abnormal production of 2-hydroxyglutarate (2-HG), an oncometabolite that competes with α-KG for TET binding[17,64,65]. In fact, patients with mutations in IDH1/IDH2 present high levels of 2-HG[66], which associate with DNA and histone hypermethylation and blocked differentiation of hematopoietic progenitor cells[61].

Likewise, mutations can be found in additional metabolic enzymes. For example, mutations in succinate dehydrogenase (SDH) and fumarate hydratase (FH) lead to the accumulation of succinate and fumarate. In a similar way as 2-HG, they can compete with α-KG to inhibit α-KG-dependent dioxygenases, such as TET enzymes, causing an increase in DNA methylation[17,66–68].

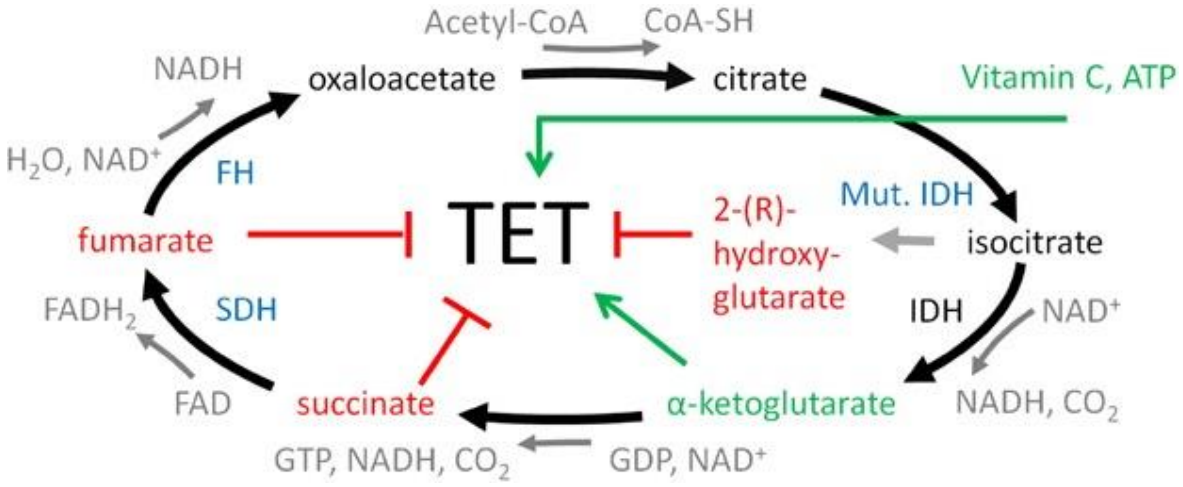

Figure 1.5 Impact of small molecule metabolites on TET activity

The TCA cycle is presented here in a simplified form, omitting enzymes and metabolites without clear link to TET activity. The co-substrate α-KG and the co-factor vitamin C as positive regulators of TET activity are highlighted in green, whereas fumarate, succinate and 2-(R)-HG as negative regulators of TET are highlighted in red. Enzymes that have been (in their mutant variants) implicated in TET inhibition are shown in blue. IDH, SDH, and FH have mitochondrial and cytoplasmic isoenzymes, encoded by separate genes in the case of IDH. Mutations in IDH, SDH and FH have been implicated in TET dysregulation (adapted from Bochtler *et al*.[34]).

1.2.2 TET-mediated DNA Demethylation Pathways

Since the discovery of TET oxidation activity as well as their related oxidized products, numerous DNA demethylation pathways involving TET enzymes have been proposed (Figure 1.6) both *in vitro* and *in vivo*.

As mentioned before, if the maintenance methylation machinery is non-functional, 5mC can be passively diluted during successive rounds of DNA replication either globally or locally. Intriguingly, some studies also implicated TET proteins in passive DNA demethylation processes (during DNA replication). They assume that TET enzymes oxidize 5mC to their respective oxidized forms, which then might be successively diluted to regenerate unmodified cytosines in a replication-dependent manner[10,17,66]. This proposed pathway is based on the lower DNMT1 activity on hemi-5hmC containing DNA (>60 fold, *in vitro*)[24,69,70] that leads to a passive dilution of the intermediate oxidized forms during replication. Although this mechanism was reported to occur during mouse erythropoiesis *in vivo*[71], some observations argue against it. For example, the DNMT1 interaction partner UHRF1 binds 5hmC[72,73], which then may target DNMT1 to hemi-5hmC containing DNA. Secondly, contrary to DNMT1, DNMT3A and DNMT3B are not sensitive to hemi-5hmC DNA and can re-methylate the hemi-5hmC DNA[24]. Thereby, further investigation is required to elucidate this potential TET demethylation replication-dependent mechanism.

Another DNA demethylation pathway reported in the mouse brain[74] involves the deamination of 5hmC to 5hmU by the AID/APOBEC family members. The 5hmU is then removed by DNA glycosylases, such as thymine DNA glycosylase (TDG), SMUG, MBD4 or NEIL[29,74–76]. Nevertheless, this mechanism

still remains controversial because TETs may also oxidize thymines to 5hmU followed by DNA demethylation through DNA glycosylases and base excision repair (BER) mechanisms[24,43].

Besides these suggested mechanisms, a biochemical demethylation pathway involving the direct decarboxylation of 5caC to unmethylated cytosine has been proposed. Nonetheless, a decarboxylase specific for 5caC in mammals has not been identified so far[24,77,78].

Extensive research suggests that an active demethylation pathway involving TET-TDG-BER represents the most likely demethylation mechanism triggered by TET enzymes in a replication-independent manner[10,17,27,79]. In this process, the oxidized forms 5fC and 5caC (generated by TET enzymes) are recognized by a specific DNA glycosylase, namely TDG (5mC and 5hmC are not recognized by TDG)[36,80,81]. TDG is a DNA mismatch repair enzyme responsible for binding and excision of mismatched pyrimidines in G:U and G:T base pairs[33]. This enzyme cleaves the N-glycosidic bond between the base and the sugar at 5fC and 5caC residues generating abasic (apyrimidinic, AP) sites[82]. Then, the abasic sites are processed by the BER machinery that generates an unmodified cytosine. The BER mechanism involves an AP endonuclease I (Apex 1) that cleaves the phosphodiester bond, generating a 3′-OH and 5′-deoxyribose phosphate. Then, the DNA polymerase β removes the sugar phosphate moiety and incorporates an unmodified cytosine. At the end, DNA ligase I or IIIα seal the nick[82,83].

Notwithstanding all the proposed mechanisms for processing of 5mC to unmodified cytosines (5C), in post-mitotic MO and MO-derived cells like iDCs, it is poorly understood which mechanisms and enzymes are involved in further processing of 5hmC to 5C.

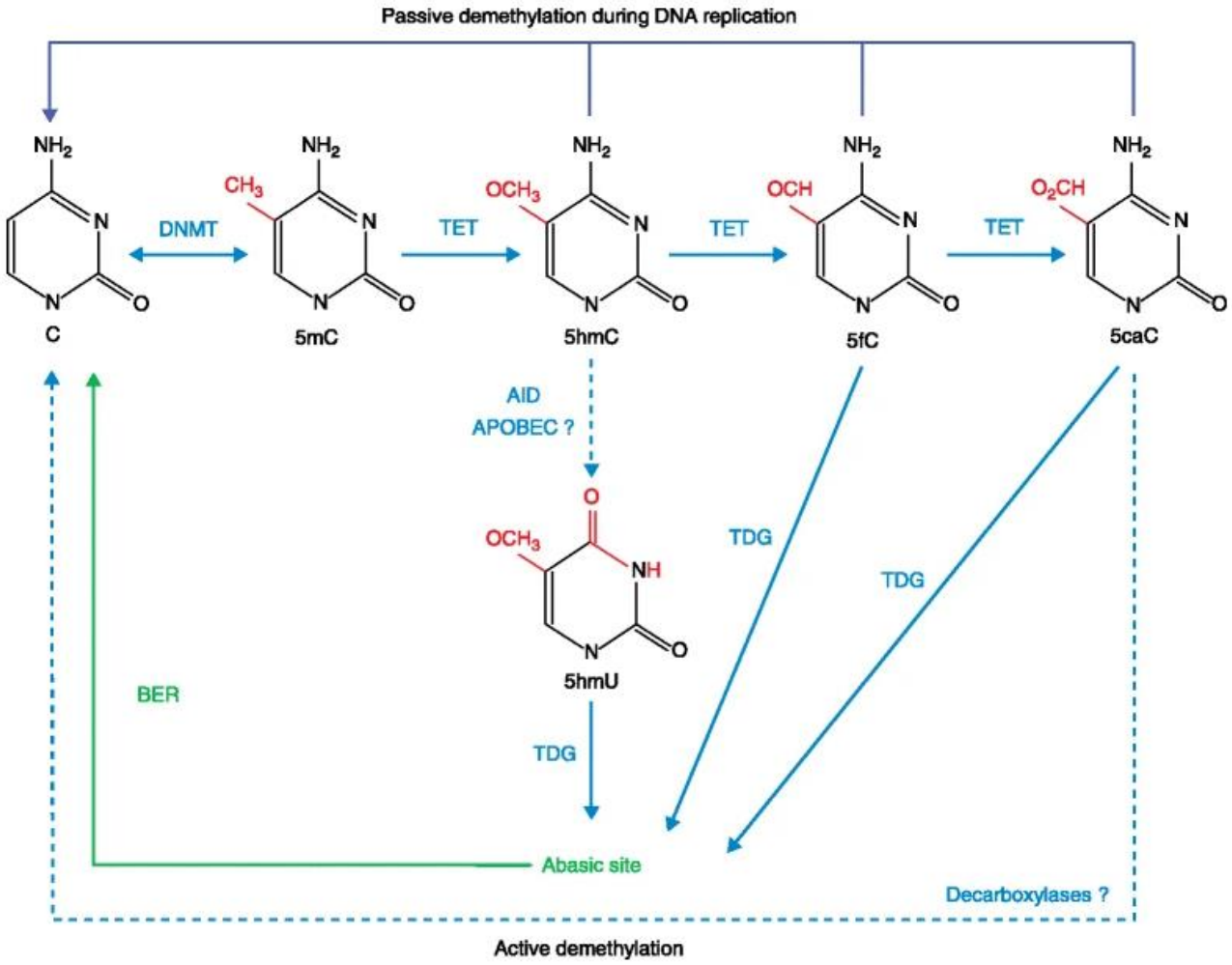

Figure 1.6 TET-mediated DNA demethylation pathways

During DNA replication, DNMT enzymes participate in the generation and maintenance of the DNA methylation patterns. In contrast, TET enzymes, which successively oxidize 5mC to 5hmC, 5fC and 5caC, are involved in DNA demethylation. Though still controversial, some reports hypothesize that TET enzymes may be implied in the passive DNA demethylation due to the lower activity of DNMT1 (maintenance methylation) on hemi-5hmC containing DNA, resulting in the passive dilution (during DNA replication) of the oxidized intermediate forms (5hmC, 5fC and 5caC) generated by TETs. Another unclear DNA demethylation pathway involves the deamination of 5hmC to 5hmU by the AID/APOBEC family members, which is then removed by TDG generating an abasic site as part of the BER that regenerates an unmodified cytosine. An alternative proposed pathway, implicates direct decarboxylation of 5caC to C. However, a decarboxylase specific for 5caC has not been identified so far. Accordingly, the pathway involving direct removal of 5fC or 5caC by TDG followed by BER towards the replacement with a C, is the most likely mechanism triggered by TET enzymes (active demethylation). AID/APOBEC, activation-induced deaminase/apolipoprotein B mRNA editing enzyme (adapted from Scourzic *et al.*[77]).

1.2.3 Transcription Factors & DNA Demethylation Recruitment

Transcription factors (TFs) are proteins that bind DNA in a sequence-specific mode and regulate transcription[84]. For this purpose, TFs bind to TF binding sites or motifs[85] (6-10 bp regions) located in regulatory DNA regions (promoter and enhancers)[86]. They are also able to interact with each other or compete for DNA binding, suggesting that their binding sites may co-localize or overlap[86,87].

Although each TF is believed to possess DNA-binding domains, under certain circumstances, they may bind DNA indirectly by interacting with another TF[86]. In fact, critical TFs in the differentiation of hematopoietic lineages like PU.1 and GATA-1 are capable to interact with each other, allowing PU.1 to bind to DNA both directly and indirectly (in cooperation with GATA-1)[86,88].

Interestingly, the same TF can regulate distinct genes in different cell types[84]. In line with this, TF networks define cell type identity by initiating and sustaining lineage-specific expression profiles[85] as observed during cell differentiation.

Biological processes like cell differentiation depend on the combined action of a set of TFs, specifically pioneer and non-pioneer TFs. The former TFs are able to bind “closed” chromatin and establish an open chromatin conformation that enables the latter TFs (lacking that feature) to bind DNA[86]. Examples of pioneer TFs include Early B-cell factor 1 (Ebf1), which is essential during B cell development[89], and CEBPα, which is required for trans-differentiation of pre-B cells into macrophages (MAC)[90]. In a similar way, pioneer factors like Oct4, Kruppel-like factor 4 (Klf4), and Sox2 mediate reprogramming of fibroblasts into induced pluripotent stem cells[91,92].

Remarkably, recent work in two pituitary lineages, melanotropes and corticotropes[93], proposed that Pax7 (pioneer TF) binds heterochromatin regardless of Tpit (non-pioneer TF), whereas the latter is required for Pax7-dependent open chromatin. Therefore, this work suggests that cooperation between pioneer and non-pioneer TFs may mediate lineage specific chromatin opening[93].

While TF binding patterns and DNA methylation strongly anti-correlate, it is still controversial whether TF binding occurs before DNA methylation or vice versa[94]. The established model suggests that DNA methylation patterns instruct TF binding with DNA methylation repressing the binding of TFs[94,95] sensitive to DNA methylation like NRF1[96]. On the other hand, recent studies propose an emerging model where TF binding controls DNA methylation patterns[97–99]. Accordingly, TFs identified as insensitive to DNA methylation[100] (e.g. OCT4[101]), bind to methylated regions and recruit TET enzymes to initiate DNA demethylation (Figure 1.7).

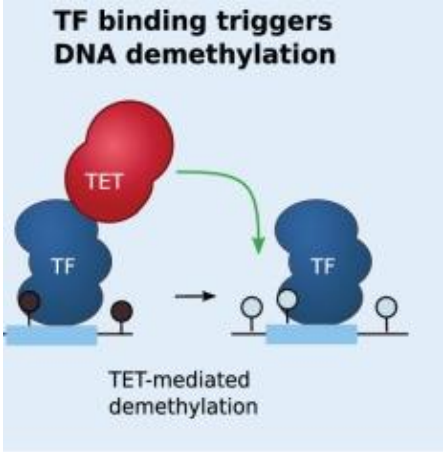

Figure 1.7 TF-mediated DNA demethylation

TFs insensitive to DNA methylation were first shown in mouse ESCs to bind methylated CpGs (black lollipops) and recruit TET proteins to induce DNA demethylation. As a consequence, TFs sensitive to DNA methylation are then capable to bind to these sites (adapted from Héberlé and Bardet[94]).

As previously mentioned, TET1 and TET3 possess a CXXC domain that mediates direct binding to DNA (methylated or unmethylated)[102], and may thus facilitate their recruitment to the genomic target sites[33]. Conversely, TET2 lacks a CXXC interaction motif, suggesting that it may depend on other TF(s) for locus specific recruitment[24] as proposed in the latter model.

As a matter of fact, multiple TFs important in cell differentiation and lineage specification were reported to interact with TET2, modulating its activity and target gene expression. For example, in AML cells, Wilms Tumor 1 (WT1) recruits TET2 to their target genes[103,104] and EBF1 was also reported to interact with TET2 in IDH-mutant cancers[105]. In a similar way, PU.1 interacts with TET2 during MO to osteoclast differentiation, which facilitates demethylation recruitment to PU.1 binding sites[17,97]. Another study showed physical interactions between Runt-related transcription factor 1 (RUNX1) and DNA demethylation machinery enzymes (TET2, TET3, TDG and GADD45) in hematopoietic cells, suggesting a potential function of RUNX1 in demethylation recruitment[106].

Chen and colleagues[107] used a mammalian two-hybrid screen and identified the SMAD nuclear interacting protein 1 (SNIP1) as a physical interactor of TET2. They postulated that SNIP1 is involved in TET2 recruitment to the promoters of c-MYC target genes, comprising genes linked with DNA damage response and cell viability. Interestingly, Sardina and collaborators[98] found that C/EBPα, Klf4, and Tfcp2l1 can recruit Tet2 to specific DNA sites, leading to enhancer demethylation and activation throughout reprogramming in mouse.

Additional interaction partners of TET proteins (reported in mouse ESCs) that may contribute to their recruitment to specific loci, include the pluripotency factor NANOG[108], PR domain zinc finger protein 14 (PRDM14)[109], Polycomb repressive complex 2 (PRC2)[110] and LIN28A[111]. Some of these studies require further investigation to clarify whether the interaction of these factors with TET proteins mediates the recruitment, or instead, if these TFs interfere with the chromatin environment for TET binding[17].

1.3 Human Mononuclear Phagocyte System

The Mononuclear Phagocyte System (MPS) encompasses the mononuclear fraction of myeloid cells, which includes MO, MO-derived cells, MAC and DC populations. This system constitutes a complex, cellular compartment of phenotypically and functionally heterogeneous cells[112,113]. However, the functional significance of their diversity as well as the mechanism of diversification remain poorly understood[114].

MPS exhibit important functions that are essential, for example, to maintain homeostasis, to resolve inflammation and to heal wounds[115], to activate innate immunity and to promote the crosstalk between innate and adaptive immunity[116]. For this purpose, cells from the MPS possess specialized functions in phagocytosis, cell recruitment, antigen presentation to T cells and cytokine production[117] that facilitate the regulation of innate and adaptive immune responses. For instance, under inflammation or infection, intermediate chemokines can migrate to these sites[118] directed by these cells and trigger an immune response. This might be observed in chronic inflammatory diseases such as rheumatoid arthritis, neurodegenerative disorders and atherosclerosis[119].

Taken together, cells from the MPS can participate in numerous activities, such as infection defence, tissue homeostasis and control of T cell immunity[113,120–122]. Additional critical homeostatic functions are associated with tissue repair, remodeling, angiogenesis and neural networking in the course of embryonic, fetal and postnatal development[123,124].

1.3.1 *In Vitro* Models

The differentiation steps of the MPS can be recapitulated *in vitro* using specific growth/differentiation factors such as granulocyte-macrophage colony stimulating factor (GM-CSF) and interleukin 4 (IL-4). Earlier work of Sallusto and Lanzavecchia[125] demonstrated that isolated peripheral blood-MO upon stimulation with GM-CSF and IL-4 acquire a DC morphology and produce inflammatory cytokines. In addition, under further maturation by adding toll-like receptor ligands, tumor necrosis factor alpha (TNFα) or lipopolysaccharide (LPS)[126], DCs display an efficient antigen cross-presentation capacity *in vitro,* which is characteristic of DCs *in vivo*[125].

Despite applicability of the *in vitro* systems to study cellular processes like cell differentiation, the *in vivo* conditions leading to the formation of mo-DC remain poorly understood and difficult to reproduce by *in vitro* experiments[127]. Nevertheless, some *in vivo* studies suggest that at steady state, mo-DC are primarily found in the intestine[128,129], while under inflammatory conditions these cells were described in atopic dermatitis and psoriasis patients[130]. In addition, mo-DCs were also detected in lung cancer and colorectal and breast tumors[129,131,132].

In contrast to *in vitro* mo-DCs, the *in vitro* differentiation of blood-MO into MAC can be explored using different strategies. The most common strategies include the cultivation of MO in the presence of

human AB-serum[133] or with either GM-CSF or macrophage colony-stimulating factor (M-CSF)[134]. In addition, these cells can be activated in response to different stimuli like LPS, interferon-gamma (IFNγ), IL-4 and IL-10, producing distinct MAC phenotypes[134]. Interestingly, mo-MAC *in vivo* studies described the presence of these cells in the intestine and in the skin at steady state as well as in tissues with poorer self-renewal abilities[129] such as in the heart[135] lung[136] and liver[137].

Generally, under proper stimulation, the *in vitro* MO-derived cells may resemble their *in vivo* counterparts as reported by Goudot and colleagues[138]. These authors showed that their system (human $CD14^+$ MO cultured with M-CSF, IL-4 and TNF-α) to generate *in vitro* mo-DC and mo-MAC (in the same culture) strictly resembled the inflammatory mo-DC and mo-MAC observed in human tumor ascites.

In this thesis, the *in vitro* model depicted in Figure 1.8 was used to differentiate peripheral blood human MO into iDC upon stimulation with IL-4 and GM-CSF cytokines.

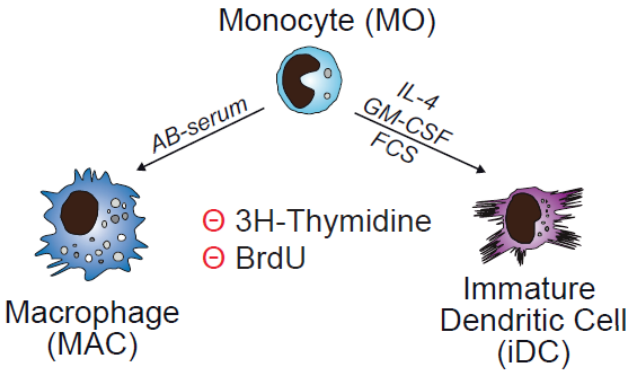

Figure 1.8 Post-mitotic differentiation model of *in vitro* MO differentiation

Schematic representation of *in vitro* MO differentiation. Peripheral human blood MO can be differentiated into MAC in the presence of human AB-serum or into iDCs upon stimulation with GM-CSF and IL-4. In addition, MO do not proliferate (as indicated by the lack of nucleotide incorporation) and thus, DNA demethylation takes place as an active process (adapted from Klug *et al.*[32]).

Based on many proliferation assays, our group previously confirmed that the differentiation of post-mitotic MO into iDC occurs independent of any replication events[31]. Therefore, this system represents an ideal model to investigate cellular processes in the absence of DNA replication like the active DNA demethylation process.

1.3.2 Monocytes

MO represent the key members of the MPS, constituting between 4-10% of total leukocytes in the blood[117]. These circulating MO have the capacity to differentiate into different phagocytes like MAC, myeloid DCs, osteoclast and microglia in the central nervous system[139].

Considering that MO circulate in the bloodstream and traffic to tissues during steady state and at higher rates during inflammation[140], they play a critical role in many diseases with an inflammatory component, such as infection, cardiovascular disease, type I diabetes and cancer[141–145].

Besides their potential role in triggering inflammation responses, further evidences suggest that MO can carry antigens to lymph nodes[146], patrol and clean up the vasculature[147,148], recognize pathogens[149] and help to kill tumor cells[140,150].

The development of MO from the common myeloid progenitor takes place in the bone marrow and is regulated by sequential expression of key TFs, including PU.1 (master regulator required for development of all myeloid cells) CCAAT/enhancer-binding protein beta (CEBPβ), IRF8 and KLF4[140,151,152]. During MO maturation KLF4 is upregulated while IRF8 is downregulated, which underpins the importance of IRF8 at an earlier step of MO development, whereas KLF4 and CEBPβ function in latter steps that lead to the development of patrolling MO[141,153,154].

In humans, MO can be divided into three subsets by their distinctive expression of the LPS-coreceptor CD14 and the Fcɣ III receptor CD16[155]. This includes classical MO ($CD14^{++}CD16^{-}$), non-classical MO ($CD14^{dim}CD16^{+}$) and intermediate MO ($CD14^{+}CD16^{+}$)[140,155]. Nonetheless, recent findings include additional surface markers for improving discrimination of MO subsets, such as HLA-DR, CCR2, CD36, and CD11c[140,156]. In addition, each MO subset is characterized by differential gene-expression patterns, transcriptional regulation and specific functions[157].

In general, classical MO encompass about 80–95% of circulating MO. These cells are highly phagocytic and are known as important scavenger cells. Intermediate MO comprise about 2–8% of circulating MO. They are fundamental to produce reactive oxygen species (ROS), antigen presentation, inflammatory responses, angiogenesis and are also involved in the proliferation and stimulation of T cells. Non-classical MO cover about 2–11% of circulating MO and their functions may include patrolling the endothelium in search of injury, secretion of inflammatory cytokines in response to infection, antigen presentation and T cell stimulation[158–160].

1.3.3 Dendritic Cells

DCs were initially discovered in 1973[161] and represent a heterogeneous class of bone marrow-derived professional antigen-presenting cells (APCs) widely distributed in both lymphoid and non-lymphoid tissues[162,163]. They originate from HSCs specialized progenitor subsets and are essential mediators between the innate and adaptive immune responses. The role of DCs in the innate immune response is to recognize and respond to pathogen-associated signals through pattern-recognition receptors, thus sharping the acute inflammatory response. In respect to the adaptive immune response, DCs participate in the processing and presentation of pathogen-derived peptides in the context of major histocompatibility complex (MHC) molecules to prime naïve T cells[127,162].

Despite immature DCs (*in vivo*) capacity to capture and process antigens, they represent poor inducers of immune responses[164]. However, upon recognition of exogenous and endogenous signals by Toll-like receptors (TLRs), they undergo a maturation process that correlates with upregulation of cell surface MHC gene products, co-stimulatory molecules (CD40, CD80, and CD86 and CD83) and

relevant chemokine receptors. Subsequently, all these stimuli improve the ability of DCs to migrate toward lymphoid tissues where they initiate antigen-specific T cell responses[162,165].

Although under normal conditions DCs are present at low density, they are critical in managing the balance between immunity and tolerance[117,166]. Indeed, DCs are able to regulate immune responses by producing both central and peripheral tolerance and controlling inflammatory responses through numerous mechanisms. Some of these mechanisms include, induction of T cell anergy and apoptosis of autoreactive T cells, expansion of regulatory T cells and limitation of other effector cell responses[164,166].

Based on their localization, phenotype and function[167], DCs can be divided into three major subsets, which include plasmacytoid DC (pDC), myeloid/conventional DC (cDC) and DCs derived from monocytes (Mo-DCs). As illustrated in Figure 1.9, these DC subsets develop under the differential expression of IRF8 and IRF4 in collaboration with specific TFs, such as PU.1, ID2, E2-2, KLF4 and BATF3[127].

pDCs are a specialized subset of DCs with a critical role in the antiviral response[168]. Under steady-state conditions, these cells possess an immature phenotype (plasma cell morphology (e.g., lack dendrites))[162] and lack antigen presenting functions[169]. However, upon proper stimulation, such as during exposure to viral stimuli, they can differentiate into immunogenic DCs and acquire the ability to secrete cytokines and to produce high levels of type I and type III interferons, thus priming T cells against viral antigens[117,127,170].

cDCs are a subset of DCs found in many lymphoid and non-lymphoid tissues. These cells upon antigen presentation to lymphocytes induce either immunity or tolerance toward that antigen[169]. There are consistent reports that further distinguish two population of cDCs, specifically cDC1 and cDC2, which play complementary functions in the delineating of immune responses[169]. Myeloid cDC1 have a high intrinsic capacity to cross-present antigens via MHC class I to activate $CD8^+$ T cells and to promote T helper type 1 (Th1) and natural killer responses through IL-12[127,171]. Regarding to the myeloid cDC2 population, these cells are equipped with a wide range of lectins, TLRs, NOD-like receptors and RIG-I-like receptors and respond well to LPS, flagellin, poly IC and R848. Although cDC2 cells are specialized in presentation to $CD4^+$ T cells, under appropriated activation, they also have the capacity of cross-presentation to CD8 cells[127].

Mo-DCs are a subset of DCs that are believed to be inflammatory DCs. These cells function primarily at the site of inflammation rather than travelling to lymph nodes[127]. In humans, inflammatory mo-DCs have been described for example, in eczema, psoriasis, allergic rhinitis, coeliac disease and inflammatory bowel disease[127].

In terms of genetic regulation, there is a specific repertoire of TFs able to regulate Mo-DC differentiation, including CEBP α and β, IRF4, STAT5 and RELB[122,167,172]. In addition, KLF4, a TF with

an important role in inducing MO lineage-associated molecules, is also considered a key switching factor regulating the differentiation of MO into Mo-DCs[167,173].

As mentioned above, the production of mo-DCs can be recapitulated *in vitro* upon stimulation of hematopoietic progenitor cells or MO with GM-CSF and IL4[167]. Under these circumstances, *in vitro* mo-DCs are characterized by upregulation of CD11c, HLA-DR, CD1a, DC-SIGN and downregulation of CD14 cell surface molecules[174].

Despite poor understanding of how the *in vitro* mo-DCs correlate with their *in vivo* counterparts, this *in vitro* system to differentiate peripheral MO into iDC represents an ideal model to investigate molecular processes occurring in the absence of cell divison like active DNA demethylation.

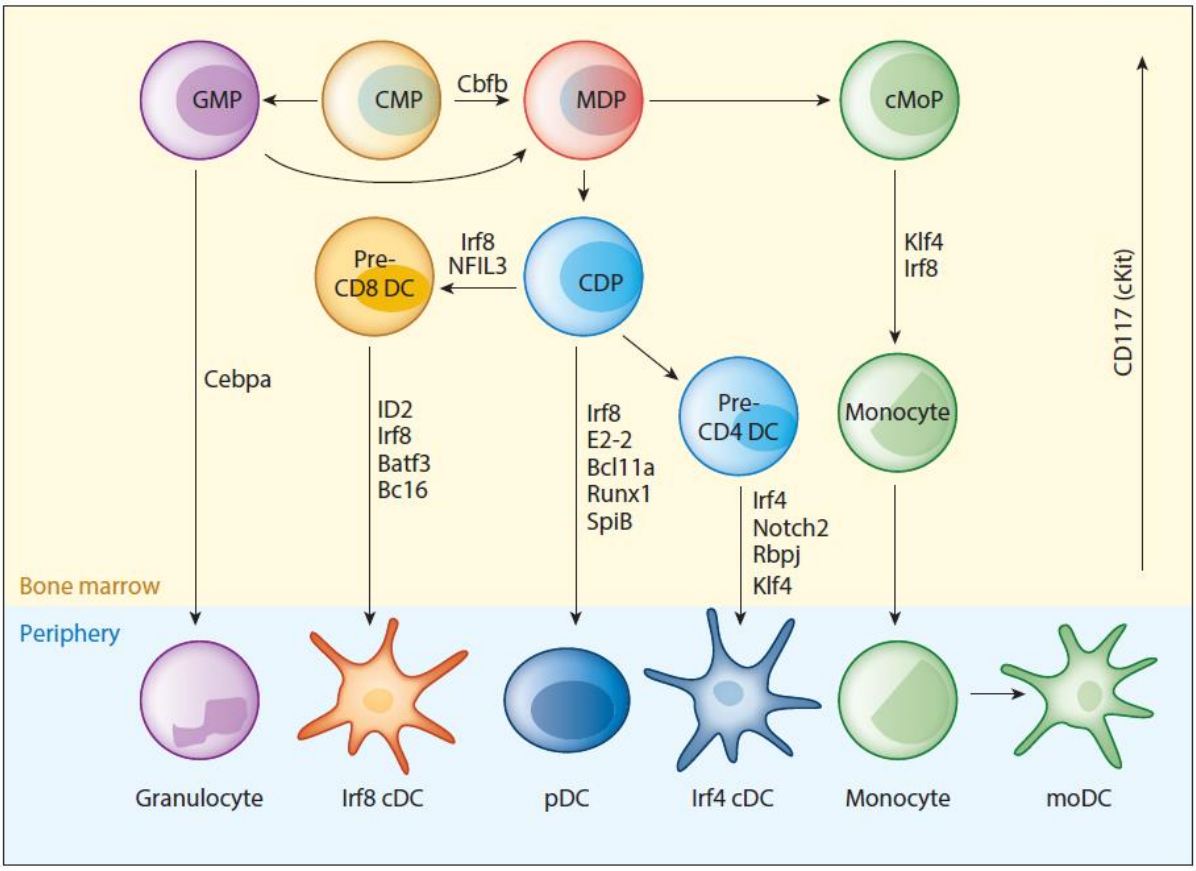

Figure 1.9 Stages and TFs in DC development

The figure depicts the myeloid lineage development from the CMP, showing TFs required for the appropriate transition between stages. Interestingly, Irf8 and Irf4 appear to be critical for Irf8 cDC (also known as cDC1), Irf4 cDC (also known as cDC2) and pDC commitment in the bone marrow. Relative level of cKit in progenitors is indicated by the vertical arrow. CDP, common dendritic cell progenitor; cMoP, committed monocyte progenitor; CMP, common myeloid progenitor; GMP, granulocyte-macrophage progenitor; MDP, macrophage-DC precursor (adapted from Murphy *et al*[172]).

2 Objectives

DNA methylation turnover plays an important role during myeloid cell differentiation. Its dysregulation, for example due to the mutation of key enzymes like DNMT3A or TET2, is associated with various myeloid malignancies.

This thesis particularly focuses on DNA demethylation processes that occur during the post-mitotic differentiation of human blood monocytes (MO). Previous work had established that TET2 is the main hydroxylase catalyzing the removal of 5mC during MO differentiation, but it has been unclear, how the demethylation machinery is recruited to its target sites, and how important this epigenetic process is for the differentiation of post-mitotic MO.

Hence, one main objective of this thesis was to establish a system to deplete TET2 in primary MO to study the consequences of TET2 depletion on local DNA methylation, chromatin accessibility and transcription in the course of MO differentiation. This should address one of the key questions in DNA methylation research: Is DNA demethylation necessary for the activation of previously DNA methylated cis-regulatory modules, or is it the simple consequence of other processes?

Another important unsolved question was, how the specificity of the enzymatic turnover is controlled. Previous work suggested that the oxidation of 5mC at regions undergoing active demethylation is an early event that may even precede the binding of transcription factors (TFs). To shed further light on the targeting mechanisms, the second major aim was to identify candidate TFs that may be involved in TET2 recruitment. Candidate factors should be identified from differentially methylated regions (DMRs) in published whole genome bisulfite sequencing data, validated using knock-down approaches and further characterized using high-throughput-based methods. Finally, biochemical approaches should be used to establish interaction partners of candidate TFs and to uncover targeting mechanisms of active DNA demethylation.

This thesis aimed at expanding our knowledge of a process that is fundamental to normal hematopoiesis. Findings may not only be important for the general understanding of MO biology but also for our understanding of how mutations in the DNA methylation machinery might alter hematopoietic cells to drive leukaemogenesis.

3 Material & Equipment

3.1 Equipment

Autoclave	Walter, Geislingen, Germany
Avanti J25	Beckman Coulter, Munich, Germany
Centrifuges	Heraeus, Hanau; Eppendorf, Germany
Electrophoresis equipment	Bio-Rad, Munich, Germany
Evos Digital Microscope	Thermo Fisher Scientific, Waltham, MA USA
Fusion Pulse	Vilber Lourmat, Eberhardzell, Germany
Gene Pulser Xcell	Bio-Rad, Munich, Germany
Heatblock	Eppendorf, Hamburg, Germany
Heat sealer	Eppendorf, Hamburg, Germany
Hemocytometer	Marienfeld, Lauda-Königshofen, Germany
Incubators	Heraeus, Hanau, Germany
J6M-E centrifuge	Beckmann, Munich, Germany
Laminar air flow cabinet	Heraeus, Hanau, Germany
Luminometer	Sirius Berthold, Oakville, Canada
MassARRAY Compact System	Sequenom, San Diego, CA, USA
MassARRAY MATRIX Liquid Handler	Sequenom, San Diego, CA, USA
MassARRAY Phusio chip module	Sequenom, San Diego, CA, USA
Mastercycler Nexus M2	Eppendorf, Hamburg, Germany
Megafuge 3.0 R	Heraeus, Osterode, Germany
Microscopes	Zeiss, Jena, Germany
Mini Centrifuge	Carl Roth, Karlsruhe, Germany
Mini-PROTEAN Tetra System	Bio-Rad, Munich, Germany
Multifuge 3S-R	Heraeus, Osterode, Germany
Multipipettor Multipette plus	Eppendorf, Hamburg, Germany
NanoDrop Spectrophotometer	PeqLab, Erlangen, Germany
NextSeq 550	Illumina, San Diego, USA
PCR-Thermocycler 4800	Perkin Elmer, Uberlingen, Germany
PCR-Thermocycler PTC-200	MJ-Research/Biometra,Oldendorf, Germany
PCR-Thermocycler Veriti 384-well	Applied Biosystems, Foster City, USA
pH-Meter	Knick, Berlin, Germany

Picofuge	Heraeus, Hanau, Germany
PowerPac Basic	Bio-Rad, Munich, Germany
Power supplies	Biometra, Gottingen, Germany
Qubit 2.0 Flurometer	Thermo Fisher Scientific, Waltham, USA
Realplex Mastercycler epGradientS	Eppendorf, Hamburg, Germany
Sonifier 250	Branson, Danbury, USA
TapeStation 2200	Agilent Technologies, Boblingen, Germany
Thermomixer	Eppendorf, Hamburg, Germany
Typhoon 9200	Molecular Dynamics, Krefeld, Germany
Vortex Genie	Scientific Industries, New York, USA
Water bath	Julabo, Seelstadt, Germany
Water purification system	Merck Millipore, Darmstadt, Germany
Whatman Fastblot B44	Biometra, Göttingen, Germany

3.2 Consumables

8-channel pipettor tips Impact 384	Thermo Fisher Scientific, Waltham, USA
Adhesive PCR sealing film	Thermo Fisher Scientific, Waltham, USA
Cell culture flasks and pipettes	Eppendorf, Hamburg, Germany
Cell culture plates (6-well)	Eppendorf, Hamburg, Germany
Centrifuge tubes (15, 50, 225 ml)	Falcon, Heidelberg, Germany
CLEAN resin	Sequenom, San Diego, CA, USA
Electroporation cuvettes (0.4 cm)	PeqLab, Erlangen, Germany
Heat sealing Film	Eppendorf, Hamburg, Germany
Loading tips for TapeStation 2200	Agilent Technologies, Böoblingen, Germany
Luminometer vials	Falcon, Heidelberg, Germany
MATRIX Liquid Handler D.A.R.Ts tips	Thermo Fisher Scientific, Hudson, NH, USA
Micro test tubes (0.2 ml)	Biozym Scientific, Oldendorf, Germany
Micro test tubes (0.5, 1.5, 2, 5 ml)	Eppendorf, Hamburg; Sarstedt, Numbrecht, Germany
Multiwell cell culture plates and tubes	Eppendorf, Hamburg, Germany
nProteinG Sepharose 4 FastFlow	GE Healthcare, Munich, Germany
PCR plate Twin.tec 96 well	Eppendorf, Hamburg, Germany
PCR plate 384 well (MassARRAY)	Thermo Fisher Scientific, Hudson, NH, USA

Petri dishes	Falcon, Heidelberg, Germany
Pierce™ Streptavidin Magnetic Beads	Thermo Fisher Scientific, Waltham, USA
Qubit assay tubes	Invitrogen, Carlsbad, USA
Sepharose Cl-4 beads	Sigma-Aldrich, Taufkirchen, Germany
Slide-A-Lyzer MINI Dialysis Devices	Thermo Fisher Scientific, Waltham, USA
Sterile combitips	Eppendorf, Hamburg, Germany
Sterile micropore filters	Merck Millipore, Darmstadt, Germany
Sterile plastic pipettes	Costar, Cambridge, USA
Syringes and needles	Becton Dickinson, Heidelberg, Germany
Teflon foils	Heraeus, Hanau, Germany

3.3 Chemicals

All reagents used were purchased from Sigma-Aldrich (Taufkirchen, Germany) or Merck Millipore (Darmstadt, Germany) unless otherwise mentioned.

3.4 Enzymes, Reagents & Kits

2-Mercaptoethanol	GE Healthcare, Chalfont St. Hiles, UK
Agencourt AMPure XP beads	Beckman Coulter, Krefeld, Germany
Alkaline Phosphatase	Thermo Fisher Scientific, Waltham, USA
Ammonium peroxodisulfate	MERCK, Darmstadt, Germany
Annexin V MicroBead Kit	Miltenyi Biotec, Gladbach, Germany
Beetle-Juice BIG KIT	PJK, Kleinblittersdorf, Germany
ECL Prime Western Blotting System	Sigma-Aldrich, Taufkirchen, Germany
Disuccinimidyl glutarate	Thermo Fisher Scientific, Waltham, USA
DNeasy Blood & Tissue Kit	Qiagen, Hilden, Germany
dNTP Mix	Agena Bioscience, San Diego, USA
dNTPs	Thermo Fisher Scientific, Waltham, USA
Dynabeads® Antibody Coupling Kit	Thermo Fisher Scientific, Waltham, USA
EZ DNA Methylation Kit	Zymo Research, Irvine, USA
Gibson Assembly Master Mix	NEB, Frankfurt, Germany
GM-CSF	Berlex, Seattle, USA
IL-4 (Human)	Promokine, Heidelberg, Germany
MinElute Gel Extraction Kit	Qiagen, Hilden, Germany
mMESSAGE mMACHINE T7 Ultra Kit	Thermo Fisher Scientific, Waltham, USA

Monarch DNA Gel Extraction Kit	NEB, Frankfurt, Germany
Monarch PCR & DNA Cleanup Kit	NEB, Frankfurt, Germany
Monarch Plasmid Miniprep Kit	NEB, Frankfurt, Germany
NEBNext Ultra II DNA Library Prep Kit	NEB, Ipswich, Massachusetts, USA
Nextera XT DNA Library Preparation Kit	Illumina, San Diego, USA
Nextera XT Index Kit v2	Illumina, San Diego, USA
NextSeq 550 High Output v2 kit (75 cycles)	Illumina, San Diego, USA
Phusion Hot Start High-Fidelity DNA Polymerase	Thermo Fisher Scientific, Waltham, USA
Plasmid Plus Midi Kit	Qiagen, Hilden, Germany
Protease inhibitor Cocktail (50x)	Thermo Fisher Scientific, Waltham, USA
Proteinase K	Thermo Fisher Scientific, Waltham, USA
QuantiFast SYBR Green Kit	Qiagen (Hilden, Germany)
ReBlot Plus Mild Solution, 10x	Merck Millipore, Darmstadt, Germany
Renilla-Juice BIG KIT	PJK, Kleinblittersdorf, Germany
Restriction endonucleases	NEB, Frankfurt; Roche, Penzberg, Germany
Reverse Transcriptase SuperScript II	Promega, Madison, USA
RNeasy Mini Kit	Qiagen, Hilden, Germany
ScriptSeq™ Complete Kit (Human/Mouse/Rat)	Epicentre, Chicago, USA
ScriptSeq™ Index PCR Primers	Epicentre, Chicago, USA
Shrimp Alkaline Phosphatase (SAP)	Sequenom, San Diego, CA, USA
T-Cleavage MassCleave Reagent kit	Sequenom, San Diego, CA, USA

3.5 Antibodies

Chromatin Immunoprecipitation (ChIP)

Anti-FLAG M2 (F3165)	Sigma-Aldrich, Taufkirchen, Germany

Western Blot

Anti-FLAG M2 (F3165)	Sigma-Aldrich, Taufkirchen, Germany
Anti-IRF4 (sc-6059X)	Santa Cruz, Heidelberg, Germany
Anti-EGR2 (sc-293195)	Santa Cruz, Heidelberg, Germany

Anti-TET2 (MABE462)	Merck, Darmstadt, Germany
Anti-Actin (A2066)	Sigma Aldrich, Taufkirchen, Germany

Co-immunoprecipitation (CoIP)

Anti-FLAG M2 (F3165)	Sigma-Aldrich, Taufkirchen, Germany
Anti-TET2 (MABE462)	Merck, Darmstadt, Germany
Anti-IgG (sc-2025)	Santa Cruz, Heidelberg, Germany
Anti-TET2 (A304-247A)	Biomol, Bethyl, Hamburg, Germany

Western Blot (Secondary Antibodies)

m-IgGk BP-HRP (sc-516102)	Santa Cruz, Heidelberg, Germany
Goat anti-rabbit (P0448)	Dako, Boblingen, Germany
Rabbit Anti-goat (P0449)	Dako, Boblingen, Germany

3.6 Antibiotics

Antibiotic	Stock concentration	Final concentration	Company
Ampicillin	100 mg/ml	100 µg/ml	Roth, Karlsruhe, Germany

3.7 Plasmids

pEF6/V5-His Topo	Thermo Fisher Scientific, Waltham, USA
psiCHECK™-2	Promega, Madison, USA

3.8 *E.coli* Strains

DH10B	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL nupG /pMON14272 / pMON7124

3.9 Cell lines

THP-1	Human acute monocytic leukemia (DSMZ ACC 16)

3.10 Oligonucleotides

Oligonucleotides were synthesized, purified and purchased from Sigma-Aldrich (Taufkirchen, Germany). T7 Promoter and BGH rev sequencing primers were obtained from Geneart (Thermo Fisher Scientific, Regensburg).

3.10.1 Real-time PCR primers

ACTB_1 forw	5'-CGAGAAGATGACCCAGATCATGTTTGAG-3'
ACTB_1 rev	5'-CAGAGGCGTACAGGGATAGCACAG-3'
IRF4_2_forw	5´-TGGGAGAACGAGGAGAAGAGCA-3´
IRF4_2_rev	5´-AAACAGTGCCCAAGCCTTGAAGAG-3´

3.10.2 Sequencing primers

T7 Promoter	5'-TAATACGACTCACTATAGGG-3'
BGH rev	5'-TAGAAGGCACAGTCGAGG-3'
3xFLAG_EGR2_biot_forw	5´-CTTTTCCCGATGATTCCAGA-3´
3xFLAG_EGR2_biot_rev	5´-CTGGATGTGCTTGTTGATGG-3´

3.10.3 EpiTYPER primers

Amplicon		**Primer Sequence (5´ - 3´)[1]**
SQ00031_HCRTR1	**10F**	aggaagagagTTTTGTAGATTAGTGAGTGAGTGGAAGG
	T7R	cagtaatacgactcactatagggagaaggctCCAAAAAAAACAACCCCTAAAATTC
SQ00044	**10F**	aggaagagagGGGTGATGTAGGGGTGAATTTTATT
	T7R	cagtaatacgactcactatagggagaaggctAACAAACACCTACCCAAAAACCC
SQ00060_PPCDC	**10F**	aggaagagagGGTATTTTGTGGTTGATTTTTTTGG
	T7R	cagtaatacgactcactatagggagaaggctACCAAACTAATATTAAACTCCTAACCTCAA
SQ00069_ITGAM	**10F**	aggaagagagTTTGTGTGTAAGTGTGTGTATATGTGTG
	T7R	cagtaatacgactcactatagggagaaggctAAAAAAAATTCCTAATTCCACCCAC
SQ00080_RAB15	**10F**	aggaagagagTGGTAATTAGGGTAGAAGGATAATGGTTAA
	T7R	cagtaatacgactcactatagggagaaggctTCTCCTCCTAACACAAACACAAACTC
Epi00109_CCL13.1	**10F**	aggaagagagTTTGTGGTTTGAATAGTTAGAAGGA
	T7R	cagtaatacgactcactatagggagaaggctCAACAAACACAAAAACACTACAAAAA
Epi00148_C1ORF78.3	**10F**	aggaagagagGGAATTTTGTTATTTTTTAGGGTGG
	T7R	cagtaatacgactcactatagggagaaggctAAAACCACCATCCTCTAACTCTC
SQ00004_TGM2_02	**10F**	aggaagagagTAGGAATTTTTATTGTTGGGTGGAGT
	T7R	cagtaatacgactcactatagggagaaggctTAAAATATAAATAAAATCCCCCACCCTTCT
SQ00012_SLC25A47	**10F**	aggaagagagTGGAGGGTTTGTTTGTGAGTTAGG
	T7R	cagtaatacgactcactatagggagaaggctACCCTATTTCCCCTCTAACTTCCCTATA
SQ00016_BAX_02	**10F**	aggaagagagTTGGGGGTTTTAGTTTATTTTTTTT
	T7R	cagtaatacgactcactatagggagaaggctACCTAAATCCAACTCTTTAATACCC
SQ00001_AK021415_01	**10F**	aggaagagagTGGGTTTTGGAATTTTATTATTTGGG
	T7R	cagtaatacgactcactatagggagaaggctAAACTCACATAAATCCATCTCCTCC
SQ00006_AK309785_01	**10F**	aggaagagagTGTTAGGTTGTTTTTGTTTTTTGATAAGTT
	T7R	cagtaatacgactcactatagggagaaggctAAATCCCCTACTTCCATTCACAATC
SQ00010_AK024310_01	**10F**	aggaagagagGGGTAGGATAGTGTATTGTAGTGGAAGT

	T7R	cagtaatacgactcactatagggagaaggctCCTTAATAATCAAAAAATAACAACAAAAC
SQ00012_DOCK2_01	**10F**	aggaagagagTGGTGAGAAGAGGTTGTTTGTTGTT
	T7R	cagtaatacgactcactatagggagaaggctAATTAATCCCCACAATCAAAATTCATTATA
SQ00018_ITGAM_03_2	**10F**	aggaagagagATTTTTTAGTGGGTATTTTTATTGGGTATT
	T7R	cagtaatacgactcactatagggagaaggctAAAAAAAACTACCTAAACTCCCTTATAC
SQ00021_JD117533_01	**10F**	aggaagagagAGAAGTAAATAAATGTTGAGAATGTGTTAG
	T7R	cagtaatacgactcactatagggagaaggctTCCAAAACTCCATACAATAAAAATAC
SQ00023_ITPR2_01_2	**10F**	aggaagagagTGAGTAAGGGGTAGGAAGGTTATTTT
	T7R	cagtaatacgactcactatagggagaaggctAACTTTTAAAACCCAAAACTTTATACTTCA
SQ00029_LOC541471	**10F**	aggaagagagTTTGTGGGTGTTTGGGTTTTTTTAT
	T7R	cagtaatacgactcactatagggagaaggctATTCCAAAATTTTATCCTCTTCAAATATCA
SQ00047_TBC1D1_01	**10F**	aggaagagagGTTATAGGTAGGTTGAGGTTAGATTTGGAA
	T7R	cagtaatacgactcactatagggagaaggctAAATCCCATCAAAATTTCAATAAATC
SQ00001_CABP4_01	**10F**	aggaagagagTGTTTTAGAAGTTTAGGTAGATGATTAGGT
	T7R	cagtaatacgactcactatagggagaaggctCTCCTTAAACTACCAACCAATAAATTAATA
SQ00005_TMIGD3_01	**10F**	aggaagagagTTTTGTTTGTTTATTAATTTTGGAGGTTTA
	T7R	cagtaatacgactcactatagggagaaggctCCTAAAAAATAATTCAATAATTCCAACTTT
SQ00008_MAN1C1_02	**10F**	aggaagagagTTTTATTAGTATAAAGGGGTTTGTTTT
	T7R	cagtaatacgactcactatagggagaaggctTCCAAATAACCTAATAATCTAAAATTCTAA
SQ00012_MYRF_03	**10F**	aggaagagagGGGGAGGGGTAGTTAATGTTTGAGT
	T7R	cagtaatacgactcactatagggagaaggctAAAAAAACTAATTCTATAACTAATCCCAAA

[1]Lowercase: reverse primers (T7R) tagged with the T7 promoter sequence; forward primers (10F) tagged with a 10 mer overhang. Uppercase: primer sequences

3.10.3.1 Genomic location

Amplicon	**Genomic location (GRCh38/hg38)**
SQ00031_HCRTR1	chr1:31625344-31625648
SQ00044	chr8:143399602-143400034
SQ00060_PPCDC	chr15:75051084-75051292
SQ00069_ITGAM	chr16:31332112-31332381
SQ00080_RAB15	chr14:64965677-64965967
Epi00109_CCL13.1	chr17:34356259-34356559
Epi00148_C1ORF78.3	chr9:129839252-129839471
SQ00004_TGM2_02	chr20:38141132-38141388
SQ00012_SLC25A47	chr14:100325124-100325387
SQ00016_BAX_02	chr19:48961384-48961645
SQ00001_AK021415_01	chr10:75070419-75070720
SQ00006_AK309785_01	chr2:218005061-218005375
SQ00010_AK024310_01	chr14:103889639-103889892
SQ00012_DOCK2_01	chr5:169805716-169805911
SQ00018_ITGAM_03_2	chr16:31332112-31332381

SQ00021_JD117533_01	chr8:8949947-8950168
SQ00023_ITPR2_01_2	chr12:26728155-26728400
SQ00029_LOC541471	chr2:111239165-111239537
SQ00047_TBC1D1_01	chr4:37954547-37954793
SQ00001_CABP4_01	chr11:67452163-67452604
SQ00005_TMIGD3_01	chr1:111491097-111491464
SQ00008_MAN1C1_02	chr1:25695856-25696293
SQ00012_MYRF_03	chr11:61749586-61750085

3.11 siRNAs (small interfering RNAs)

Chemically modified siRNAs were designed, synthesized, purified and purchased from Axolabs (Kulmbach, Germany). siRNAs for initial testing were delivered in ultrapure water at a concentration of 50 µM. Tested siRNAs were then delivered as lyophilized powder that was dissolved in ultrapure water at a concentration of 10 µg/µl (stock solution) and stored at -80 °C.

IRF4 siRNA sequences (for testing in THP-1 cell line)

siRNA	Strand	Sequence[1]
siRNA_1406	Sense	5´-gcuAcGAuuuAccaGAAcAdTsdT-3´
	Antisense	5´-UGUUCUGGuAAAUCGuAGCdTsdT-3´
siRNA_2102	Sense	5´-agccucAcAcGuAaAAGAAdTsdT-3´
	Antisense	5´-UUCUUUuACGUGUGAGGCUdTsdT-3´
siRNA_2309	Sense	5´-gaAGccAGuuAGuaAAcuudTsdT-3´
	Antisense	5´-AAGUUuACuAACUGGCUUCdTsdT-3´
siRNA_3724	Sense	5´-ccAAGcGGAuGcuccAuuudTsdT-3´
	Antisense	5´-AAAUGGAGcAUCCGCUUGGdTsdT-3´
siRNA_4223	Sense	5´-guGAccGAcucAuuuAcAAdTsdT-3´
	Antisense	5´-UUGuAAAUGAGUCGGUcACdTsdT-3´
siRNA_4228	Sense	5´-cgAcucAuuuAcAacuGAAdTsdT-3´
	Antisense	5´-UUcAGUUGuAAAUGAGUCGdTsdT-3´
siRNA_4316	Sense	5´-uauuGGGuAuGAAcuAAAAdTsdT-3´
	Antisense	5´-UUUuAGUUcAuACCcAAuAdTsdT-3´
siRNA_4449	Sense	5´-gaGcGAGGGcAuAaAuAcAdTsdT-3´
	Antisense	5´-UGuAUUuAUGCCCUCGCUCdTsdT-3´

[1]Chemical modification pattern:
A, G, U, C: RNA Nucleotide
a, g, u, c: 2´-O-Methyl-Nucleotide
dT: desoxy-T residue
s: Phosphorothioate

siRNA sequences (siRNAs used in knock-down experiments)

siRNA	Strand	Sequence[1]
siRNA_CTRL_luciferase	Sense	5´-cuuAcGcuGAGuAcuucGAdTsdT-3´
	Antisense	5´-UCGAAGuACUcAGCGuAAGdTsdT-3´
siRNA_TET2_1212	Sense	5´-accucAGGGcAGAucAAuudTsdT-3´
	Antisense	5´-AAUUGAUCUGCCCUGAGGUdTsdT-3´
siRNA_EGR2_1132 (a)	Sense	5´-cucuAcAAuccGuaAcuuudTsdT-3´
	Antisense	5´-AAAGUuACGGAUUGuAGAGdTsdT-3´
siRNA_EGR2_2665 (b)	Sense	5´-guAAAuGGGuuGccuuAuudTsdT-3´
	Antisense	5´-AAuAAGGcAACCcAUUuACdTsdT-3´
siRNA_IRF4_2931 (a)[2]	Sense	5´-caGGAuAuuuAcuauuAcudTsdT-3´
	Antisense	5´-AGuAAuAGuAAAuAUCCUGdTsdT-3´
siRNA_IRF4_2384 (b)[2]	Sense	5´-uguccuAcAAucuaGuAAudTsdT-3´
	Antisense	5´-AUuACuAGAUUGuAGGAcAdTsdT-3´

[1]Chemical modification pattern:
A, G, U, C: RNA Nucleotide
a, g, u, c: 2´-O-Methyl-Nucleotide
dT: desoxy-T residue
s: Phosphorothioate

[2]siRNAs with the best silencing efficiency after testing (sequences not shown in the previous siRNA (IRF4) list)

3.12 Molecular Weight Standards

1 kb Plus DNA Ladder	Thermo Fisher Scientific, Waltham, USA
50 bp DNA Ladder	NEB, Frankfurt
Precision Plus Protein Kaleidoscope Standards (10–250 kD)	Bio-Rad, Munich, Germany

3.13 gBlocks® Gene Fragments

gBlocks gene fragments were synthesized, high pressure liquid chromatography purified and purchased from Integrated DNA Technologies (IDT, San Jose, USA).

gBlocks Gene Fragments - Luciferase Reporter Assays

psiCHECK2-IRF4.1 (5´- 3´)

TGCATCTCGAGGAACCTCTGCTAGCCAGACAACTATATTATTTTGCTCAACAAAACAGTGGACATTTCCTGAGGGGCTACGATTTACCAGAACACATCAGCA
ATCCAGAAGATTACCACAGGCGGCCGCGTGCA

psiCHECK2-IRF4.2 (5´- 3´)

TGCATCTCGAGTGTAAATTGAAGAAGCCTCACACGTAAAAGAAATGTATTAATGTATGTAGGAGCTGCAGTTCTTGTGGAAGACACTTGCTGAGTGAAGGAA
ATGAATCTTTGACTGAAGCCGTGCCTGTAGCCTTGGGGAGGCCCATCCCCCACCTGCCAGCGGTTTCCTGGTGTGGGTCCCTCTGCCCCGCCCTCCTTCCCA
TTGGCTTTCTCTCCTTGGCCTTTCCTGGAAGCCAGTTAGTAAACTTCCTATTTTCTTGAGTCAAAAAACATGAGCGCTACTCTTGGATGGGACATTTTTGTC
TGTCCTACAATCTAGTAATGTCTAAGTAATGGTTAAGTTTTCTTGTTTCTGCATCTTTTTGACCCTCATTCTTTAGAGATGCTAAAATTCTTCGCATAAAGA
AGAAGAAATTAAGGAACATAAATCTTAATACTTGAACTGTTGCCCTTCTGTCCAAGTACTTAACTATCTGTTCCCTTCCTCTGTGCCACGCTCCTCTGTTTG
TTTGGCTGTCCAGCGATCAGCCATGGCGACACTAAAGGAGGAGGAGCCGGGGACTCCCAGGCTGGAGAGCACTGCCAGGACCCACCACTGGAAGCAGGATGG
AGCTGACTACGGAACTGCACACTCAGTGGGCTGTTTCTGCTTATTTCATCTGTTCTATGCTTCCTCGTGCCAATTATAGTTTGACAGGGCCTTAAAATTACT

TGGCTTTTTCCAAATGCTTCTATTTATAGAATCCCAAAGACCTCCACTTGCTTAAGTATACCTATCACTTACATTTTTGTGGTTTTGAGAAAGTACAGCAGT
AGACTGGGGCGTCACCTCCAGGCCGTTTCTCATACTACAGGATATTTACTATTACTCCCAGGATCAGCAGGCGGCCGCGTGCA

psiCHECK2-IRF4.3 (5´- 3´)

TGCATCTCGAGTGTCTTCTTAATTCTCCAAGCGGATGCTCCATTTCAATTGCTTTGTGACTTCTTCTTCTTTGTTTTTTTAAATATTATGCTGCTTTAACAG
TGGAGCTGAATTTTCTGGAAAATGCTTCTTGGCTGGGGCCACTACCTCCTTTCCTATCTTTACATCTATGTGTATGTTGACTTTTTAAAATTCTGAGTGATC
CAGGGTATGACCTAGGGAATGAACTAGCTATGAAATACTCAGGGTTAGGAATCCTAGCACTTGTCTCAGGACTCTGAAAAGGAACGGCTTCCTCATTCCTTG
TCTTGATAAAGTGGAATTGGCAAACTAGAATTTAGTTTGTACTCAGTGGACAGTGCTGTTGAAGATTTGAGGACTTGTTAAAGAGCACTGGGTCATATGGAA
AAAATGTATGTGTCTCCCAGGTGCATTTCTTGGTTTATGTCTTGTTCTTGAGATTTTGTATATTTAGGAAAACCTCAAGCAGTAATTAATATCTCCTGGAAC
ACTATAGAGAACCAAGTGACCGACTCATTTACAACTGAAACCTAGGAAGCCCCTGAGTCCTGAGCGAAAACAGGAGAGTTAGTCGCCCTACAGAAAACCCAG
CTAGACTATTGGGTATGAACTAAAAAGAGACTGTGCCATGGTGAGAAAAATGTAAAATCCTACAGTGAAATGAGCAGCCCTTACAGTATTGTTACCACCAAG
GGCAGGTAGGTATTAGTGTTTGAAAAAGCTGGTCTTTGAGCGAGGGCATAAATACAGCTAGCCCCAGGGGTGGGCGGCCGCGTGCA

gBlocks Gene Fragments – ChIP-seq and/or CoIP experiments

pEF6-3xFLAG-EGR2 (5´- 3´)

Fragment 1:

TATAGGGAGACCCAAGCTGGCTAGGTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCG
ATTACAAGGATGACGATGACAAAATGACCGCCAAGGCCGTAGACAAAATCCCAGTAACTCTCAGTGGTTTTGTGCACCAGCTGTCTGACAACATCTACCCGG
TGGAGGACCTCGCCGCCACGTCGGTGACCATCTTTCCCAATGCCGAACTGGGAGGCCCCTTTGACCAGATGAACGGAGTGGCCGGAGATGGCATGATCAACA
TTGACATGACTGGAGAGAAGAGGTCGTTGGATCTCCCATATCCCAGCAGCTTTGCTCCCGTCTCTGCACCTAGAAACCAGACCTTCACTTACATGGGCAAGT
TCTCCATTGACCCTCAGTACCCTGGTGCCAGCTGCTACCCAGAAGGCATAATCAATATTGTGAGTGCAGGCATCTTGCAAGGGGTCACTTCCCCAGCTTCAA
CCACAGCCTCATCCAGCGTCACCTCTGCCTCCCCCAACCCACTGGCCACAGGACCCCTGGGTGTGTGCACCATGTCCCAGACCCAGCCTGACCTGGACCACC
TGTACTCTCCGCCACCGCCTCCTCCTCCTTATTCTGGCTGTGCAGGAGACCTCTACCAGGACCCTTCTGCGTTCCTGTCAGCAGCCACCACCTCCACCTCTT
CCTCTCTGGCCTACCCACCACCTCCTTCCTATCCATCCCCCAAGCCAGCCACGGACCCAGGTCTCTTCCCAATGATCCCAGACTATCCTGGATTCTTTCCAT
CTCAGTGCCAGAGAGACCTACATGGTACAGCT

Fragment 2:

AGACTATCCTGGATTCTTTCCATCTCAGTGCCAGAGAGACCTACATGGTACAGCTGGCCCAGACCGTAAGCCCTTTCCCTGCCCACTGGACACCCTGCGGGT
GCCCCCTCCACTCACTCCACTCTCTACAATCCGTAACTTTACCCTGGGGGGCCCCAGTGCTGGGGTGACCGGACCAGGGGCCAGTGGAGGCAGCGAGGGACC
CCGGCTGCCTGGTAGCAGCTCAGCAGCAGCAGCAGCCGCTGCCGCAGCTGCCTATAACCCACACCACCTGCCACTGCGGCCCATTCTGAGGCCTCGCAAGTA
CCCCAACAGACCCAGCAAGACGCCGGTGCACGAGAGGCCCTACCCGTGCCCAGCAGAAGGCTGCGACCGGCGGTTCTCCCGCTCTGACGAGCTGACACGGCA
CATCCGAATCCACACTGGGCATAAGCCCTTCCAGTGTCGGATCTGCATGCGCAACTTCAGCCGCAGTGACCACCTCACCACCCATATCCGCACCCACACCGG
TGAGAAGCCCTTCGCCTGTGACTACTGTGGCCGAAAGTTTGCCCGGAGTGATGAGAGGAAGCGCCACACCAAGATCCACCTGAGACAGAAAGAGCGGAAAAG
CAGTGCCCCCTCTGCATCGGTGCCAGCCCCCTCTACAGCCTCCTGCTCTGGGGGCGTGCAGCCTGGGGGTACCCTGTGCAGCAGTAACAGCAGCAGTCTTGG
CGGAGGGCCGCTCGCCCCTTGCTCCTCTCGGACCCGGACACCTTGATAATCTAGAGGGCCCGCGGTTCGAAGGTAAGCCTATCCCTAACCCTCTCCTC

pEF6-3xFLAG-IRF4 (5´- 3´)

ATTGCGGATCCGCCACCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAAATGAACCTGGAGGGCG
GCGGCCGAGGCGGAGAGTTCGGCATGAGCGCGGTGAGCTGCGGCAACGGGAAGCTCCGCCAGTGGCTGATCGACCAGATCGACAGCGGCAAGTACCCCGGGC
TGGTGTGGGAGAACGAGGAGAAGAGCATCTTCCGCATCCCCTGGAAGCACGCGGGCAAGCAGGACTACAACCGCGAGGAGGACGCCGCGCTCTTCAAGGCTT
GGGCACTGTTTAAAGGAAAGTTCCGAGAAGGCATCGACAAGCCGGACCCTCCCACCTGGAAGACGCGCCTGCGGTGCGCTTTGAACAAGAGCAATGACTTTG
AGGAACTGGTTGAGCGGAGCCAGCTGGACATCTCAGACCCGTACAAAGTGTACAGGATTGTTCCTGAGGGAGCCAAAAAAGGAGCCAAGCAGCTCACCCTGG
AGGACCCGCAGATGTCCATGAGCCACCCCTACACCATGACAACGCCTTACCCTTCGCTCCCAGCCCAGCAGGTTCACAACTACATGATGCCACCCCTCGACC
GAAGCTGGAGGGACTACGTCCCGGATCAGCCACACCCGGAAATCCCGTACCAATGTCCCATGACGTTTGGACCCCGCGGCCACCACTGGCAAGGCCCAGCTT
GTGAAAATGGTTGCCAGGTGACAGGAACCTTTTATGCTTGTGCCCCACCTGAGTCCCAGGCTCCCGGAGTCCCCACAGAGCCAAGCATAAGGTCTGCCGAAG
CCTTGGCGTTCTCAGACTGCCGGCTGCACATCTGCCTGTACTACCGGGAAATCCTCGTGAAGGAGCTGACCACGTCCAGCCCCGAGGGCTGCCGGATCTCCC
ATGGACATACGTATGACGCCAGCAACCTGGACCAGGTCCTGTTCCCCTACCCAGAGGACAATGGCCAGAGGAAAAACATTGAGAAGCTGCTGAGCCACCTGG
AGAGGGGCGTGGTCCTCTGGATGGCCCCCGACGGGCTCTATGCGAAAAGACTGTGCCAGAGCAGGATCTACTGGGACGGGCCCCTGGCGCTGTGCAACGACC
GGCCCAACAAACTGGAGAGAGACCAGACCTGCAAGCTCTTTGACACACAGCAGTTCTTGTCAGAGCTGCAAGCGTTTGCTCACCACGGCCGCTCCCTGCCAA
GATTCCAGGTGACTCTATGCTTTGGAGAGGAGTTTCCAGACCCTCAGAGGCAAAGAAAGCTCATCACAGCTCACGTAGAACCTCTGCTAGCCAGACAACTAT
ATTATTTTGCTCAACAAAACAGTGGACATTTCCTGAGGGGCTACGATTTACCAGAACACATCAGCAATCCAGAAGATTACCACAGATCTATCCGCCATTCCT
CTATTCAAGAATGATAATCTAGAGTCAT

pEF6-3xFLAG-NAB2 (5´- 3´)

TATAGGGAGACCCAAGCTGGCTAGGTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCG
ATTACAAGGATGACGATGACAAAATGCATAGGGCTCCCAGCCCGACTGCTGAACAGCCTCCCGGGGGAGGCGACAGTGCAAGAAGAACCTTGCAACCGAGAT
TGAAACCCAGTGCCCGCGCCATGGCCCTGCCGCGGACTCTGGGAGAGTTGCAACTTTACCGCGTGCTGCAACGAGCCAATCTGTTGTCCTACTATGAAACGT
TTATCCAGCAGGGTGGAGACGATGTTCAGCAGTTGTGCGAAGCAGGTGAGGAAGAATTTCTCGAGATTATGGCGCTTGTAGGTATGGCAACGAAGCCGCTTC
ACGTGCGAAGGCTTCAGAAAGCATTGCGAGAATGGGCGACCAACCCGGGTTTGTTTAGTCAGCCCGTACCTGCTGTACCTGTCTCTAGTATTCCCCTTTTCA
AGATTAGCGAGACGGCTGGTACGCGCAAGGGATCAATGAGTAATGGGCATGGCTCTCCCGGAGAGAAGGCAGGTAGCGCGCGGAGCTTTTCTCCAAAGTCCC
CGTTGGAACTTGGTGAAAAGTTGTCTCCACTTCCAGGTGGGCCGGGTGCCGGAGATCCAAGAATCTGGCCAGGACGCTCCACACCAGAAAGTGATGTCGGGG
CGGGTGGCGAAGAGGAAGCCGGTTCACCACCATTCTCCCCTCCAGCTGGTGGAGGTGTACCTGAAGGCACCGGAGCCGGCGGGCTGGCGGCAGGTGGAACCG
GTGGTGGTCCCGATCGCTTGGAGCCTGAAATGGTCCGGATGGTTGTGGAATCTGTAGAACGGATCTTTCGATCTTTCCCAAGGGGTGATGCGGGGGAAGTCA
CTAGTTTGTTGAAGCTGAACAAGAAACTCGCTAGATCTGTGGGCCATATATTCGAGATGGATGATAACGATTCTCAAAAGGAGGAGGAAATCAGAAAATACT
CTATAATTTACGGCCGATTCGACTCTAAGCGGCGGGAAGGAAAGCAGCTGTCATTGCATGAACTTACTATTAACGAAGCCGCCGCACAATTCTGCATGCGGG
ATAACACCTTGCTGCTGCGCAGAGTCGAACTCTTCAGTTTGAGCCGGCAGGTCGCTCGAGAGTCAACCTACCTGTCCTCACTCAAAGGGTCCAGACTGCACC

CAGAAGAACTCGGGGGGCCCCCACTCAAAAAGCTTAAGCAGGAGGTGGGGGAACAATCTCACCCCGAGATCCAGCAACCTCCTCCCGGTCCAGAGAGCTATG
TGCCCCCATATCGCCCGAGTCTTGAGGAGGATAGCGCCTCCCTCAGTGGTGAGTCCTTGGACGGACACCTTCAGGCAGTGGGGTCCTGTCCTCGACTTACTC
CTCCGCCAGCAGACCTGCCACTGGCGCTGCCCGCTCATGGCCTGTGGAGTAGACACATTCTCCAACAAACTCTTATGGATGAGGGATTGCGACTGGCTCGGC
TTGTCTCCCACGACAGGGTGGGCCGGCTCTCACCGTGCGTACCAGCGAAACCCCCGCTGGCTGAATTTGAAGAAGGCCTCCTCGATAGGTGCCCCGCGCCCG
GCCCACACCCTGCGCTTGTTGAGGGTAGGCGATCCAGCGTAAAGGTTGAAGCAGAGGCCTCCAGACAATGATCTAGAGGGCCCGCGGTTCGAAGGTAAGCCT
ATCCCTAACCCTCTCCTC

gBlocks Gene Fragments – BioID experiments

pEF6-3xFLAG-EGR2-linker-BirA* (5´- 3´)

TATAGGGAGACCCAAGCTGGCTAGGTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCG
ATTACAAGGATGACGATGACAAAATGATGACTGCCAAAGCGGTGGACAAGATTCCAGTTACACTCTCCGGATTTGTACACCAATTGAGTGATAACATATACC
CTGTCGAGGATCTCGCAGCCACATCCGTTACTATATTCCCAAATGCTGAATTGGGTGGGCCGTTTGACCAGATGAACGGAGTCGCCGGTGATGGAATGATTA
ACATTGATATGACCGGTGAGAAGAGATCACTTGACTTGCCATATCCTTCCAGTTTCGCCCCAGTATCCGCTCCCAGGAACCAAACCTTCACTTACATGGGTA
AGTTCTCAATTGATCCGCAGTATCCAGGAGCTAGTTGCTATCCTGAAGGAATAATCAACATAGTTTCTGCTGGCATATTGCAAGGTGTTACCTCACCTGCGT
CCACTACAGCGTCCTCTAGTGTAACCTCAGCCAGCCCGAATCCTCTCGCGACGGGCCCACTGGGGGTATGTACAATGTCCCAGACTCAGCCTGATTTGGATC
ACCTTTACTCTCCACCACCCCCGCCTCCACCCTACAGTGGTTGCGCAGGGGACTTGTATCAAGATCCGAGTGCATTCCTGTCTGCCGCGACAACATCTACCT
CCTCAAGTCTCGCCTACCCACCTCCTCCCTCCTACCCCTCTCCCAAGCCAGCCACCGACCCGGGGCTTTTCCCGATGATTCCAGATTACCCTGGTTTCTTTC
CATCACAATGCCAGAGGGATCTCCACGGGACTGCCGGACCAGACAGAAAGCCTTTCCCATGCCCATTGGACACACTTCGAGTGCCACCGCCATTGACGCCTC
TTTCCACAATTCGAAATTTCACACTCGGAGGTCCAAGTGCGGGCGTCACCGGCCCTGGTGCCAGTGGCGGATCAGAGGGACCGCGCCTCCCGGGTTCCAGTA
GTGCGGCAGCCGCTGCAGCTGCCGCTGCTGCATATAATCCGCATCACCTTCCACTGCGACCCATACTGAGACCGCGCAAGTACCCAAATAGGCCGTCCAAGA
CCCCTGTTCACGAACGCCCATACCCATGTCCAGCGGAGGGATGTGACCGCCGGTTTTCTCGCAGTGATGAACTCACTAGGCACATTCGGATCCATACAGGTC
ATAAACCATTTCAATGCAGGATATGCATGCGGAATTTCAGTAGGAGCGACCATCTGACGACTCACATTAGAACGCACACTGGAGAAAAGCCATTCGCATGTG
ACTACTGTGGTCGCAAGTTCGCACGGTCCGACGAGCGGAAGCGACACACAAAGATTCATCTCCGCCAGAAAGAGCGCAAAAGTAGCGCTCCATCCGCTTCAG
TCCCGGCTCCTAGCACGGCAAGTTGTAGCGGTGGAGTTCAGCCTGGAGGGACATTGTGCTCCTCAAACTCCAGTAGCCTCGGAGGAGGTCCACTTGCCCCAT
GTAGCAGCAGAACCCGGACGCCTACTATGTCTGTGGACTCAAGCCTGCCCAGCCCCAACCAGCTGAGCAGCCCCAGCCTGGGTTTCGACGGCCTGCCCGGCC
GGAAGGACAACACCGTGCCCCTGAAGCTGATCGCCCTGCTGGCCAACGGCGAGTTCCACTCTGGCGAGCAGCTGGGAGAGACCCTGGGAATGAGCAGAGCCG
CCATCAACAAGCACATCCAGACACTGAGAGACTGGGGAGTGGACGTGTTCACCGTGCCTGGCAAGGGCTACAGCCTGCCTGAGCCTATCCAGCTGCTGAACG
CCAAGCAGATCCTGGGACAGCTGGATGGCGGAAGCGTGGCCGTGCTGCCTGTGATCGACTCCACCAATCAGTACCTGCTGGACAGAATCGGAGAGCTGAAGT
CCGGCGACGCCTGCATCGCCGAGTACCAGCAGGCTGGCAGAGGAGGCAGAGGACGGAAGTGGTTCAGCCCATTCGGAGCCAACCTGTACCTGTCCATGTTCT
GGAGACTGGAGCAGGGACCTGCTGCTGCCATCGGACTGAGTCTGGTGATCGGAATCGTGATGGCCGAGGTGCTGAGAAAGCTGGGAGCCGACAAGGTGAGAG
TGAAGTGGCCTAATGACCTGTACCTCCAGGACCGCAAGCTGGCTGGCATCCTGGTGGAGCTGACAGGCAAGACAGGCGATGCCGCTCAGATCGTGATCGGAG
CCGGAATCAACATGGCCATGAGAAGAGTGGAGGAGAGCGTGGTGAACCAGGGCTGGATCACCCTGCAGGAGGCTGGCATCAACCTGGACCGGAACACCCTGG
CCGCCATGCTGATCAGAGAGCTGAGAGCCGCTCTGGAGCTGTTCGAGCAGGAGGGACTGGCTCCTTACCTGAGCAGATGGGAGAAGCTGGACAACTTCATCA
ACAGACCTGTGAAGCTGATCATCGGCGACAAGGAAATCTTCGGCATCTCCAGAGGAATCGACAAGCAGGGAGCTCTGCTGCTGGAGCAGGACGGAATCATCA
AGCCCTGGATGGGCGGAGAAATCTCCCTGAGAAGCGCAGAGAAGCTCGAGTGATCTAGAGGGCCCGCGGTTCGAAGGTAAGCCTATCCCTAACCCTCTCCTC

pEF6-3xFLAG-NLS-BirA* (5´- 3´)

TATAGGGAGACCCAAGCTGGCTAGGTAAGCTTGGTACCGAGCTCGGATCCGCCACCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCG
ATTACAAGGATGACGATGACAAACCGAAAAAGAAGCGCAAGGTTAAGGACAACACCGTGCCCCTGAAGCTGATCGCCCTGCTGGCCAACGGCGAGTTCCACT
CTGGCGAGCAGCTGGGAGAGACCCTGGGAATGAGCAGAGCCGCCATCAACAAGCACATCCAGACACTGAGAGACTGGGGAGTGGACGTGTTCACCGTGCCTG
GCAAGGGCTACAGCCTGCCTGAGCCTATCCAGCTGCTGAACGCCAAGCAGATCCTGGGACAGCTGGATGGCGGAAGCGTGGCCGTGCTGCCTGTGATCGACT
CCACCAATCAGTACCTGCTGGACAGAATCGGAGAGCTGAAGTCCGGCGACGCCTGCATCGCCGAGTACCAGCAGGCTGGCAGAGGAGGCAGAGGACGGAAGT
GGTTCAGCCCATTCGGAGCCAACCTGTACCTGTCCATGTTCTGGAGACTGGAGCAGGGACCTGCTGCTGCCATCGGACTGAGTCTGGTGATCGGAATCGTGA
TGGCCGAGGTGCTGAGAAAGCTGGGAGCCGACAAGGTGAGAGTGAAGTGGCCTAATGACCTGTACCTCCAGGACCGCAAGCTGGCTGGCATCCTGGTGGAGC
TGACAGGCAAGACAGGCGATGCCGCTCAGATCGTGATCGGAGCCGGAATCAACATGGCCATGAGAAGAGTGGAGGAGAGCGTGGTGAACCAGGGCTGGATCA
CCCTGCAGGAGGCTGGCATCAACCTGGACCGGAACACCCTGGCCGCCATGCTGATCAGAGAGCTGAGAGCCGCTCTGGAGCTGTTCGAGCAGGAGGGACTGG
CTCCTTACCTGAGCAGATGGGAGAAGCTGGACAACTTCATCAACAGACCTGTGAAGCTGATCATCGGCGACAAGGAAATCTTCGGCATCTCCAGAGGAATCG
ACAAGCAGGGAGCTCTGCTGCTGGAGCAGGACGGAATCATCAAGCCCTGGATGGGCGGAGAAATCTCCCTGAGAAGCGCAGAGAAGCTCGAGTGATCTAGAG
GGCCCGCGGTTCGAAGGTAAGCCTATCCCTAACCCTCTCCTC

3.14 Databases & Software

Adobe Illustrator	Adobe, San Jose, California, USA
Bcl2fastq v2.20.0.422	Illumina, San Diego, USA
BEDtools2 v2.27.1	Quinlan Laboratory, Virginia, USA
BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Bowtie2 v2.3.4	www.bowtie-bio.sourceforge.net/index.shtml
Enrichr	http://amp.pharm.mssm.edu/Enrichr/
EpiTYPER v1.2, agena	Sequenom, San Diego, CA, USA
ExPASy	http://web.expasy.org/
FastQC v0.11.7	https://www.bioinformatics.babraham.ac.uk/projects/download.html
GENtle v1.9.4	University of Cologne, Cologne, Germany
HOMER v4.11	http://homer.ucsd.edu/homer/index.html
Illumina Experiment Manager v1.14.0	Illumina Inc., San Diego, USA
Illumina Sequencing Analysis Viewer v2.1.18	Illumina Inc., San Diego, USA
Inkscape 0.92.5	https://inkscape.org/
Integrative Genomics Viewer (IGV) v2.4.6	Broad Institute, Cambridge, USA
IGVTools v2.3.98	Broad Institute, Cambridge, USA
NCBI	https://www.ncbi.nlm.nih.gov/
Metascape	http://metascape.org/
MethPrimer	http://www.urogene.org/methprimer
Microsoft Office 2016	Microsoft Corporation, Redmond, USA
PubMed	www.ncbi.nlm.nih.gov/entrez
R 3.4.3	http://www.r-project.org/
R Studio v0.97.551	http://www.rstudio.com/
SAMtools v1.9	https://sourceforge.net/projects/samtools/files/samtools
SnapGene Viewer v2.8.2	GSL Biotech LLC, Chicago, USA
STAR v2.5.3a	https://github.com/alexdobin/STAR
STRING v11.0	https://string-db.org/
TapeStation A.01.05 (SR1)	Agilent Technologies, Boblingen, Germany
Typer 4.0	Sequenom, San Diego, CA, USA
UCSC Genome Browser	www.genome.ucsc.edu

4 Methods

4.1 Cell Culture

4.1.1 THP-1 Cell line

THP-1 cell line was cultured at a density of $0,5x10^6$/ml in RPMI-1640 (Gibco, Waltham, USA), supplemented with 10% FCS (fetal calf serum), L-glutamine (2 mM), sodium pyruvate (1 mM), antibiotics (50 U/ml penicillin, 50 µg/ml streptomycin), vitamins (2 ml), non-essential amino acids (1x) and ß-mercaptoethanol (50 µM). Media supplements were purchased from Gibco (Waltham, USA) and Merck Millipore (Darmstadt, Germany; L-glutamine), respectively. FCS was heat inactivated for 20 min at 56°C before use. Each batch of FCS or culture medium was tested before use.
Cell passaging was performed every 2-3 days in fresh medium and cells were maintained at general cell culture conditions in a standard incubator (37°C, 5% CO_2 and 95% relative humidity).

4.1.2 Isolation of Primary Human Monocytes by Elutriation

Peripheral blood mononuclear cells (PBMCs) from healthy donors were enriched and collected by leukapheresis, followed by density gradient centrifugation over Ficoll/Hypaque. MO were then isolated from PBMCs by counter current centrifugal elutriation. Elutriation was performed in a J6M-E centrifuge equipped with a JE 5.0 elutriation rotor and a 50 ml flow chamber (Beckman, Munich, Germany). After sterilizing the system with 6% H_2O_2 for 20 min, the system was washed with 1x PBS. Following calibration at 2500 rpm with Hanks solution, MNCs were loaded at a flow rate of 52 ml/min. Fractions were sequentially collected according with their respective flow rate (see Table 4.1). MO from the last fraction were centrifuged (300xg, 4°C, 8 min) and washed with 1x PBS. After counting, MO (purity ranging from 85% to 95%) were resuspended in 40 ml of RPMI medium.

Table 4.1 Elutriation parameters and cell types

Fraction	Volume (ml)	Flow rate (ml/min)	Main cell type recovered
Ia	1000	52	platelets
Ib	1000	57	B and T lymphocytes, NK cells
IIa	1000	64	
IIb	500	74	
IIc	400	82	
IId	400	92	
III	800	130	MO

4.1.3 IL4/GM-CSF-mediated Monocyte Differentiation

Peripheral human blood MO were cultured at a density of $1x10^6$/ml in RPMI-1640 (Thermo Fisher Scientific, Waltham, USA) containing the supplements aforementioned in section 4.1.1. To promote the differentiation of MO towards iDCs, culture medium was supplemented with 20 U/ml recombinant

human IL-4 (Promokine, Heidelberg, Germany) and 280 U/ml GM-CSF (Berlex, Seattle, USA). Cells were kept in a standard incubator at 37°C, 5% CO_2 and 95% of relative humidity. iDCs were cultured for 7 days and harvested at different time points (18 h, 66 h, 4 d and 7 d) during differentiation, unless otherwise stated.

4.1.4 Assessing Cell Number and Viability

The total number of cells and their viability was determined by Trypan blue exclusion. For this purpose, the cell suspension was diluted in Trypan blue solution (0.2% (w/v) trypan blue in 0.9 % NaCl solution) and immediately counted in a Neubauer hemocytometer. When counting the cells in the microscope, the dead cells appear dark blue and thus easily distinguishable from the live cells. The concentration of viable cells was determined according with the following formula:

Number of viable cells/ml (C): $\mathbf{C = N \times D \times 10^4}$

N: average number of unstained cells per corner square (1 mm^2 containing 16 sub-squares)

D: dilution factor

4.1.5 Transient Transfections of Primary Human Cells and THP-1 Cell line

Immune cells activation after transfection is one of the main challenges when introducing foreign genetic material into these cells. Therefore, during transfection (by electroporation) were used mRNAs generated by *in vitro* transcription (IVT) (see section 4.5.3), which were reported to be well tolerated by DCs[175] and backbone modified siRNAs (see section 3.11) with no effect on cell activation or survival.

The siRNAs when injected into a cell, they form double stranded RNA molecules with the target mRNAs, thus blocking their further processing and translation into a protein. Consequently, siRNAs were used to transiently knock-down (KD) the expression of their target genes. In contrast, the ivt mRNAs were used to overexpress the fusion proteins, previously generated by molecular cloning (see section 4.3).

For electroporation, cells were initially washed one time with RPMI (Gibco, Thermo Fisher Scientific, Waltham, USA; without phenol red) and Opti-MEM (Gibco, Thermo Fisher Scientific, Waltham, USA; without phenol red), followed by centrifugation (500xg, 10 min, RT) between washes. Afterwards, cells were gently resuspended in Opti-MEM at a final concentration of $3x10^6$ cells in a final volume of 200µl, unless otherwise mentioned. Per electroporation in a cuvette (0.4cm electroporation cuvettes (PeqLab, Erlangen, Germany)) were combined 200 µl of resuspended cells with 3 µg siRNA (1 µg siRNA/$1x10^6$ cells) or with 5 µg (3xFLAG-EGR2; 3xFLAG-IRF4) and 3 µg (3xFLAG-PU.1) of ivt mRNA (for ChIP-seq experiments). Of note, for siRNA testing using the luciferase reporter assay, siRNAs at a final concentration of 40 pmol or 120 pmol were used. For analysis of EGR2, Cterm-3xFlag-EGR2-BirA*, 3xFlag-NLS-BirA* and NAB2 fusion proteins kinetic were used 20 µg, 32 µg, 12 µg and 22 µg

of ivt mRNA (per 12x10^6 cells), respectively. For BioID (proximity-dependent biotin identification) and CoIP (co-immunoprecipitation) procedures the amounts of ivt mRNA utilized are given in sections 4.6.4 & 4.6.5, respectively.

Using the electroporator GenePulser Xcell (BioRAD), cells were electroporated according with the following conditions: 3 square waves, 400 V, 5 ms, 1 pulse interval. Immediately after electroporation, cells were placed in cell flasks containing pre-warmed RPMI medium supplemented as described in section 4.1.1 for THP-1 and in section 4.1.3 for iDCs. Cells were cultured at a density of 1x10^6/ml in a standard incubator (37°C, 5% of CO2, 95% of relative humidity) and harvested at the respective time points.

4.1.6 Luciferase Reporter Assay

For transfection, 1 µg of plasmid DNA (fused *Renilla* luciferase:gene of interest) together with 40 or 120 pmol of the respective siRNA were introduced in THP-1 cells by electroporation. Twenty-four hours post-transfection, cells were transferred to 14 ml polystyrene round-bottom tubes, centrifuged at 300×g for 10 min and washed with 1x PBS. After, cells were lysed by adding 300 µl of 1x lysis buffer (Promega) and incubated for 15 min at RT. Then, cell lysate was precleared by centrifugation and the supernatant was used to measure the luciferase and firefly activities using the Dual-Luciferase Reporter Assay System (Promega) according with the manufacturer's instructions. For data analysis, *Renilla* luciferase values were normalized to firefly luciferase values.

4.2 Bacterial Culture

4.2.1 Cultivation of *E.coli* strains

Required solutions:

LB-medium:	10 g Bacto Tryptone 10 g NaCl 5 g Yeast extract Add ddH2O up 1000 ml, adjust pH to 7.5, autoclave
LB-agar plates:	10 g Bacto Tryptone 10 g NaCl 5 g Yeast extract 15 g Agar Add ddH2O up 1000 ml, adjust pH to 7.5, autoclave, cool to 50°C and add the appropriate antibiotic. Pour the agar

solution into 10 cm petri dishes and store upside down at 4°C

E.coli strains were streaked out in solid LB-agar including the appropriate antibiotics (see section 3.6) to create a selection pressure for the bacteria growth, followed by overnight incubation at 37°C. In the next day, the single colonies in the LB-agar plates were picked out and transferred to LB medium supplemented with the same antibiotics as used before, and grown overnight at 37°C with shaking at 200 rpm. The bacteria suspensions containing the correct constructs were long-term stored at -80°C in 20% glycerol, by adding 600 µl liquid culture to 200 µl of 80% glycerol.

4.3 Molecular Cloning: upstream and downstream experiments

4.3.1 Restriction Endonuclease Digestion

Appropriated restriction enzymes were used to confirm the presence and orientation of an individual insert within a plasmid, due to their ability to cleave the DNA at specific sequences. Besides that, restriction enzymes were also used to linearize plasmid DNA for downstream applications, namely *in vitro* mRNA transcription (see section 4.5.3). In general, 1 µg DNA was digested with 1 U enzyme in a 10-50 µl reaction by incubation at 37°C or 65°C for 1 h. All the restriction enzymes and respective buffers were purchased from Sigma-Aldrich (Taufkirchen, Germany) or New England Biolabs (NEB, Frankfurt, Germany).

4.3.2 Dephosphorylation of DNA with Alkaline Phosphatase

To prevent self-ligation, digested vectors were treated with 1 U of AP (calf intestinal alkaline phosphatase) at 37°C for 30 min.

4.3.3 Agarose Gel Electrophoresis

Required buffers:

TAE (50x):	242.3 g (2 M)	Tris
	20.5 g (250 mM)	NaOAc/HOAc, pH 7.8
	18.5 g (0.5 M)	EDTA, pH 8.0
	Add ddH2O to 1000 ml	
DNA loading dye (DNA LD 5x):	500 µl (50 mM)	Tris (1 M), pH 7.8
	500 µl (1%)	SDS (20%)
	1 ml (50 mM)	EDTA (0.5 M), pH 8.0
	4 ml (40%)	Glycerol (100%)

10 g (1%)	Bromophenol blue
Add ddH2O to 10 ml, store at 4°C	

Agarose gel electrophoresis was routinely used to separate DNA fragments based on their size and charge. The agarose gel percentage varied between 0.8% - 2% depending on the fragments size (usually from 100 bp to 25 kb). The required agarose amount was weighted and added to the correct amount of 1x TAE buffer. The slurry was heated in a microwave until the agarose was completely dissolved. After cooling the solution to 50°C – 60°C, the gel was stained with Ethidium bromide (10 mg/ml) and poured into a gel tray. After agarose polymerization, the gel was mounted in the electrophoresis tank and covered with 1x TAE running buffer. Before loading the DNA samples into the gel, they were diluted 4:1 with DNA loading dye (5x). Depending on the size and the desired resolution, gels were run at 40-150 V for 30 min to 3 h.

4.3.4 Purification of DNA Fragments by Gel Extraction

The DNA fragments of interest were run in an agarose gel at a low voltage and for a longer time to allow more precise bands separation. After visualization of the gel under UV illumination, the bands containing the desired fragments were excised. Fragments were then purified by gel extraction using the Monarch DNA Gel Extraction Kit (NEB, Frankfurt, Germany) according to the manufacturer's instructions.

4.3.5 Gibson Assembly

There are different cloning methods to insert recombinant DNA into a vector. One of these methods is named Gibson assembly® and was initially described by Gibson *et al.*[176], allowing the generation of DNA constructs in a single round of cloning without using restriction enzymes. This method facilitates the assembly of multiple DNA fragments (up to 6 different fragments ranging from 500 bp to 32 kb) in a single isothermal reaction.

For the assembly reaction, the gBlocks Gene Fragments (chemically synthesized, double-stranded DNA, see section 3.13) and the linearized vector (see section 4.3.4) were required. In order to assemble the gBlocks Gene Fragments into the vector, overlapping sequences (between 30 to 40 bp) with the vector and between gene fragments (in case of multiple DNA fragments for assembly) were also included in the ordered gBlocks.

Per reaction, around 50-100 ng of linearized plasmid DNA and a 2-3-fold molar excess of insert fragments was used. Reactions were prepared as shown in Table 4.2 and according with manufacturer's instructions (NEB):

Table 4.2 Gibson Assembly setup and parameters

Number of fragments plus plasmid	2 - 3 fragments
Quantity	0.02-0.5 pmole ea.
Gibson Assembly Master Mix (2x), NEB	10 µl
ddH2O	Add to 20 µl

The reactions were incubated for 1 h at 50°C, resulting in different DNA fragments joined together. An overview and detailed protocol description of this method can be found in the instruction manual of Gibson Assembly® Master Mix from NEB.

4.3.6 Transformation of Chemically Competent *E.coli*

Required Medium:

SOC medium:	20 g	(2%)	BactoTrypton (Difco)
	5 g	(0.5%)	BactoYeastExtract (Difco)
	0.6 g	(10 mM)	NaCl
	0.2 g	(3 mM)	KCl
	Add ddH2O up 1L, autoclave and add:		
	10 ml	(10 mM)	MgCl2 (1 M, sterile)
	10 ml	(10 mM)	MgSO4 (1 M, sterile)
	10 ml	(20 mM)	Glucose (2 M, sterile)

Per transformation, 50 µl of chemically competent *E.coli* were gently mixed on ice with 1-25 ng plasmid DNA, following incubation on ice for 30 min. Cells were heat-shocked in a heat block at 42°C for 90 sec and cooled on ice for 2 min before adding 250 µl of pre-warmed SOC (Super Optimal broth with Catabolite repression) medium. Bacteria were vigorously shacked at 37°C for 1 h, then 50 µl to 100 µl of transformation reaction were plated in a LB-agar including ampicillin. The plates were incubated overnight at 37°C, thus generating single colonies.

4.3.7 Plasmid DNA Isolation

Single colonies from LB-agar plates and resistant to ampicillin (see section 4.3.6) were picked out and placed in 4 ml LB-medium (supplemented with ampicillin). After shaking overnight at 37°C, 2 ml liquid bacterial culture were used for plasmid isolation with the Monarch Plasmid Miniprep Kit (NEB, Frankfurt, Germany) according to manufacturer's instructions. To isolate larger amounts of ultrapure DNA (e.g. transfection experiments), plasmids were isolated using the endotoxin-free QIAGEN Plasmid Plus Midi Kit (Hilden, Germany) according to manufacturer's instructions.

4.3.8 Sanger Sequencing and Data Analysis

After plasmid DNA isolation (see section 4.3.7), the insert DNA sequence was checked by the Sanger sequencing method, which was performed by Thermo Fisher (Regensburg, Germany). To analyse and align the sequence files, the GENtle software or the BLAST function of the UCSC genome browser were used.

4.3.9 Purification of DNA using Phenol-Chloroform Extraction and Ethanol Precipitation

Before using the plasmid DNA for *in vitro* mRNA transcription (see section 4.5.3), it was initially linearized with restriction enzymes (see section 4.3.1) and purified according with the following protocol. Linearized plasmids DNA were vigorously mixed with 1 Volume (V) Phenol-Chloroform-Isoamylalcohol (25:24:1, pH 8) and centrifuged for 5 min at 13000 rpm (RT). The upper aqueous phase containing the DNA was carefully transferred to a new tube. The precipitation was repeated twice with 1 V of Chloroform p.a. (Merck, Darmstadt, Germany) to remove residual phenol. After extraction, the DNA-containing aqueous phase was transferred into a new RNase-free Eppendorf cup and precipitated with 0.5 V 5 M NH_4OAc (pH 5.2) and 2.5 V 100% EtOH for 1 h at -80°C. After centrifugation at 13000 rpm for 30 min (4°C), the supernatant was removed and the precipitated DNA was washed with 70% ice-cold EtOH. Pellet was air-dried and dissolved in RNase-free water at a concentration of 1 µg/µl. Purified linearized DNA was stored at 4°C for later use.

4.4 DNA-based Methods

4.4.1 Isolation and Quantification of Genomic DNA

Genomic DNA was isolated using the DNeasy Blood & Tissue Culture Kit (Qiagen) according to the manufacturer's instructions. Genomic DNA was quantified with the NanoDrop (peqLab) and used for DNA methylation analysis (see section 4.4.3).

4.4.2 Real-Time PCR

This method was used to evaluate the knock-down efficiency of different siRNAs targeting a specific gene. In the PCR step, a SYBR green dye was used for real time quantification of double-stranded DNA product in the reaction. PCR was performed using the QuantiFast SYBR Green Kit from Qiagen (Hilden, Germany) in 96-well plates adapted to the Eppendorf Realplex Mastercycler EpGradient S (Eppendorf, Hamburg, Germany) according with the following setup (Table 4.3) and PCR program (Table 4.4).

Table 4.3 Reaction setup for real-time PCR

Component	Volume [µl]	Final Concentration
SYBR Green mix (2x)	5	1x
H2O	2	-
Primer forward (10 µM)	0.5	0.5 µM
Primer reverse (10 µM)	0.5	0.5 µM
DNA	2	

Table 4.4 PCR program for real-time PCR

Stage	Temperature	Time	Cycles
Initial melting	95°C	5 min	
Melting	95°C	8 sec	45
Annealing & Extension	60°C	20 sec	
Melting	95°C	15 sec	Final cycle
Annealing & Extension	60°C	15 sec	
Melting curve		10 - 20 min	
	95°C	15 sec	

In addition to the unknown samples, the standard curve dilutions (1:10, 1:50, 1:100, 1:1000) were also measured and later used for quantification purposes. The Realplex software automatically calculated DNA amounts based on the generated slope and intercept parameters. Specific amplification was controlled in the melting-curve analysis. The data was exported from the software and analysed using Excel. All samples were measured in duplicate and for relative quantification they were normalized to β-actin (housekeeping gene).

4.4.3 DNA Methylation Analysis using the MassARRAY® System (Sequenom)

The MassARRAY system is a matrix-assisted laser desorption/ionization time of flight mass spectrometer (MALDI-TOF MS) used to detect and quantify DNA methylation of multiple CpGs in genomic regions between 100 to 600 bp.

This technology is formally known as EpiTYPER, a bisulfite-based method involving several biochemical steps. Briefly, the bisulfite converted DNA is amplified by PCR followed by treatment with a shrimp alkaline phosphatase (SAP) to dephosphorylate unincorporated dNTPs from PCR. After, the T7 promoter-tagged reverse primer used in the PCR reaction allows the *in vitro* transcription of the PCR product by T7 RNA polymerase, yielding a single stranded RNA product. Then, the RNA is cleaved by a RNase A enzyme, resulting in a specific fragmentation of the RNA, and the digestion products are separated on mass with the mass spectrometer (MALDI-TOF MS). Depending on the number of methylated CpG sites within a RNA fragment, the difference in mass results in a 16 Da shift in the mass spectrum.

A complete description of the method can be found in Ehrich *et al.*[177] or in the EpiTYPER User Guide (www.sequenom.com).

Primer Design

The MethPrimer web tool (http://www.urogene.org/methprimer/) was used to design primers capable of amplifying bisulfite converted DNA. To allow the reverse transcription by T7 RNA polymerase, all reverse primers were tagged at the 5' end with the T7 promoter sequence; a 10 mer overhang was added to all forward primers to balance for melting temperature differences (see section 3.10.3). Primers were ordered in 96-well format at 100 µM concentration (Sigma-Aldrich).

Bisulfite Treatment

The treatment of genomic DNA with sodium bisulfite leads to deamination of all unmethylated cytosines residues to uracil, whereas the methylated cytosines (5-methylcytosine) remain unaffected. Consequently, the 5-methylcytosines are transcribed to guanines, while unmethylated cytosines are transcribed to adenines.

Bisulfite conversion was performed with the EZ DNA Methylation Kit (Zymo Research) according to the manufacturer's instructions. The protocol was optimized with the following conditions as shown in the Table 4.5.

Table 4.5 Cycling protocol for bisulfite treatment

Cycle step	Temperature	Time	Number of cycles
Denaturation	95°C	30 sec	20 x
Sulfonation	50°C	15 min	
Cooling	4°C	Hold	-

PCR, SAP treatment, *in vitro* transcription, RNA specific cleavage and MALDI-TOF mass spectrometry were performed according to the EpiTYPER application guide and protocols (Sequenom).

EpiTYPER Analysis

Raw data was processed using the EpiTYPER software (v1.2, Sequenom) generating quantitative results for each RNA product.

DNA methylation ratios across amplicons and samples were exported to a text file and further analyses were performed using R software.

Heatmaps presenting averaged CpG methylation levels across amplicons and samples were plotted using the heatmap2 function of the gplots package in R. Barplots of individual CpGs across samples were plotted in R using the ggplot2 package.

4.5 RNA-based Methods

4.5.1 Isolation of Total RNA and Quality Control

Total RNA was isolated and purified using the Qiagen RNeasy Mini Kit (Hilden, Germany) according with manufacturer's instructions. RNA was quantified with the NanoDrop (peqLab) and the quality was assessed using the RNA ScreenTape Kit (Agilent, Böblingen, Germany) according to the manufacturer's instructions. The RNA was then used for reverse transcription PCR (section 4.5.2) or RNA sequencing (section 4.7.3).

4.5.2 Reverse Transcription PCR

In the reverse transcription reaction a transcriptase is used to transcribe RNA to complementary DNA (cDNA), allowing for the quantification of mRNA transcripts. For this purpose, the reaction was setup as shown in Table 4.6:

Table 4.6 Reaction setup for reverse transcription PCR

Component	Amount
RNA	1 µg
Random Decamers	1 µl
dNTP's (10 mM)	1 µl
H_2O	15 µl

After incubation at 65°C for 5 min, samples were shortly placed on ice, followed by the addition of 4 µl of 5x MMLV (Moloney Murine Leukemia Virus Reaction Buffer) (Promega) and further incubation for 2 min at 42°C. While the samples were in the cycler, 1 µl of MMLV Reverse Transcriptase (Promega) was added to each sample, followed by an initial incubation at 42°C for 50 min and then at 70°C for 15 min. The synthesized cDNA was stored at -20°C and later used for quantitative real-time PCR (see section 4.4.2).

4.5.3 *In Vitro* Synthesis of Capped mRNA

The *in vitro* synthesis of single stranded RNA (capped and polyadenylated*)* was performed using the mMESSAGE mMACHINE T7 Ultra Kit from Thermo Fisher Scientific (Waltham, USA) according to manufacturer's instructions. In general, the T7 2 x NTP/ARCA and 10 x T7 Reaction Buffer were thawed at RT to prevent precipitation. The following components were pipetted according to the listed order in Table 4.7.

Table 4.7 Reaction setup for *in vitro* mRNA transcription

Component	Amount
H_2O (Nuclease free)	x µl (add to 20 µl)
T7 2 x NTP/ARCA	10 µl
10 x T7 Reaction Buffer	2 µl
Linear DNA template (1 µg)	x µl
T7 Enzyme Mix	2 µl

The reaction was gently mixed and incubated for 1-3 h (depending on the transcript size) in a heat block at 37°C. In the next step, 1 µl TURBO DNase was added and the reaction was further incubated for 30 min at 37°C. For poly (A)-tailing production, the specific reaction components were mixed as shown in Table 4.8.

Table 4.8 Reaction setup for poly (A)-tailing production

Component	Amount
mMessage mMachine T7 Ultra reaction	20 µl
H_2O (Nuclease free)	36 µl
5 x E-PAP Buffer	20 µl
25 mM $MnCl_2$	10 µl
ATP solution	10 µl
E-PAP enzyme	4 µl

The reaction was incubated for 1 h at 37°C yielding a poly (A)-tail between 50 to 100 bases. Synthesized transcripts were then purified using the Qiagen RNeasy Mini Kit (Hilden, Germany) following manufacturer's instructions, with slight modifications. The transcripts were initially mixed with 350 µl of RLT buffer and 250 µl of absolute ethanol. After loading the sample into the column and centrifugation at 11000 rpm for 30 sec (RT), the samples were washed with 500 µl of RW1 buffer and twice with 500 µl of RPE buffer, always followed by centrifugation (11000 rpm for 30 sec, RT) between washes. At the end, the samples were eluted with 50 µl of RNase free water and centrifuged at 13000 rpm, for 1 min, RT.

After purification, transcripts were quantified using the NanoDrop (peqLab) and the quality was assessed using the RNA ScreenTape Kit (Agilent, Böblingen, Germany) according to the manufacturer's instructions.

4.6 Protein-based Methods

4.6.1 Preparation of Whole Cell Lysates

Required buffer:

SDS Sample Buffer (2x):	10 ml (150 mM)	Tris (1.5 M), pH 6.8
	6 ml (1.2%)	SDS (20%)
	30 ml	Glycerol
	15 ml	β-mercaptoethanol
	1.8 mg	Bromophenol blue
	Add ddH_2O to 100 ml; aliquot in 10 ml stock solution (-20°C) and working solution (4°C)	

For whole cell lysates preparation, cells were washed with 1x PBS and after centrifugation, the supernatant was completely removed. The cell pellet containing around $3x10^6$ cells was resuspended in 200 µl of SDS sample buffer (2x) and the samples immediately incubated in a heat block at 95°C for 10 min, shaking at 1000 rpm. After incubation, cell lysates were shortly vortexed and stored at -80°C for use in downstream applications, as SDS-PAGE (see section 4.6.2) and Western blotting (see section 4.6.3).

4.6.2 Discontinuous SDS-PAGE

Required solutions and buffers:

Tris/HCl, pH 8.8:	90.83 g (1.5 M)	
	Add ddH_2O to 500 ml	
Tris/HCl, pH 6.8:	30 g (0.5 M)	
	Add ddH_2O to 500 ml	
SDS (10%):	100 g (10%)	SDS
	Add ddH_2O to 1000 ml, adjust pH to 7.2	
Ammonium Persulfate (APS):	1 g (10%)	APS
	Add ddH_2O to 10 ml	
Laemmli Buffer (5x):	15 g (40 mM)	Tris

216 g (0.95 M)	Glycine
15 g (0.5%)	SDS
Add ddH_2O to 3000 ml	

Protein samples were separated in a discontinuous gel system consisting of two layers of gel with different salt and polyacrylamide concentrations. Depending on the size of the proteins to be resolved, different acrylamide (AA) gel concentrations were prepared. Gels were prepared in the day before, as described in Table 4.9 & Table 4.10. After pouring the first gel (separating gel) into the chamber, isopropanol (70%) was immediately added on the top to facilitate the generation of a straight gel. After polymerization for 30 min, the isopropanol was removed and the stacking gel was added and a comb inserted into the top. After polymerization for 30 min, the gels were wrapped in wet paper and kept at 4°C.

Table 4.9 Separating gel preparation using different acrylamide (AA) concentrations

Component	Final AA concentration 12%	Final AA concentration 10%	Final AA concentration 8%
Tris/ HCl pH 8.8	5 ml	5 ml	5 ml
SDS (10%)	200 µl	200 µl	200 µl
AA (30%)	8 ml	6.66 ml	5.33 ml
H_2O	6.6 ml	7.92 ml	9.25 ml
TEMED	20 µl	20 µl	20 µl
APS (10%)	200 µl	200 µl	200 µl
Final Volume	20 ml	20 ml	20 ml

Table 4.10 Stacking gel preparation using 5% acrylamide (AA) concentration

Component	Final AA concentration 5%
Tris/ HCl pH 6.8	630 µl
SDS (10%)	50 µl
AA (30%)	830 µl
H_2O	3.4 ml
TEMED	5 µl
APS (10%)	50 µl
Final Volume	5 ml

Gel Running

The gel was mounted in the electrophoresis tank, which was filled out with 1x Laemmli buffer. The protein samples (20-40 µl) were loaded into the gel and in parallel a precision plus standard was included in order to determine the size of the detected proteins. The gel was initially run at 80 V for approximately 30 min or until the bands reached the surface of the stacking gel. Then, the voltage was increased to 120 V and the gel was run for an additional 1-2 h until the bromophenol blue was detected nearly to the gel bottom.

4.6.3 Western Blot

Required solutions and buffers:

Anode Buffer A:	36.3 g (0.3 M)	Tris
	200 ml (20%)	Methanol
	Add ddH_2O to 1000 ml	
Anode Buffer B:	3.03 g (25 mM)	Tris
	200 ml (20%)	Methanol
	Add ddH_2O to 1000 ml	
Cathode Buffer C:	5.20 g (4 mM)	ε-Amino-n-capron acid
	200 ml (20%)	Methanol
	Add ddH_2O to 1000 ml	
TBS (10x):	45.8 g (100 mM)	Tris/HCl, pH 8
	175.5 g (1.5 M)	NaCl
	Add ddH_2O to 2000 ml	
Washing Buffer (1x TBST):	100 ml	TBS (10x)
	1 ml (0.05%)	Tween-20
	Add ddH_2O to 1000 ml	
Blocking Buffer:	5.0 g (5%)	nonfat dried milk
	100 ml	TBS

Blotting

Proteins were blotted electrophoretically onto a PVDF membrane (merck) using a three-buffer semi-dry system. First, three Whatman3MM filter paper soaked with buffer A (bottom, on the anode) were placed in the blotting machine, followed by another three Whatman3MM filter paper soaked with buffer B. The hydrophobic PVDF membrane was first immersed in isopropanol and after in buffer B solution and then placed on the top of the previous paper layers. 600 µl of buffer B solution were spread out on the top of the membrane and the acrylamide gel containing the resolved proteins was added to the top. At the end three Whatman 3MM filter papers soaked with buffer C were placed on top of the gel followed by the cathode. Protein transfer was performed for 1 h at 11 V.

Washing and Staining

After blotting, the membranes were washed for 3 min with water and immediately blocked with 5% non-fat milk in 1x TBS for 1 h at RT (shaking at 45-60 rpm). After blocking, membranes were washed and incubated with anti-TET2 (MABE462, Merck, 1:1000), anti-IRF4 (sc-6059X, Santa Cruz, 1:5000), anti-EGR2 (sc-293195, Santa Cruz, 1:200), anti-FLAG M2 (F3165, Sigma, 1:2000) or anti-Actin (A2066, Sigma, 1:2000) overnight at 4°C. After washing one time for 10 min with 1x TBS, the membranes were incubated for 1 h at RT with a horseradish-peroxidase (HRP)-coupled secondary antibody (Goat anti-rabbit, P0448, Dako, 1:5000; Rabbit Anti-goat, P0449, Dako, 1:5000; m-IgGk BP-HRP, sc-516102, Santa Cruz, 1:5000). Then, membranes were washed eight times with TBST before using the ECL Prime Western Blotting System (Sigma-Aldrich, Taufkirchen, Germany) to visualize the bound antibody. Blots were developed using the Fusion Pulse imaging system from Vilber Lourmat (Eberhardzell, Germany). For re-blotting, blots were stripped with 1 x ReBlot Plus Mild Antibody Stripping Solution (Merck) for 15 min at RT.

4.6.4 Proximity-dependent Biotin Identification (BioID)

Required solutions and buffers:

Cell buffer Mix (CBM):	0.5 ml (10 mM)	HEPES/KOH (1 M), pH 7.9
	4.25ml (85 mM)	KCl (1 M)
	0.1 ml (1 mM)	EDTA (500 mM), pH 8.0
	Add ddH_2O to 50 ml	
	Add just prior to use to 1 ml of CBM:	
	10 µl (1 mM)	PMSF (100 mM)
	5 µl (1 mM)	Sodium-o-vanadate (200 mM)
	20 µl (1x)	50x Protease inhibitor cocktail
Lysis buffer 1A (L1A):	0.9 ml	CBM (including inhibitors)
	0.1 ml (0.4%)	ddH_2O
Lysis buffer 1B (L1B):	0.9 ml	CBM (including inhibitors)
	0.1 ml (1%)	NP-40 (10%)
Lysis buffer 2 (L2):	2.5 ml (50 mM)	Tris/HCl, pH 7.4 (1 M)
	2 ml (0.4%)	SDS (10%)
	0.5 ml (5 mM)	EDTA pH 8.0 (0.5 M)
	5 ml (500 mM)	NaCl (5 M)

	Add ddH_2O to 50 ml	
	Add just prior to use to 1 ml of L2:	
	10 µl (1 mM)	PMSF (100 mM)
	5 µl (1 mM)	Sodium-o-vanadate (200 mM)
	20 µl (1x)	50x Proteaseinhibitor cocktail
Dilution buffer (DB):	2.5 ml (50 mM)	Tris/HCl, pH 7.4 (1 M)
	Add ddH_2O to 50 ml	
	Add just prior to use to 1 ml of DB:	
	10 µl (1 mM)	PMSF (100 mM)
	5 µl (1 mM)	Sodium-o-vanadate (200 mM)
	20 µl (1x)	50x Proteaseinhibitor cocktail
Triton X-100 (10%):	1 ml (10%)	Triton X-100
	Add ddH_2O to 10 ml	
Wash buffer 1 (WB1):	5 ml (100 mM)	Tris/HCl, pH 8.0 (1 M)
	12.01 g (4 M)	Urea (60.06 g/mol)
	Add ddH_2O to 50 ml	
Wash buffer 2 (WB2):	0.395 g (100 mM)	NH4HCO3 (79.06 g/mol)
	Add ddH_2O to 50 ml	

BioID is a powerful method to study physiologically relevant protein interactions in living cells or tissues. This proximity labelling method was initially established by Roux *et al.*[178] and it is based on the fusing of a bait protein to a promiscuous biotin ligase from *E.coli* (BirA*, harboring a R118G point mutation). After supplementation of the culture medium with biotin, the biotin ligase will biotinylate the vicinal proteins, which can then be selectively isolated and identified by mass spectrometry.
For this purpose, constructs containing a protein of interest and a flexible linker region fused to a biotin ligase were ordered from IDT (San Jose, USA) (see section 3.13) and then cloned using the Gibson Assembly approach (see section 4.3.5). After transformation and plasmid DNA isolation (see sections 4.3.6 and 4.3.7), linearized plasmids were used for *in vitro* transcription (see section 4.5.3) and the resulting mRNA transcripts utilized in transfection experiments.

Briefly, 50 µg of 3xFlag-NLS-BirA* (construct used as reference to calculate the required amount of mRNA for the other construct) and 135 µg of C-term-3xFlag-EGR2-BirA*, were introduced by electroporation (see section 4.1.5) in iDC ($50x10^6$ cells per construct). Three hours after transfection, the cell culture medium was supplemented with 50 µM biotin and five hours later, cells were harvested for subsequent lysis and purification of biotinylated proteins. Of note, the optimal time points for addition of biotin and harvesting the cells were initially determined by Western blot.

Before lysis, cell pellets were washed twice with ice-cold PBS and then resuspended in 1250 µl of L1A, followed by addition of an equal amount of L1B. Suspensions were incubated on ice for 10 min before nuclei were spun down at 700×g for 5 min at 4°C. In a next step, nuclei were resuspended in 1 ml L2 buffer and sonified to fragment the genomic DNA. All sonification steps were carried out with a constant duty cycle, output control 2 for 10 sec using a Branson Sonifier 250 (Danbury, USA). After sonification in L2, a second sonification in Triton X-100 (10%) at a final concentration of 2% was also performed. Then, an equal amount of DB was added and suspensions were again sonified. Lysates were centrifuged at 11.000×g for 15 min at 4°C and supernatants were transferred into a new tube. To completely remove residual unbound biotin, a dialysis step using Slide-A-Lyzer MINI Dialysis Devices (Thermo Fisher Scientific, Waltham, USA) was performed overnight at 4°C according to the manufacturer's instructions. In the next day, Pierce™ Streptavidin Magnetic Beads (Thermo Fisher Scientific, Waltham, USA) were prewashed according to manufacturer's instructions and 150 µl prewashed beads were added to each lysate. Then, lysates were incubated on a rotating wheel at RT for 1-2 h. Following incubation, beads were washed on a magnet for three times with WB1 and WB2 (400 µl each) at RT and transferred to a new tube with the last washing step. Protein-conjugated beads were finally resuspended in 50 µl of WB2 and sent to the Zentrallabor fur Proteinanalytik (ZfP) of the LMU (Prof. Dr. Axel Imhof) for subsequent on-bead digestion and mass spectrometry analysis.

4.6.5 Co-immunoprecipitation (CoIP)

CoIP associated with western blot is a powerful method to uncover physiological relevant protein-protein interactions[179].

CoIP and reciprocal CoIP of TET2-containing complexes were performed using the Dynabeads antibody coupling kit provided by Thermo Fisher Scientific. Accordingly, 5 µg of anti-TET2 (MABE462, Merck), anti-FLAG M2 (F3165, Sigma) or anti-IgG (sc-2025, Santa Cruz) antibody were covalently coupled to 1 mg of Dynabeads® M-270 Epoxy beads (per IP) according with the manufacturer instructions.

Per IP experiment were used $15x10^6$ transfected THP-1/iDC cells. Before transfection, 3xFLAG EGR2 and 3xFLAG NAB2 constructs were ordered from IDT (San Jose, USA) (see section 3.13), cloned using the Gibson Assembly approach (see section 4.3.5), transformed (see section 4.3.6), linearized (see section 4.3.1) and used for *in vitro* transcription (see section 4.5.3).

Briefly, 26 µg of 3xFLAG-EGR2 and 29 µg of 3xFLAG-NAB2 mRNA transcripts were introduced by electroporation (see section 4.1.5) into dendritic cells (day 3) or THP-1 cell line. After 3 h incubation at 37 °C, transfected (THP-1 and iDCs-d3) cells were harvested and washed with 15 ml ice-cold PBS (including PMSF). Following centrifugation at 1200 rpm for 5 min at 4°C, cells were lysed in 1 ml of Dyna-IP buffer (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1% Triton X-100, 2 mM EDTA, 1 x protease inhibitor cocktail) and incubated on ice for 30 min with gentle shaking.

After incubation, cell lysates were sonified three times, 10 sec each followed by 30 sec rest on ice, using a Branson Sonifier 250 (constant duty cycle, output control 2) and insoluble material was removed by centrifugation at 13000 rpm for 7 min at 4°C. Afterwards, the lysate protein concentration was determined using the Qubit Protein Assay Kit and 10 µg of each lysate were kept as input control. Then, lysates were incubated with the aforementioned anti-TET2, anti-FLAG M2 or anti-IgG-coupled Dynabeads for 2 h and 30 min at RT on a rotating wheel. After lysate incubation, beads were washed on a magnet three times each with PBS including 0.02% Tween 20 and ultrapure water respectively. In a next step, protein complexes were eluted from the beads with 50 µl 2 x SDS sample buffer without 2-mercaptoethanol followed by incubation at 95°C for 10 min. Then, the supernatants were collected on a magnetic rack and 2.5 µl of 2-mercaptoethanol added to each sample. Approximately 25 µl of IP samples along with 10 µg of input samples were loaded into a 8% polyacrylamide gel (see section 4.6.2) and the western blot was performed as described in section 4.6.3, using the following antibodies: anti-TET2 (A304-247A, Bethyl, 1:1000) and anti-FLAG M2 (F3165, Sigma, 1:2000).

4.7 High Throughput Sequencing-based Methods

4.7.1 Chromatin Immunoprecipitation coupled with NGS (ChIP-seq)

Required solutions and buffers:

DSG (2 mM):	50 mg DSG	Thermo Fisher Scientific
	300 µl DMSO	Sigma Aldrich
	Dissolve DSG in DMSO and add 76,3 ml PBS	
Formaldehyde, methanol-free:	16% (w/v)	Thermo Fisher Scientific
Glycine (20x):	9.85 g (2,625 M)	Glycine
	Add ddH_2O to 50 ml	
Cell Buffer Mix (CBM):	1 ml (10 mM)	HEPES/KOH (1 M), pH 7.9
	4.25 ml (85 mM)	KCl (3 M)

	200 µl (1 mM)	EDTA (500 mM, pH 8.0)
	Add ddH_2O to 97 ml	
	Add just prior to use to 1 ml of CBM:	
	10 µl (1 mM)	PMSF (100 mM)
	20 µl (1x)	50x Protease inhibitor cocktail
Nuclear Lysis Buffer (NL):	5 ml (50 mM)	Tris/HCl (1M), pH 7.4
	5 ml (1%)	SDS (20%)
	1.43 ml (0.5%)	Empigen BB (35%)
	2 ml (10 mM)	EDTA (500 mM, pH 8.0)
	Add ddH_2O to 97 ml	
	Add just prior to use to 1 ml of NL:	
	10 µl (1 mM)	PMSF (100 mM)
	20 µl (1x)	50x Protease inhibitor cocktail
Dilution Buffer (DB):	2 ml (20 mM)	Tris/HCl (1 M), pH 7.4
	2 ml (100 mM)	NaCl (5 M)
	400 µl (2 mM)	EDTA (500 mM, pH 8.0)
	5 ml (0.5%)	TritonX-100 (10%)
	Add ddH_2O to 97 ml	
	Add just prior to use to 1 ml of DB:	
	10 µl (1 mM)	PMSF (100 mM)
	20 µl (1x)	50x Protease inhibitor
Wash Buffer I (WBI):	2 ml (20 mM)	Tris/HCl (1 M), pH 7.4
	3 ml (150 mM)	NaCl (5 M)
	500 µl (0,1%)	SDS (20%)
	10 ml (1%)	Triton X-100 (10%)
	400 µl (2 mM)	EDTA (500 mM, pH 8.0)
	Add ddH_2O to 100 ml	

Wash Buffer II (WBII):	2 ml (20 mM)	Tris/HCl (1 M), pH 7.4
	10 ml (500 mM)	NaCl (5 M)
	10 ml (1%)	TritonX-100 (10%)
	400 µl (2 mM)	EDTA (500 mM, pH 8.0)
	Add ddH_2O to 100 ml	
Wash Buffer III (WBIII):	1 ml (10 mM)	Tris/HCl (1 M), pH 7.4
	10 ml (250 mM)	LiCl (5 M)
	10 ml (1%)	NP-40 (10%)
	10 ml (1%)	Deoxycholat (10%)
	200 µl (1 mM)	EDTA (500 mM, pH 8.0)
	Add ddH_2O to 100 ml	
TE Buffer:	1 ml (10 mM)	Tris (1 M), pH 8.0
	0.2 ml (1 mM)	EDTA (500 mM, pH 8.0)
	Add ddH_2O to 100 ml	
Elution Buffer (EB):	500 µl (0.1 M)	NaHCO3 (1 M)
	250 µl (1%)	SDS (20%)
	Add ddH_2O to 5 ml	
Sepharose Cl-4B:	25 µl or 50 µl/IP	
nProtein A/G Sepharose:	25 µl or 50 µl/IP	
	wash 3x with TE, pH 8.0, before use	

ChIP-seq is a technique widely used to study protein interactions with DNA at the genome-wide level. This method is based on the principle that formaldehyde directly targets amino groups on amino acids and nucleotides, resulting in a covalent crosslink between adjacent proteins and DNA. For TFs that bind DNA as dimers, an additional fixation step with DSG (disuccinimidyl glutarate; Thermo Fisher Scientific, Waltham, USA) to crosslink protein-protein interactions is also performed.

Due to the lack of suitable ChIP-grade antibodies for most of the TFs of interest, 3xFLAG tag fusion proteins were generated as described before (see sections 4.2 and 4.3). Then, the respective ivt mRNAs (see section 4.5.3) were introduced by electroporation into iDCs (day 4 of differentiation) (see section 4.1.5).

Dual Crosslinking with DSG (disuccinimidyl glutarate) and Formaldehyde

Cells were harvested 6 h after transfection and centrifuged for 10 min at 300xg, 4°C. After centrifugation cells were resuspended in 1x PBS with 2 mM of DSG and crosslinked for 30 min at RT. Then cells were fixed with 1% formaldehyde for 10 min at RT and after incubation, the reaction was stopped with 0.125 M glycine. Followed by centrifugation at 300xg, 4°C for 10 min, cells were washed twice with 1x PBS including 1 mM PMSF (centrifugation at 300xg, for 8 min, at 4°C) and pellet was stored at -80°C.

Sonification of Fixated Chromatin

The fixated chromatin was resuspended in L1A buffer followed by lysis in buffer L1B. Then, samples were incubated on ice for 10 min and nuclei was pelleted at 700xg for 5 min. This pellet was resuspended in nuclear lysis buffer L2 and chromatin was sheared using a Branson Sonifier 250 (Danbury, USA) using the following settings:

Branson Sonifier 250 settings	
Duty cycle	Constant
Output control	2
Time	10 sec (repeat 5x)
Ice incubation	30 sec (repeat between each sonification)

To clear the lysate, the solution was centrifuged at 13000 rpm, 4°C for 5 min and the supernatant transferred into a new Eppendorf. After centrifugation, 5% of the lysate was saved as input control and 20 µl were used to check fragmentation (ideally fragments size around 200-500 bp) on the agarose gel.

Immunoprecipitation

In order to remove all unspecific binding fragments, the lysate was precleared with 50 µl Sepharose CL-4B beads (GE Healthcare, Munich, Germany) (blocked with 0.5% BSA and 20 µg glycogen) including dilution buffer for 2 h. Then, samples were immunoprecipitated overnight with 2.5 µg of antibody against FLAG M2 (Sigma Aldrich, F3165). In parallel, nProtein G Sepharose beads (GE Healthcare, Munich, Germany) were blocked with 0.5% BSA and 20 µg glycogen overnight at 4°C. In order to pull down antibody-chromatin complexes, blocked Protein G coated beads were added to the lysate antibody mix and rotated for 2-3 h at 4°C. After the antibody-chromatin-complex bound to the beads, they were thoroughly washed twice with WB I, WB II and WB II, and three times with TE buffer (400 µl each). At the end, DNA was eluted in 200 µl EB and 10 µl 5 M NaCl were added to reverse crosslinking, which was then conducted at 65°C overnight.

Chromatin Purification

DNA was treated with 7 µl RNase (10 µg/µl) and 5 µl Proteinase K (20 µg/µl) for 1 h at 37°C to degrade the remaining RNA and proteins from the samples. DNA was purified using the Monarch PCR & DNA Cleanup Kit (NEB, Frankfurt, Germany) according to the manufacturer's instructions and eluted in 50µl EB to be used for library preparation.

Generation of DNA Libraries for Next Generation Sequencing (NGS)

To enable sequencing on an illumina platform, adaptors with specific barcodes were ligated to enriched DNA fragments. Library preparation was carried out using the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, Frankfurt, Germany) according to manufacturer's instructions. The quality of dsDNA libraries was analysed using the High Sensitivity D1000 ScreenTape Kit (Agilent) and concentrations were determined with the Qubit dsDNA HS Kit (Thermo Fisher Scientific). Sequencing was performed at the Biomedical Sequencing Facility (BSF) of the Research Center for Molecular Medicine of the Austrian Academy of Sciences (Ce-M-M, Vienna, Austria) using an Illumina HiSeq 3000 sequencer.

ChIP-seq Analysis

Reads (single-end) were aligned to the human genome (GRCh38/hg38) using bowtie2[180] in very sensitive mode, keeping only reads that map to a single unique genomic location for further analysis (MAPQ > 10). Initial quality control was performed by calculating the fraction of reads in peaks (FRIP) by running HOMER's[181] (v4.9) findPeaks program in ''factor'' or "histone" mode using default parameters and the appropriate matching background data set (either ChIP input, genomic DNA or control ChIP). For further analyses, chromosome scaffolds were removed. TF ChIP-seq peaks were called using HOMER's findPeaks program in ''factor'' mode with -fdr 0.00001 to identify focal peaks. Peak sets were filtered by subtracting blacklisted genomic regions[182], and by filtering out regions with a mappability <0.8. The latter was annotated to peak regions from mappability tracks generated with the GEM package[183] using HOMER's annotatePeaks.pl.

4.7.2 Assay for Transposase-Accessible Chromatin Sequencing (ATAC-seq)

Required buffers and solutions:

Digitonin:	1% (w/v)	Promega
DNase:	35.5 U/µl	Sigma-Aldrich
NP-40:	10% (w/v)	Sigma-Aldrich
Tween-20:	10% (w/v)	Sigma-Aldrich
Resuspension Buffer (RSB):	0.5 ml (10 mM) 0.1 ml (10 mM) 0.15 ml (3 mM) Add ddH_2O to 50 ml	Tris/HCl (1 M), pH 7.4 NaCl (5 M) MgCl2 (1 M)
ATAC lysis buffer:	970 µl 10 µl (0.1%) 10 µl (0.1%) 10 µl (0.01%)	RSB NP-40 (10%) Tween-20 (10%) Digitonin (1%)
RSB-Tween-20:	990 µl 10 µl (0.1%)	RSB Tween-20 (10%)
Transposition Mixture:	25 µl (1x) 2.5 µl (100 nM) 16.5 µl 0.5 µl (0.01%) 0.5 µl (0.1%) Add sterile ddH_2O to 50 µl	2x TD buffer (Nextera, Illumina) transposase (Nextera, Illumina) PBS Digitonin (1%) Tween-20 (10%)

The ATAC-seq technique is widely used to study chromatin accessibility and to infer about nucleosome positioning (method overview in Figure 4.1).

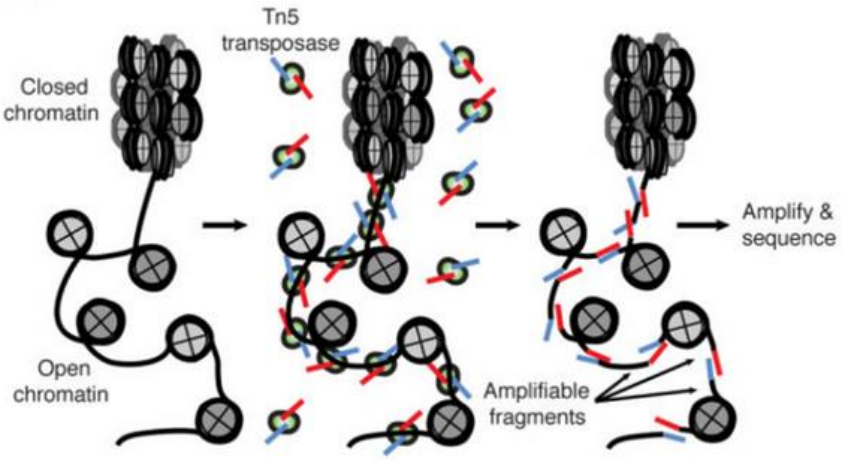

Figure 4.1 Overview of ATAC-seq library preparation

In this method, a hyperactive Tn5 transposase is used to simultaneously fragment and insert sequencing adapters into nucleosome-free regions (open chromatin). Then, DNA fragments are amplified and subjected to high-throughput sequencing, yielding reads that indicate regions of increased chromatin accessibility (adapted from Buenrostro *et al*.[184]).

ATAC-seq was carried out as described by Corces *et al.*[185]. Briefly, cells were harvested and washed with 1xPBS. Due to the low cell viability, which was below than 50%, dead cells were excluded using the magnetic bead depletion based on Annexin V MicroBead Kit (Miltenyi Biotec, Bergisch Gladbach) according with the manufacturer's instructions. To further increase the percentage of viable cells, a treatment with DNase at a final concentration of 200 U/ml in RPMI medium coupled with 30 min incubation at 37°C was performed. After incubation, cells were washed twice with ice-cold 1x PBS in order to remove DNase before proceeding to the transposition reaction. After counting the cells, 50,000 viable cells were centrifuged at 2700 rpm at 4ºC for 5 min in a fixed angle centrifuge. The supernatant was discarded and cells were resuspended in 50 µl of cold ATAC lysis buffer by pipetting up and down three times. Next, the lysate was incubated on ice for 3 min, washed out with 1 ml of cold RBS-Tween-20 buffer and mixed by inverting the tube three times. The nuclei pellet was recovered by centrifugation at 2700 rpm for 10 min at 4ºC in a fixed angle centrifuge. The supernatant was carefully removed and the cell pellet was resuspended in 50 µl of transposition mixture by pipetting up and down 6 times. The reaction was incubated in a heat block at 37°C for 30 min and with constant mixing (1000 rpm). Then, reaction was cleaned up using the NEB Monarch PCR & DNA Cleanup Kit according with the manufacturer's instructions. The samples were eluted in 21 µl of elution buffer and 10 µl of purified DNA were used to PCR amplification. To generate a library compatible with next generation sequencing on an Illumina platform, Nextera XT i7- and i5-index primers (Illumina, San Diego, USA) were used in the PCR reaction mixture as shown in Table 4.11. After mixing all the components, the PCR program was conducted as indicated in Table 4.12.

Table 4.11 PCR reaction setup for DNA library preparation

Component	Volume	Final Concentration
5x HF buffer	10 µl	1x
H_2O	10 µl	-
dNTPs (10 mM)	1.5 µl	0.3 mM
Betaine (5 M, Sigma-Aldrich)	13 µl	1.3 M
Nextera XT i7 index (25 µM, Illumina)	2.5 µl	1.25 µM
Nextera XT i5 index (25 µM, Illumina)	2.5 µl	1.25 µM
Phusion HF Polymerase (2 U/µl, Thermo Fisher)	0.5 µl	1 U
DNA	10 µl	-

Table 4.12 PCR program for DNA library generation

Temperature	Time	Cycles
72°C	5 min	1
98°C	30 sec	1
98°C	10 sec	12
63°C	30 sec	
63°C	1 min	
72°C	1 min	1
4°C	forever	-

DNA was purified and size selected using magnetic beads (Agencourt AMPure XP; Beckman Coulter, Krefeld, Germany). In terms of DNA purification the ratio sample:beads was 1:1.8 and for size selection (for fragments of 250 to 600 bp) the ratio was 1:0.55. Briefly, the addition of the beads to the samples was always followed by incubation for 10 min at RT. Then samples were placed in a magnetic rack for 3 min and washed twice with 70% ethanol. After, beads were air dried and eluted in water (upon fragments size selection) or in 15 µl of EB, corresponding to the ultimate DNA purification step.
ATAC libraries were analysed using the High Sensitivity D1000 ScreenTape Kit (Agilent) and sequenced using an Illumina NextSeq550 sequencer according to Illumina's instructions.

ATAC-seq Analysis

Reads (paired-end) were aligned to the human genome (GRCh38/hg38) using bowtie2 in very-sensitive and no-discordant modes, keeping only reads that map to a single unique genomic location for further analysis (MAPQ > 10). Initial quality control was performed by calculating the fraction of reads in peaks by running HOMER's findPeaks program in using parameters "-region -size 150". ATAC-seq peak regions were called by combining two different approaches: The basic peak region

set was called using HOMER's findPeaks program in 'region'' mode using parameters "-size 150 -minDist 250 -L 2 -fdr 0.00001" to identify regions of variable length by stitching nucleosome-size peaks. To exclude shallow peak regions, only those were kept that overlapped a second peak set that was generated in ''factor'' mode using parameters "-size 250 -minDist 250 -L 2 -fdr 0.00001" to identify focal peaks. Statistically significant differences in read counts across peaks between sets of replicate ATAC-seq experiments were determined with quantile (0.95) normalization and GC correction using edge R 3.20.8[186] with the cqn package[187] in R (3.4.3). Read coverage across individual peaks sets was calculated using HOMER's annotatePeaks.pl with parameters "-hist 25 -ghist" using merged replicate ATAC-seq data sets and plotted in R using the image function. Average coverage data and 95% confidence intervals were calculated in R and the ggplot2 package was used to draw histograms. Scatterplots were drawn in R using the ggplot2 package and corresponding correlation coefficients were calculated in R.

Motif Analysis

De novo motif discovery in peaks or regions was performed with HOMER's findMotifsGenome.pl program and parameters "-len 7,8,9,10,11,12,13,14 -h". For searches in ChIP-seq peaks we used a 200 bp, peak-centered window, while for DMR or differential ATAC regions the given region sizes were used. *De novo* motifs were further filtered using HOMER's compareMotifs.pl and parameters "-reduceThresh .75 -matchThresh .6 -pvalue 1e-12 -info 1.5". Motif log-odds scores for EGR, ETS and STAT motifs across DMR sets were calculated using HOMER's annotatePeaks program.

Footprints across motif-centered peak sets (centered using HOMER's annotatePeaks.pl program) were generated using the -hist option of HOMER's annotatePeaks.pl with parameters "-hist 1 -len 1" and plotted in R using the ggplot2 package.

4.7.3 RNA-sequencing (RNA-seq)

The RNAseq is used to genome-wide quantification of gene expression with high sensitivity and accuracy. For this purpose, the total cellular RNA was isolated, quantified and quality assessed as described in section 4.5.1. Only high quality RNAs with a RIN (RNA integrity number) equal or higher than 7 were used. Generation of libraries for Illumina sequencing was carried out using the ScriptSeq™ Complete Kit (Human/Mouse/Rat) – Low Input from Epicentre (Chicago, USA) according to the instructions provided by the manufacturer. The quality of dsDNA libraries was checked with the High Sensitivity D1000 ScreenTape Kit (Agilent) and concentrations were determined with the Qubit dsDNA HS Kit (Thermo Fisher Scientific). Sequencing was performed using an Illumina NextSeq550 sequencer according to Illumina's instructions.

RNA-seq Analysis

Sequencing reads were mapped to the human genome using STAR v2.5.3a[188]. The human GRCh38 genome index incorporated gene annotation from GENCODE 44 (release 27) was used to support in spliced alignment. Tables of raw uniquely mapped read counts per human gene were generated during mapping using the built-in --quantMode GeneCounts option in STAR. Differential expression analysis was carried out on raw gene counts using edgeR 3.20.8[186] in R (3.4.3). Pairwise comparisons of indicated data sets were done using the quasi-likelihood test.

Heatmaps of differentially expressed genes used batch-corrected, normalized and scaled CPM (counts per million) data and were generated using the heatmap.2 function of the gplots package in R. Dimensionality reduction based on the tSNE algorithm was done using the Rtsne package and visualized using the ggplot2 package in R. Statistically significant enriched Gene Ontology terms were identified using Metascape[189] and barplots of significance levels were generated in R.

5 Results

Our group previously established a post-mitotic *in vitro* model of MO differentiation as an ideal system to study active DNA demethylation[31]. Later on, this system was used to demonstrate that DNA demethylation during differentiation of MO is preceded by the local appearance of 5hmC[32]. Moreover, using a siRNA-mediated strategy to knock-down (KD) TET2 in primary MO, our group also showed that this enzyme is essential for the DNA demethylation process[32]. Nevertheless, transfected MO cultures showed massive cell death after 48h, making functional studies difficult.

Taking the previous work into consideration, backbone-modified siRNAs (chemical modification patterns are shown in section 3.11) were used in this thesis to transiently knock-down TET2, as well as other TFs that might be involved in the demethylation process. Unlike the siRNAs mentioned before, backbone-modified siRNAs are well tolerated by MO with no significant effect on cell survival. Similarly, during differentiation, iDCs develop a normal morphology and are phenotypically indistinguishable from non-transfected cells. Hence, the work presented in this thesis contains the first reported data generated by our group using this improved siRNA-based system to knock-down target genes.

For easier understanding of the present work, preliminary results generated before in our group will be shown. When this is the case, it will be clearly stated.

5.1 Genome-wide DNA methylation changes during MO differentiation

As already stated, the previously established *in vitro* model of MO differentiation was used to differentiate iDCs in culture for seven days. To obtain a global view of the epigenomic changes occurring during the differentiation process, cells were harvested and used in downstream applications as depicted in Figure 5.1.

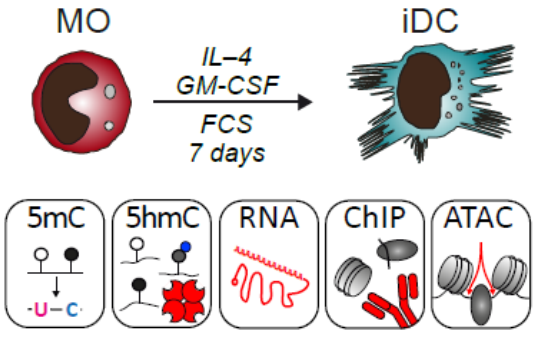

Figure 5.1 Schematic of the experimental setup for *in vitro* MO differentiation into iDCs and downstream methodologies

Primary human MO were differentiated into iDCs (for 7d) upon stimulation with IL-4 and GM-CSF cytokines, including supplementation of culture medium with FCS. Then, iDCs were harvested and used in the following applications: 5mC (methylation analysis), 5hmC (hydroxymethylation analysis), RNA-seq, ChIP-seq and ATAC-seq.

5.1.1 Identification of iDC-specific differentially methylated regions (DMRs)

To capture the entire dynamic changes in DNA methylation, published whole genome bisulfite sequencing (WGBS) data sets of MO (n=4) and iDC (n=6) were used. The analysis of these data sets allowed the identification of a group of regions that reproducibly lost DNA methylation during iDC differentiation (DC-specific DMR) as illustrated in Figure 5.2 (preliminary data generated in our group). Of note, the WGBS data was validated in our group by random selection of a number of regions showing dynamic changes in methylation and comparing them with the DNA methylation patterns determined by MassARRAY EpiTYPER. Accordingly, there was a tight correlation between EpiTYPER and WGBS data at demethylated regions (data previously generated by our group and not shown in this thesis).

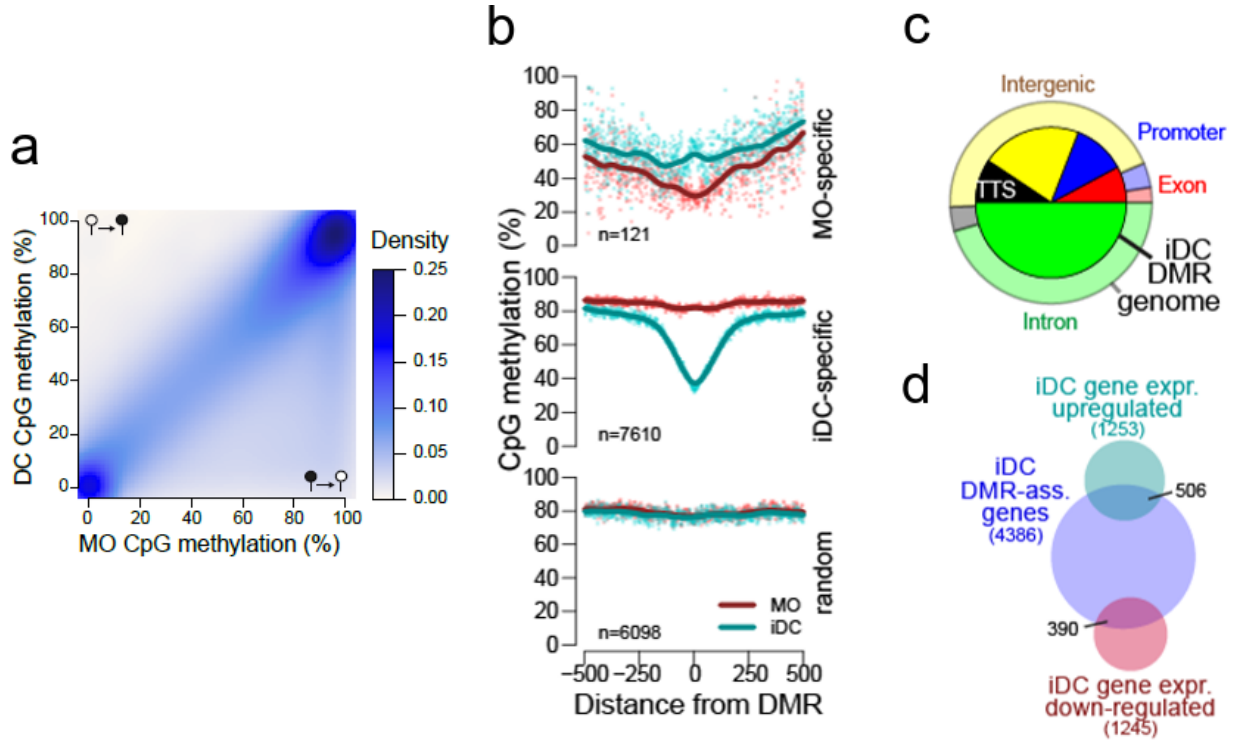

Figure 5.2 Identification of iDC-specific DMRs

a Scatterplot showing the density distribution of average DNA methylation ratios (WGBS) in MO (n=4) and iDCs (n=6). Black lollipops indicate 5mC and white lollipops unmethylated 5C. **b** Genomic distance distribution of averaged DNA methylation ratios (at single CpG resolution) centered on MO-specific, DC-specific DMR and random regions. Lines represent spline curves across individual data points. Sample types are indicated by coloring. **c** Pie chart illustrating the genomic location distribution of iDC-specific DMRs (inner circle) relative to the entire genome. **d** Venn diagram depicting the overlap between iDC DMR-associated genes and genes up- or downregulated during MO differentiation.

As shown in the scatterplot from Figure 5.2a, the majority of differentially methylated sites reside in the lower right half of the diagram, suggesting that demethylation is the predominant event during DC differentiation.

Figure 5.2b shows a total of 7610 regions (middle panel) that reproducibly lost DNA methylation during iDC differentiation (DC-specific DMR) whereas only 121 regions (top panel) gained DNA methylation (MO-specific DMR). On opposite to the latter regions, a total of 6098 random regions (bottom panel) did not exhibit any changes in DNA methylation ratios upon MO differentiation.

Figure 5.2c illustrates the genomic location distribution of iDC-specific DMRs compared to the entire genome. Accordingly, iDC-specific DMRs were enriched in promoters, exons and transcription termination sites (TTS). The Venn diagram in Figure 5.2d shows the overlap between iDC DMR-associated genes and genes that were found to be upregulated (1253 genes) or downregulated (1245 genes) during MO differentiation into iDC. As indicated, only a small number of iDC-DMR associated genes (896 genes) overlapped with genes that were significantly regulated during MO differentiation, indicating that active demethylation processes occur independent of transcriptional changes.

5.1.2 Genome distribution of 5hmC & ATAC signals in MO & iDC

In this thesis, further analyses were focused on iDC-specific DMRs (7610) as described in section 5.1.1. Figure 5.3 shows the dynamic changes in 5hmC enrichment (data generated by Sandra Schmidhofer) and open chromatin (ATAC-seq, see section 4.7.2) across DC-specific DMRs in MO and iDC.

As shown in Figure 5.3a, the majority of hydroxymethylated sites reside in the upper left half of the diagram with a broader distribution of 5hmC enrichment in iDCs at day 7 (bottom panel) compared to iDCs after 18 h (top panel) of culture. These data suggest that DNA demethylation is a continuous process during differentiation of monocytic lineage.

Figure 5.3b & Figure 5.3c show the distribution of 5hmC and ATAC signals across iDC-specific DMRs. As expected, the coverage of 5hmC signals is not uniformly distributed across the different stages of MO differentiation with higher 5hmC coverage in iDC-18h compared to MO. Interestingly, 5hmC coverage is broadening in iDC-d7, suggesting that demethylation spreads from the center. Parallel to the 5hmC turnover we also observe an increase of chromatin accessibility (as measured by ATAC-seq).

The Venn diagram in Figure 5.3d indicates that only 3751 DMRs overlapped with MO- and/or iDC-specific ATAC-seq peaks, while the remaining DMRs (3859) did not overlap either with MO or iDC ATAC-seq peaks. These data suggest that only half of the DMR get detectably accessible during MO differentiation.

Figure 5.3e displays the IGV genome browser tracks of two example regions for active demethylation characterized by a continuous increase on 5hmC and gains in chromatin accessibility (ATAC) in iDCs.

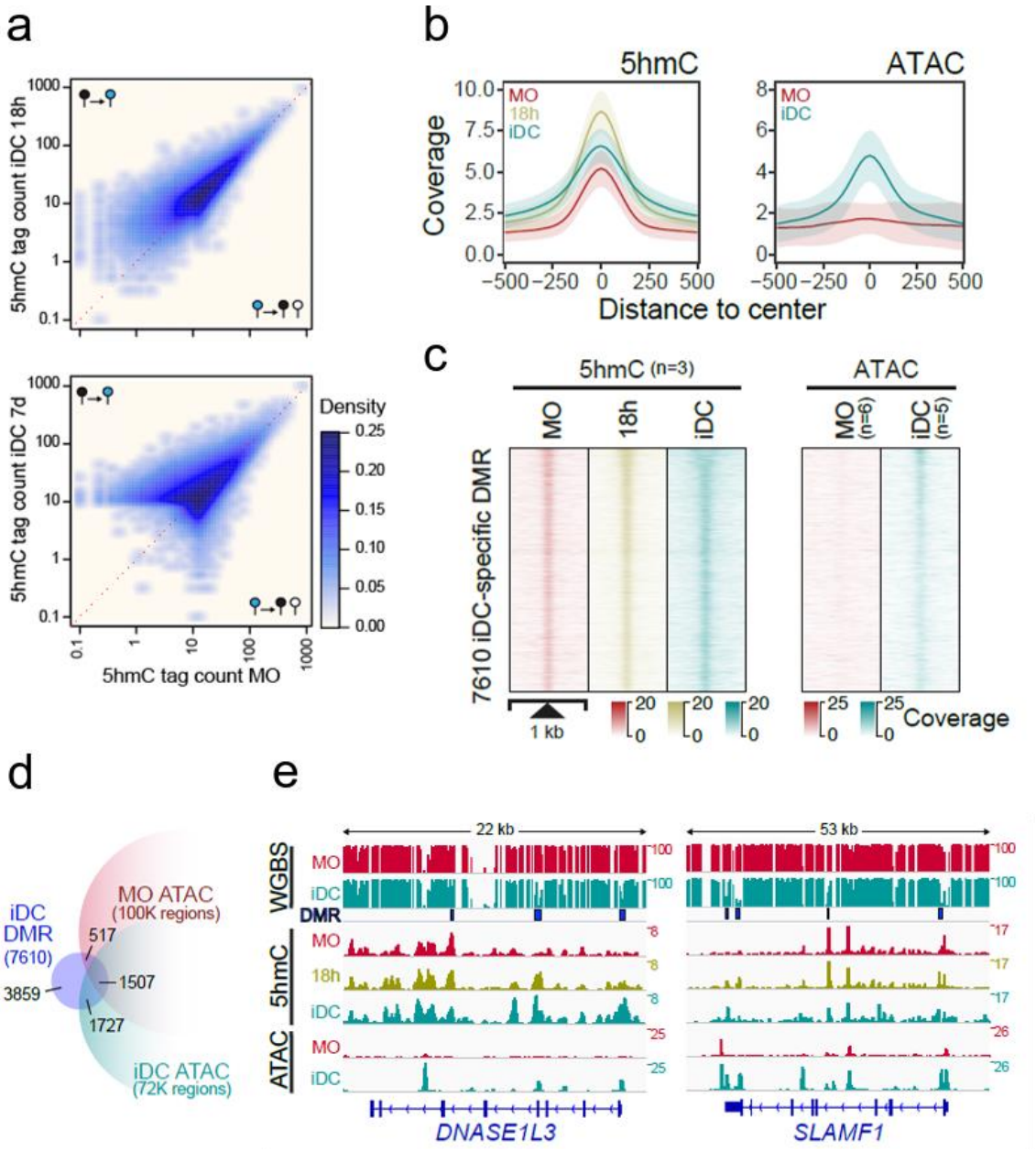

Figure 5.3 DNA demethylation during MO differentiation correlates with dynamic changes in 5hmC and increasing chromatin accessibility

a Scatterplots showing the density distribution of 5hmC enrichment of MO compared to iDCs at 18h (top panel) or iDCs at d7 (bottom panel) across a merged set of 5hmC peaks. Black lollipops indicate 5mC, blue lollipops 5hmC, and white lollipops unmethylated 5C. **b** Genomic distance distribution of 5hmC enrichment and ATAC coverage centered on iDC-specific DMR. **c** Distribution of the 5hmC and ATAC-seq signals across iDC-specific DMRs in MO and iDCs. **d** Venn diagram showing the number of iDC-DMRs overlapping with MO- and/or iDC-specific ATAC-seq peaks. **e** IGV genome browser tracks of example regions (DNASE1L3 and SLAMF1 loci) including CpG methylation ratios (WGBS), 5hmC and ATAC-seq coverage in MO and iDCs as well as the genomic locations of DMRs.

5.2 Role of TET2 in iDC differentiation & associated epigenetic processes

As previously mentioned, an improved siRNA-mediated approach to knock-down target genes was established in our group. Considering that, a significant fraction of siRNAs designed for a specific gene is not always effective, the gene silencing efficiency of the siRNAs should be first evaluated. For the TET2 knock-down experiments, a siRNA set (ten siRNAs) was previously tested (data not shown) by Julia Wimmer and Claudia Kiesewetter and the siRNA with the best silencing efficiency was used in further experiments.

In this thesis, the TET2 enzyme was transiently knocked-down in MO using a siRNA against TET2. A siRNA against luciferase (gene not present in the human genome) was also included as control (siCTRL). Then, its impact on iDC differentiation, gene expression (RNA-seq), chromatin accessibility (ATAC-seq) and DNA methylation (EpiTYPER) was measured using high-throughput techniques (see sections 4.4.3, 4.7.2 & 4.7.3).

Figure 5.4 illustrates the general experimental setup to knock-down TET2 during MO to iDC differentiation, including the depletion effects on cell morphology/viability and the TET2 knock-down efficiency on protein level.

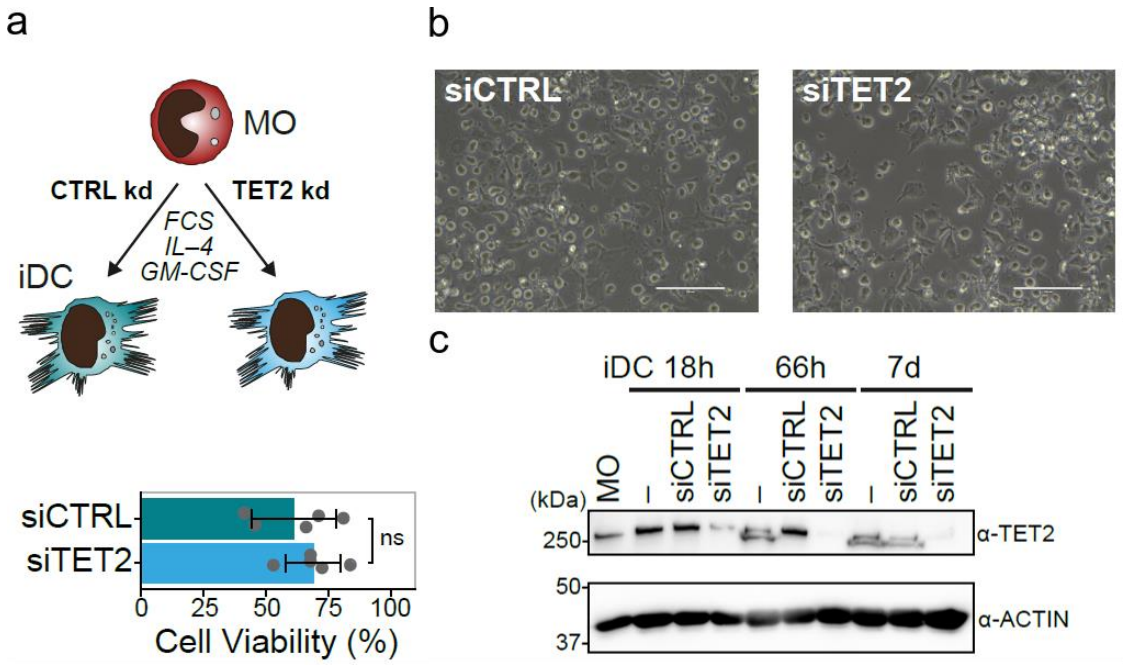

Figure 5.4 Effects of TET2 KD on iDC

a Schematic of the experimental setup and effect of siTET2 treatment on iDC viability (measured at d7). **b** Light microscopy images of iDC (d7) treated with siCTRL or siTET2 (40x magnification, bar represents 100 µm). **c** TET2 protein expression levels in MO, iDCs and siTET2- or siCTRL-treated iDCs at the indicated time points (18h, 66h and 7d). Blots were stained with α-TET2 and α-ACTIN antibodies (protein loading control).

As shown in both Figure 5.4a & Figure 5.4b, TET2 depletion had no impact on either cell survival or cell morphology when compared with control sample (siCTRL-treated iDC).

Figure 5.4c depicts the TET2 knock-down efficiency on protein level as determined by Western blot (see section 4.6.3). As expected, TET2 was detected in MO, iDC and siCTRL-treated iDC at the indicated time points (18 h, 66 h and 7 d). Comparing with the protein loading control (actin) we verified a progressive reduction in siTET2-treated iDC on TET2 expression with almost complete protein depletion after three days of culture.

5.2.1 Effects of TET2-depletion on DNA methylation

The MassARRAY EpiTYPER platform was used to detect and quantify DNA methylation ratios upon TET2 KD. For this purpose, cells were harvested at distinct time points (0 h, 18 h, 66 h and 7 d) during iDC differentiation and DNA was purified and quantified (see section 4.4.1). EpiTYPER experiments were performed as described in section 4.4.3 using appropriated primers (primer sequences can be found in section 3.10.3) to amplify bisulfite-converted DNA.

In general, different actively demethylated regions were analysed with this technique, and ten representative regions were selected to present in this thesis (Figure 5.5 & Figure 5.6).

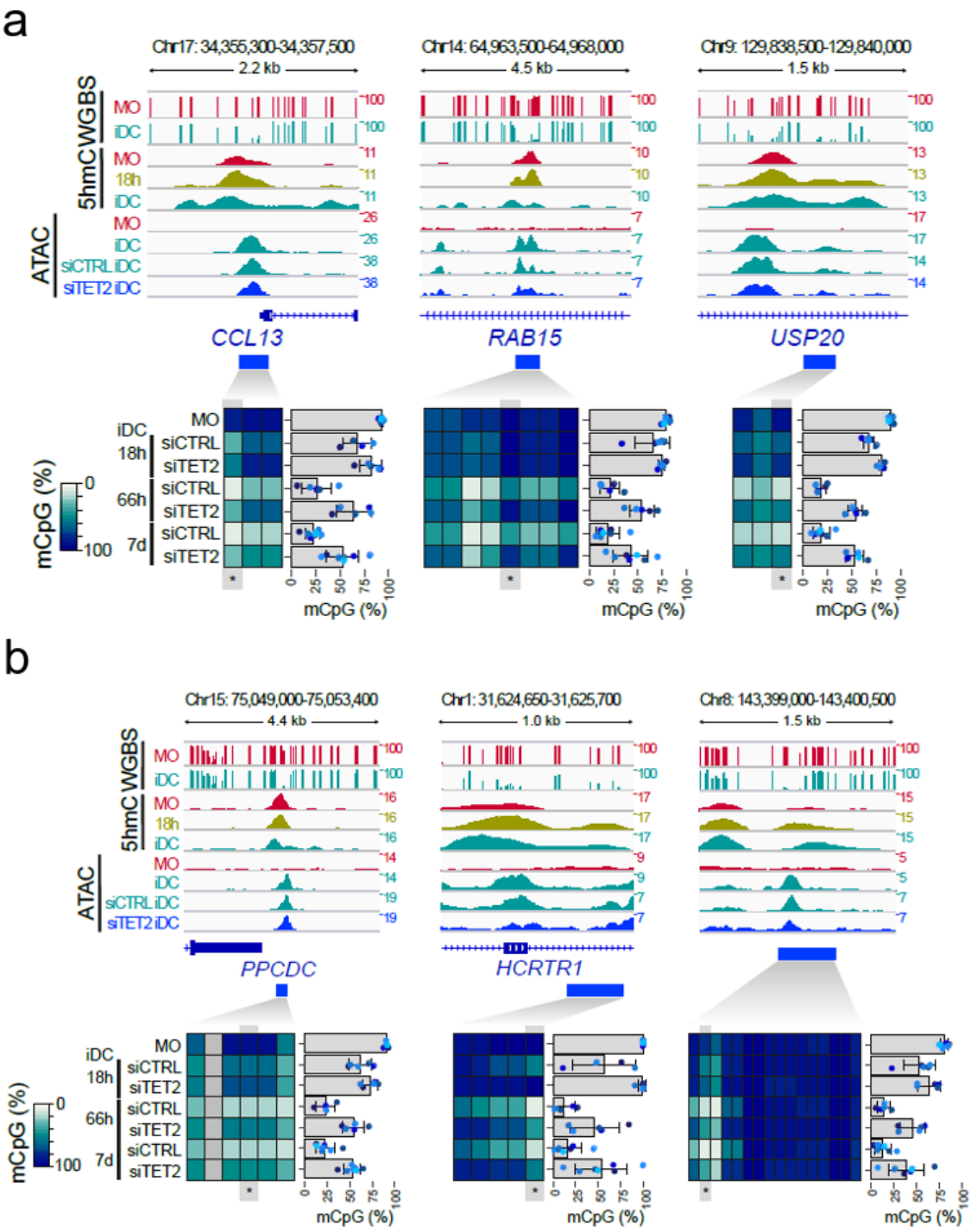

Figure 5.5 Impact of TET2 KD on DNA methylation

a-c IGV genome browser tracks for the indicated WGBS, 5hmC and ATAC data sets at regions around example DMRs. Corresponding DNA methylation ratios for the indicated regions are given below. For heatmaps, methylation ratios (representing means of n ≥ 5 in **a**,**b** and n=3 in **c**) are indicated by coloring (white: no methylation, dark blue: 100% methylation) with each column representing a single CpG. For each region the data of a single CpG (highlighted and marked by asterisks) is shown. Bars represent means ± SD, individual data points are shown as colored dots (each color represents a different donor).

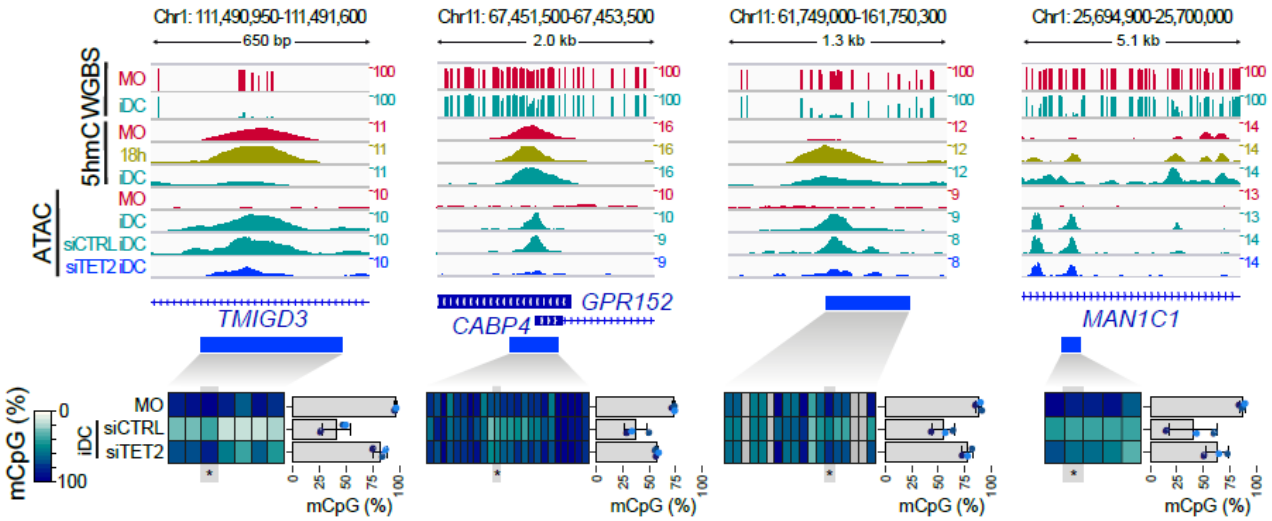

Figure 5.6 Impact of TET2 KD on DNA methylation and chromatin accessibility

IGV genome browser tracks for the indicated WGBS, 5hmC and ATAC data sets at regions around example DMRs. Corresponding DNA methylation ratios for the indicated regions are given below. For heatmaps, methylation ratios (representing means of n=3) are indicated by coloring (white: no methylation, dark blue: 100% methylation) with each column representing a single CpG. For each region the data of a single CpG (highlighted and marked by asterisks) is shown. Bars represent means ± SD, individual data points are shown as colored dots (each color represents a different donor).

Figure 5.5 & Figure 5.6 illustrate the IGV genome browser tracks (top panels) for actively demethylated regions, including different epigenomic data sets (WGBS, 5hmC and ATAC). Each individual region exhibits enrichment in 5hmC and increase in chromatin accessibility in iDCs. The mild effects of TET2 KD on open chromatin are also depicted on the bottom tracks.

Below each IGV genome browser track a heatmap is given, which provides the mean DNA methylation ratios (EpiTYPER) for the indicated region upon different treatments (MO, siCTRL- and siTET2) and time points (18h, 66h and 7d) in Figure 5.5 and 7d only in Figure 5.6.

As expected, in both cases (Figure 5.5 & Figure 5.6) the DNA demethylation was delayed in siTET2-treated iDC (higher methylation ratios compared to siCTRL sample). This delay on demethylation was observed for all the indicated DMRs with single CpGs (highlighted in the barplots) showing changes in the mean methylation ratios equal and greater than 25% upon TET2 depletion. Changes in chromatin accessibility were only observed in few regions (Figure 5.6).

5.2.2 Effects of TET2-depletion on gene expression

While the effects of TET2 KD on DNA (de)methylation are well-documented, its impact on MO-derived iDC gene expression is poorly understood. Consequently, cells were harvested at day 7 (day 0 for MO) of differentiation and RNA was isolated and purified as previously described (see section 4.5.1). Afterwards, RNA-seq was performed in siTET2- and siCTRL-treated iDC as described in section 4.7.3. Mock (electroporation without siRNA), MO and iDC samples were used as additional controls. The gene expression data is summarized in Figure 5.7.

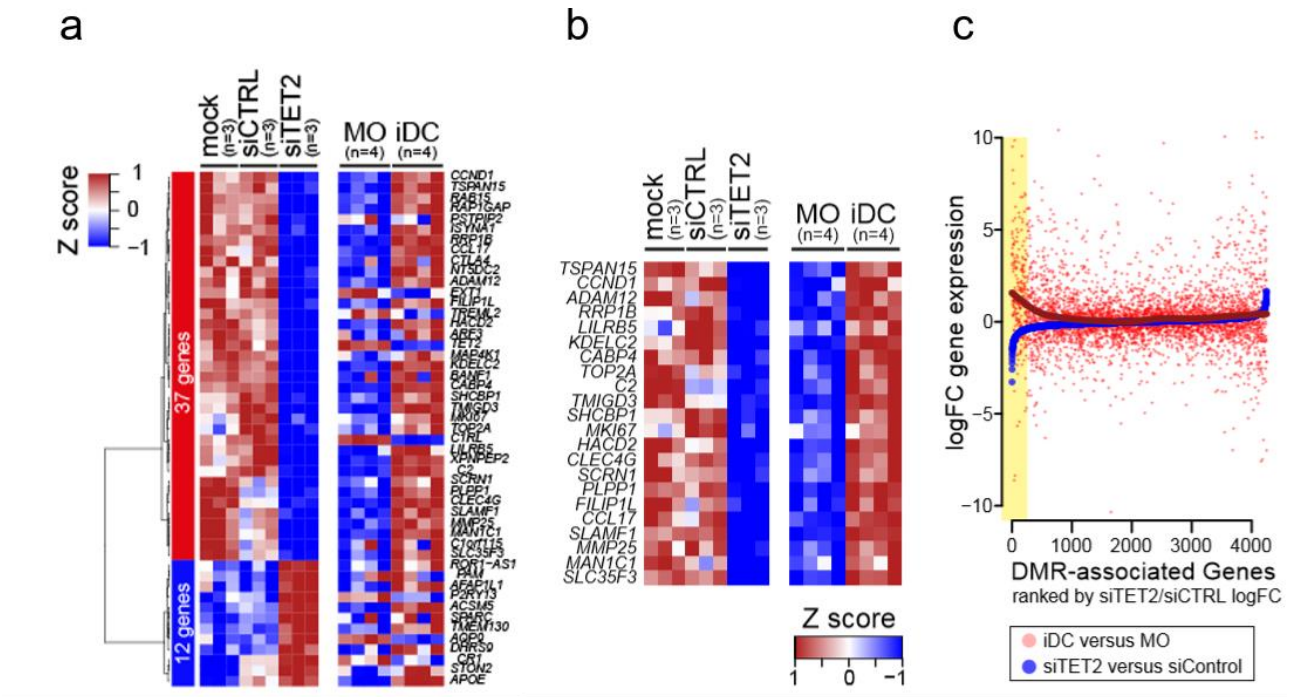

Figure 5.7 Impact of TET2 KD on gene expression

a, b Heatmaps of differential expressed genes between control- (siCTRL) and siTET2-treated cells (left heatmap) and between MO and iDCs (right heatmap). Only genes that were significantly different (FDR ≤ 0.1, RPKM ≥ 1) in both comparisons are shown. Each column represents an individual donor and the expression data is scaled. The heatmap showing all differentially expressed genes of the TET2 knock-down experiments is shown in **a**. **c** Dot plot showing log fold change (logFC) in gene expression of 4386 genes associated with iDC-specific DMR. Genes were ranked by their logFC in gene expression upon TET2 knock-down (blue dotted line). Highlighted in yellow are the genes downregulated upon siTET2 treatment. Red dots represent logFC during MO to iDC differentiation. As indicated by the dark red spline curve, the highlighted genes tend to be upregulated during MO differentiation.

Figure 5.7a depicts all significantly and differentially expressed genes affected by the TET2 KD, which are distributed into two major clusters (downregulated: 37 genes; upregulated: 12 genes; FDR ≤ 0.1, RPKM ≥ 1). From these 37 downregulated genes upon TET2 KD, a large fraction (22 genes) was normally induced during iDC differentiation as shown in Figure 5.7b. The same pattern was observed in Figure 5.7c, where the downregulated genes upon TET2 KD (highlighted in yellow) tended to be upregulated during MO differentiation as indicated by the dark red spline curve. Overall, these data suggest a mild but reproducible effect of TET2 KD on gene expression.

5.2.3 Effects of TET2-depletion on chromatin accessibility

To check whether the open chromatin dynamics would be affected by the depletion of TET2, cells were harvested at day 7 of differentiation and used for ATAC-seq experiments as described in section 4.7.2. The ATAC-seq data is summarized in Figure 5.8.

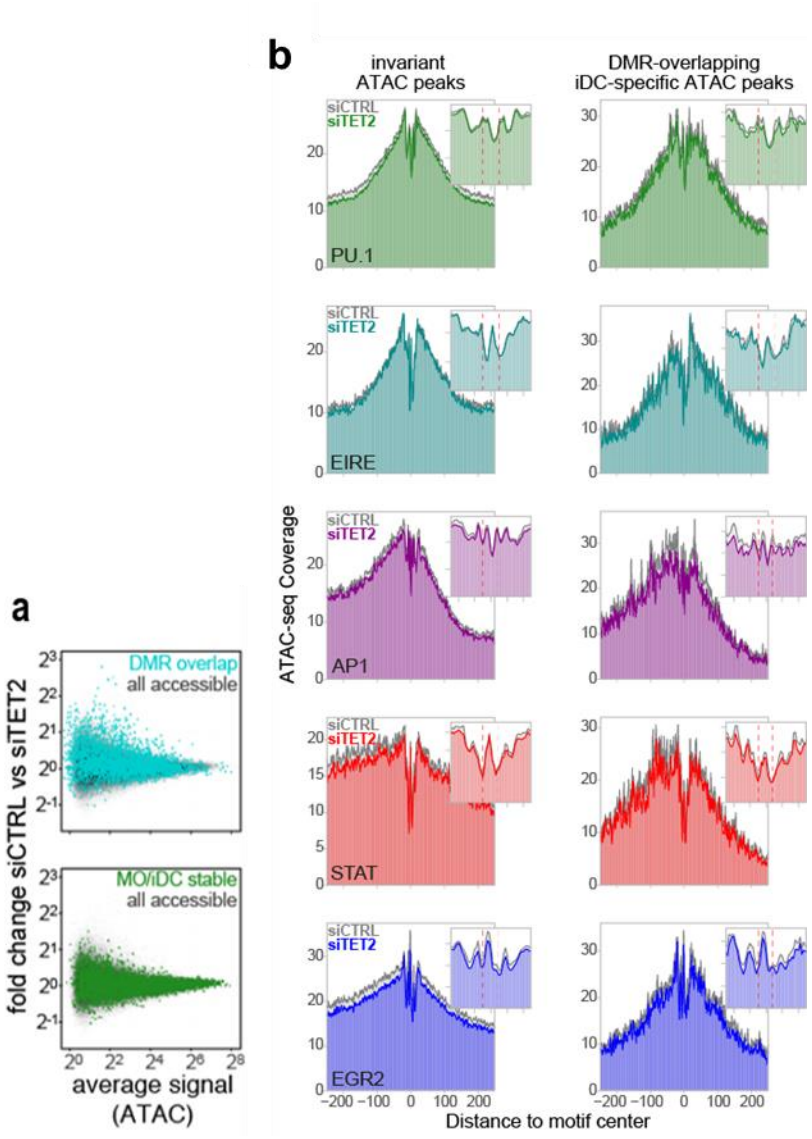

Figure 5.8 Effect of TET2 KD on chromatin accessibility

a MvA plots for ATAC-seq data comparing control- (siCTRL) and siTET2-treatment. Coloring indicates ATAC peaks overlapping DMR (top panel) or ATAC peaks that remain stable during MO differentiation. **b** ATAC-seq footprints across motif-centered, ATAC peaks that were either invariant during MO differentiation (control, left panels), or overlapping DMR and induced accessibility during MO differentiation (right panels). Footprints of control cells (siCTRL) are shown in gray, while footprints of TET2-deficient cells (siTET2) are colored. Smaller histograms in the upper right corner zoom into the central part of the main graph. The position of each motif (as labelled in the bottom left corner of the right histogram) is indicated by two vertical dashed lines.

As illustrated in Figure 5.8, only few regions were significantly different in TET2 KD cells when looking at ATAC peaks that remain stable during MO differentiation (bottom panel). However, at DMR overlapping ATAC-peaks (top panel) some regions loose accessibility upon TET2 KD compared to siCTRL samples (several points dispersed at the upper part of the plot). In fact, the ATAC-seq footprints across motifs enriched through iDC-DMRs (see Figure 5.8b & Figure 5.9b) revealed a small reduction of ATAC signals across PU.1 and AP1 motifs at DMR in TET2-depleted cells. Of note, the binding sites for these TFs are protected from transposase activity (decrease on ATAC signal) by bound TF.

Taken together, these data indicate a mild effect of TET2 KD on chromatin accessibility.

5.3 Transcription factors associated with active demethylation events

As previously described in section 5.2, TET2 KD in MO was clearly associated with a reduction on DNA demethylation as reflected by the higher DNA methylation ratios in siTET2-treated iDC (compared to siCTRL samples). In addition, TET2-depletion impaired chromatin accessibility in few regions and caused slight changes on iDC transcriptional programs.

Despite the mild effects of TET2 KD in our system, the results were reproducible among donors, supporting a critical role of TET2 in the DNA demethylation process during iDC differentiation.

In a next step, we intended to identify potential TFs that are involved in TET2 recruitment to its target loci during IL-4/GM-CSF-driven MO differentiation. For this purpose, the following strategy (described below) was used and the main results are shown in Figure 5.9.

As observed in Figure 5.3b,c, 5hmC signals are clearly detected in MO, suggesting that certain DMRs already initiated the DNA demethylation process at an earlier stage of MO development. Based on those preliminary findings, we checked the distribution of DNA methylation ratios in MO centered on iDC-specific DMRs to identify DMRs that are early or late demethylated. As illustrated in Figure 5.9a, this strategy allowed us to divide our DMRs into quintiles (represented by five pie charts) with decreasing methylation ratios in MO.

Confirming our initial hypothesis, DNA methylation at iDC-specific DMR in MO represent a mixture of highly methylated regions (*de novo* demethylated) and regions that started the demethylation process at earlier stages of MO development (progressive demethylated) (Figure 5.9a,b).

Afterwards, *de novo* motif searches across the DMR quintiles described in Figure 5.9a revealed five motifs corresponding to known candidate TF, including a PU.1 consensus motif, a composite ETS/IRF motif, as well as AP1, STAT and EGR motifs (Figure 5.9b).

Remarkably, partially methylated regions (indicated by the light red pies) were highly enriched for constitutive TFs motifs, such as PU.1 and AP1, which are present in both MO and iDC. Conversely, *de novo* demethylated sites (indicated by the dark red pies) were more enriched for STAT and EGR motifs, factors that are strongly induced upon IL4-GM-CSF-driven MO differentiation.

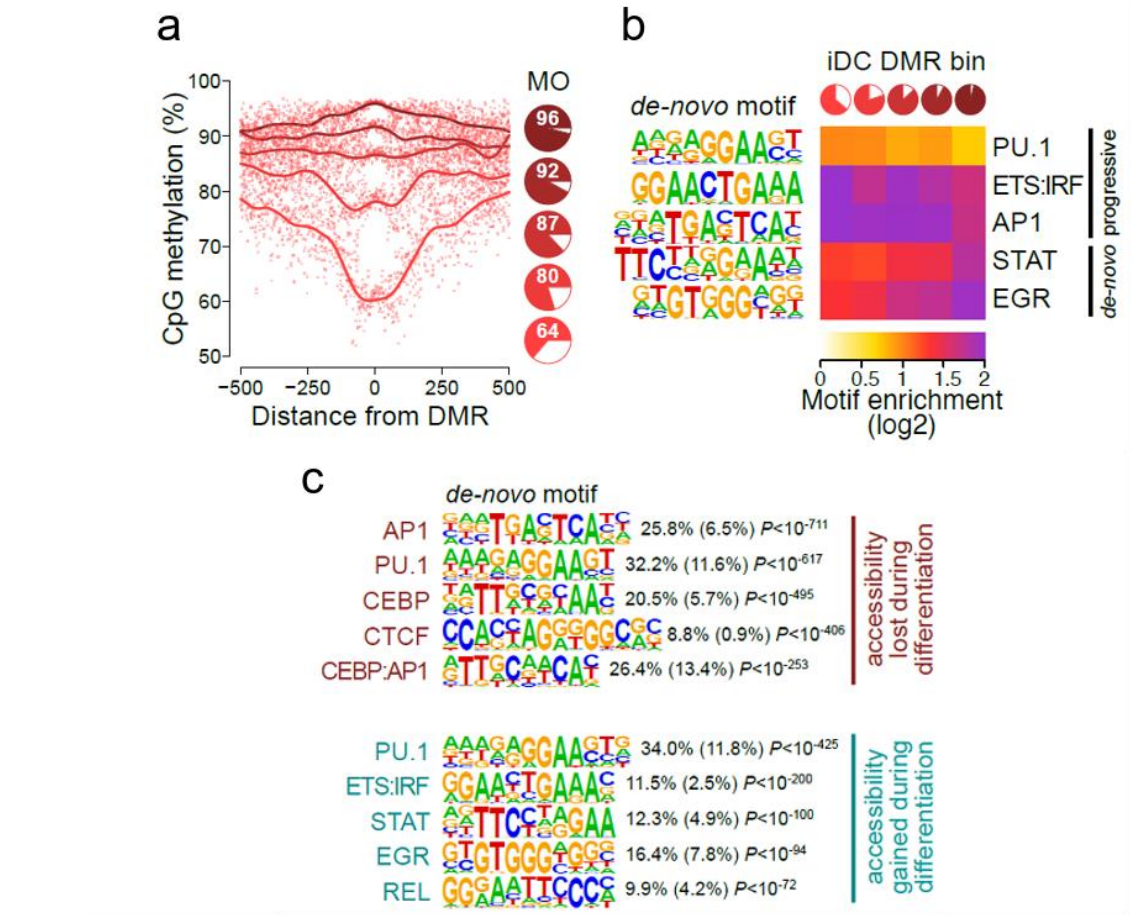

Figure 5.9 TF signatures across progressive and *de novo* demethylated iDC-DMRs

a Genomic distance distribution of mean methylation ratios in MO centered on iDC-specific DMRs. Pie charts represent the mean methylation level (in red) of DMR quantiles sorted by their mean methylation ratio in MO. **b** Heatmap of motif enrichment for the indicated TF contingent on DMR quintiles based on the methylation state of MO (pie charts shown on the top and in Figure 5.9a). **c** *De novo* identified sequence motifs across MO- and iDC-specific ATAC-peaks that loose (upper panel) or gain (lower panel) chromatin accessibility during differentiation. The fraction of motifs in peaks (background values are in parenthesis) and the significance of motif enrichment (hypergeometric test) are given for the top five motifs corresponding to known motif families.

Figure 5.9c depicts *de novo* motif analysis across regions that loose (MO-ATAC peaks) or gain (iDC-ATAC peaks) chromatin accessibility during differentiation. Interestingly, the motif signature across these regions resembled the one described in Figure 5.9b. Hence, regions that lost chromatin accessibility during differentiation were enriched for constitutive TFs (top panel), while regions that gained chromatin accessibility were enriched for TFs induced in iDCs (bottom panel).

In a next step, we checked (in MO and iDC) the gene expression levels of two candidate family TFs (see Figure 5.9), namely IRF and EGR (data shown in Figure 5.10).

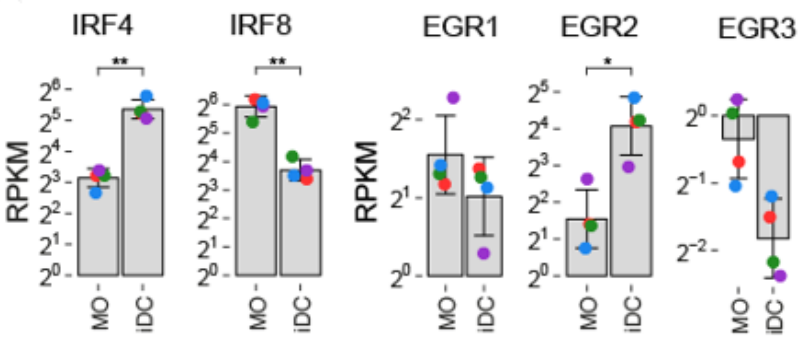

Figure 5.10 Gene expression levels of candidate TF family members in MO and iDC

RPKM were converted from edgeR normalized and batch-corrected CPM. Bars represent means ± SD, individual data points are shown as colored dots (each color represents a different donor). Significance levels correspond to q-values derived from differential gene analyses using cqn and edgeR. (**q<0.01; *q<0.05; glmQLF test and BH correction for multiple testing).

As observed in Figure 5.10, IRF4 and EGR2 are transcriptionally (and significantly) activated in iDCs compared to IRF8 or EGR1/EGR3 TFs. As a result, IRF4 and EGR2 represented valid candidates to drive *de novo* demethylation processes, and thus were investigated in more detail.

5.3.1 Distribution of iDC-specific DMR into distinct groups

To profile the genome-wide binding of IRF4 and EGR2, ChIP-seq was performed not only for these TFs but also for PU.1, a co-factor of IRF4.

To compensate for the absence of ChIP-grade antibodies against EGR2 and IRF4, we previously generated FLAG-tagged versions (described in sections 4.3 & 4.5.3) of EGR2, IRF4 and PU.1 (obtained from Julia Minderjahn) for which a ChIP-grade anti-FLAG antibody is available. Then, these constructs were introduced into iDCs (day 4 of differentiation) via mRNA electroporation (see section 4.1.5). Cells were harvested six hours after transfection and subjected to ChIP-seq experiments (including dual crosslinking) as described in section 4.7.1. Since our group possesses a proper ChIP-grade antibody against endogenous PU.1, ChIP-seq experiments for the native PU.1 protein (generated by Julia Minderjahn) were also included in our analyses (see Figure 5.11a).

As illustrated in Figure 5.11a, there is a similar distribution of ChIP-seq peaks between endogenous PU.1 and its FLAG-tagged version, suggesting that the binding of the latter does not interfere with the normal binding patterns of wild type PU.1.

Figure 5.11a also summarizes the distribution of ChIP-seq data for EGR2 and IRF4 FLAG-tagged versions as well as 5hmC enrichment (generated by Sandra Schmidhofer) and ATAC-seq coverage across the 7610 iDC-specific DMRs. Based on TF binding (ChIP-seq) and chromatin accessibility (ATAC-seq) patterns, the iDC-specific DMRs were divided into three groups. The “TF peak” group (blue coloring) encompasses the DMRs that overlapped with one or more TF peaks. The “open” group (yellow coloring) comprises the DMRs that exhibited open chromatin but no detectable binding by

PU.1, IRF4 or EGR2. The “no peak” group (red coloring) includes the remaining DMR that were neither accessible nor bound by any TF.

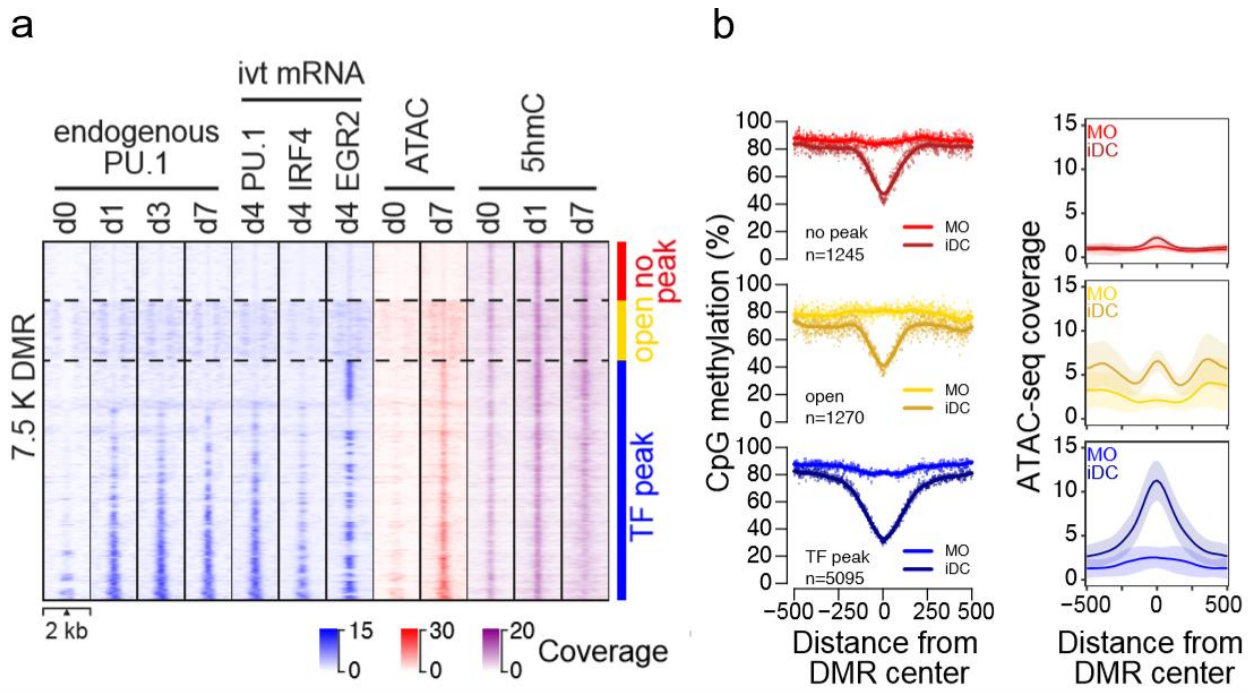

Figure 5.11 Distribution of iDC-DMRs into three groups based on epigenomic landscapes

a Distribution of ChIP-seq (native PU.1, and FLAG-tagged versions of PU.1, IRF4 and EGR2 derived from transfected, *in vitro* transcribed (ivt) mRNA), ATAC-seq and 5hmC signals across the 7610 iDC-specific DMR at the indicated time points of iDC differentiation. DMRs are divided into three majors groups characterized by detectable TF binding (blue, “TF peak”), the presence of open chromatin without detectable TF binding (yellow, “open”), and the absence of open chromatin and lack of detectable TF binding (red, “no peak”). **b** Genomic distance distribution of averaged DNA methylation ratios (left panel) and ATAC-seq coverage (right panel) in MO and iDCs centered on iDC-specific DMR in “TF peak”, “open” and “no peak” groups.

As presented in Figure 5.11b (left panel), all DMR groups were actively demethylated during MO differentiation, even though the “TF peak” subset of regions (bottom panel) presented a more pronounced reduction in DNA methylation ratios compared to the “open” (middle panel) and “no peak” (top panel) groups. The number of iDC-specific DMRs included within each individual DMR group is depicted in the lower left corner of each panel.

In Figure 5.11b (right panel), the ATAC-seq pattern (middle panel) in the “open” group suggests that DMR accessible sites (detected as peaks in the plot) are found in close proximity (separated by a single nucleosome) with neighboring accessible sites. As expected, the “no peak” group (top panel) did not show signs of open chromatin and the “TF peak” group (bottom panel) displayed a robust increase on chromatin accessibility upon differentiation (in iDCs).

In addition, *de novo* motif searches as well as motif co-associations networks across the three DMR groups were also determined as given in Figure 5.12.

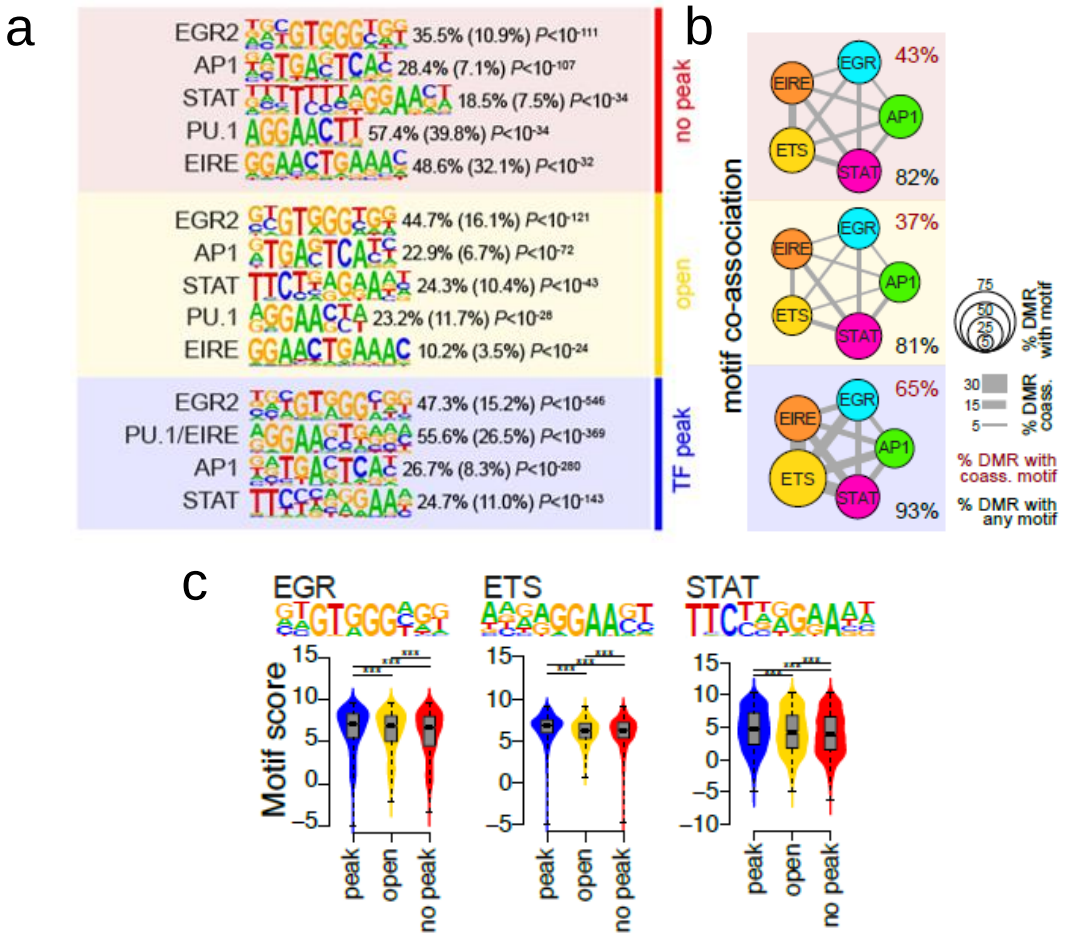

Figure 5.12 TF motifs associated with DMR groups ("TF peak", "open" and "no peak")

a *De novo* identified sequence motifs across iDC-specific DMR in "TF peak", "open" and "no peak" groups as defined in Fig. 5.11a. **b** Motif co-association networks for "TF peak", "open" and "no peak" groups. The size of each node represents the motif enrichment (fraction of peaks) and co-associated TF motifs are indicated by coloring. Edge thickness indicates the frequency of motif co-association. The fraction of DMR with co-associated motifs and the fraction of DMR with any motif are given above and below each network, respectively. **c** EGR, ETS and STAT motif log odds score distribution is shown for "TF peak", "open" and "no peak" groups. The median of the specific distribution across DMR is depicted inside the bean with a conventional boxplot. The asterisks on the top indicate the significance of the motif score enrichment between the different groups (***P<0.001; Mann-Whitney test).

Despite a similar TF signature across the three DMR groups (*de novo* motif analyses, Figure 5.12a), the strength of TF motif co-associations across these regions was distinct. In particular, the "TF peak" group exhibited a higher frequency of motif co-associations (edge thickness) compared to "no peak" and "open" groups (Figure 5.12b). In addition, motif scores for EGR, ETS and STAT were generally higher in the "TF peak" group (Figure 5.12c) compared to the other DMR groups, suggesting that this subset contains higher affinity motifs as well as more opportunities for TF cooperativity.

5.4 Effects of key TF-depletion on iDC differentiation & associated epigenetic processes

As described in section 5.3, IRF4 and EGR2 were considered valid candidates to drive *de novo* demethylation processes through MO differentiation. As a consequence, both TFs were transiently (and individually) knocked-down in MO and their functional consequences on iDC differentiation, gene expression, chromatin accessibility and DNA demethylation were determined. Of note, in this set of experiments two different siRNAs targeting the same gene, specifically siRNA.a and siRNA.b were used as an additional control. The results are summarized in the next chapters.

5.4.1 Impact of IRF4 knock-down on iDC differentiation & gene expression

Several groups have previously investigated the importance of IRF4 on DC development and function, although most of these experiments were performed in murine models[190–192]. In this thesis, the role of IRF4 in iDC differentiation was determined using primary human MO and our well-established *in vitro* model of IL-4/GM-CSF-mediated MO differentiation.

Initially, ten siRNAs against human IRF4 were designed/ordered from Axolabs and further tested using a dual-luciferase reporter assay system.

As illustrated in Figure 5.13, three Gblock gene fragments including the different IRF4 siRNAs were designed and cloned into the *Renilla* luciferase vector. Each fusion construct (three in total) was electroporated together with the respective siRNA (in THP-1 cells) and twenty-four hours post-transfection, *Renilla* and firefly luciferase activities were measured as described in section 4.1.6. A decrease in the *Renilla* Luminescence (Renilla luciferase signal) indicates that the siRNA is specific to the target mRNA. Compared to control (*Renilla* luciferase vector without target gene fragment) all ten siRNAs were specific to IRF4, but the siRNAs 2931 (siIRF4.a) and 2384 (siIRF4.b) exhibited the best silencing efficiency. In a next step, the knock-down efficiency of both siRNAs at the mRNA and protein expression level was determined as depicted in Figure 5.14.

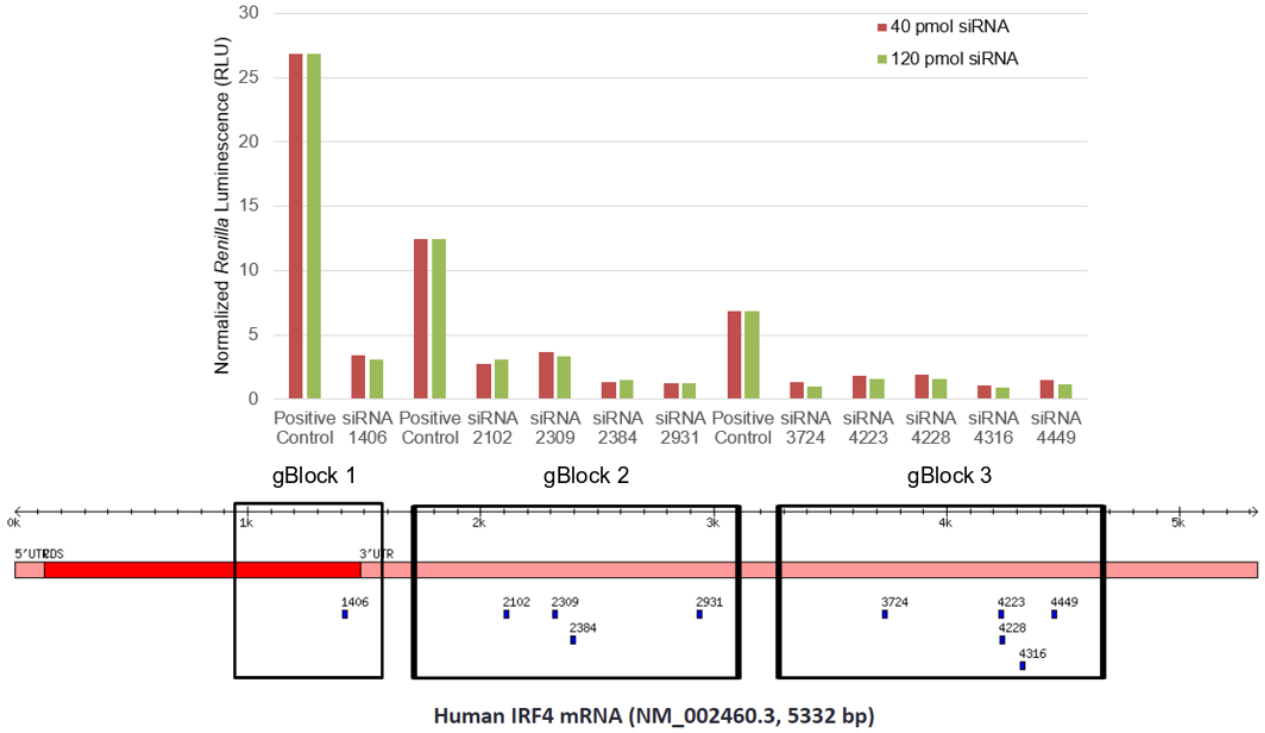

Figure 5.13 IRF4 siRNAs selection based on luciferase activity

Illustration of the *Renilla* luciferase data normalized to firefly luciferase data. Reporter plasmid alone was used as a positive control and two different concentrations (40 pmol and 120 pmol) of each siRNA (ten siRNAs) were tested. Below plot are depicted the positions for the ten siRNAs across the human IRF4 mRNA. Boxes correspond to the genomic region included in each predesigned gBlock (gBlocks were individually cloned into the *Renilla* luciferase vector).

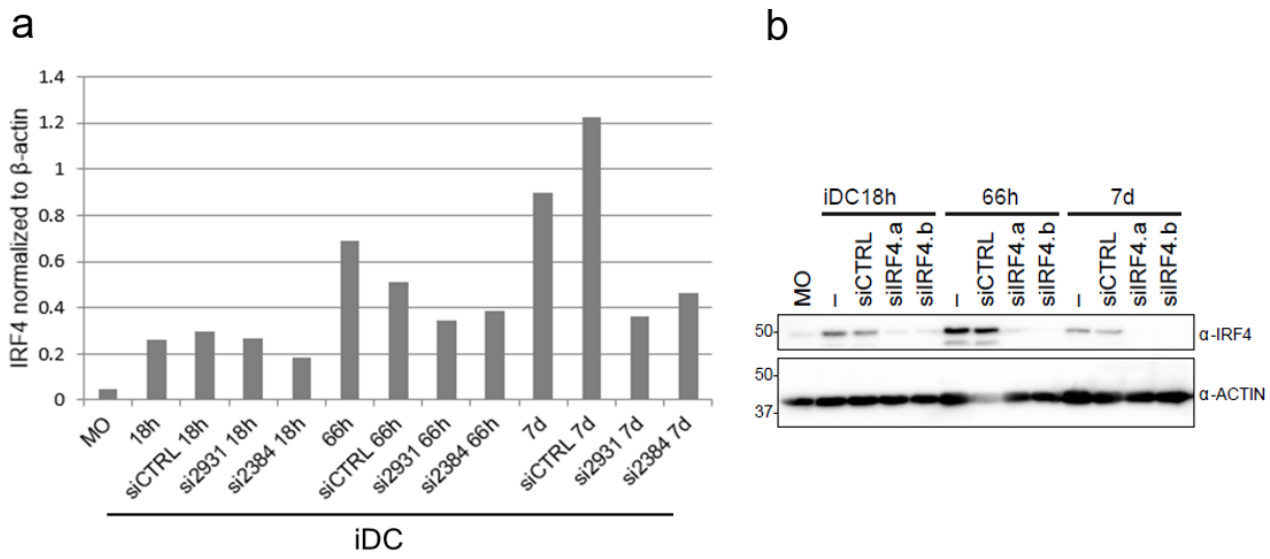

Figure 5.14 IRF4 KD efficiency at the mRNA and protein level

Quantification of IRF4 mRNA (**a**) and protein (**b**) expression levels in MO, iDCs and siIRF4- (two different siRNAs) or siCTRL-treated iDCs at the indicated time points (18h, 66h and 7d). IRF4 real-time PCR data (n=2) in **a** was normalized to the β-actin (housekeeping gene). Blots in **b** were stained with α-IRF4 and α-ACTIN antibodies (the latter represents the protein loading control).

As shown in Figure 5.14a IRF4 was upregulated during in IL4-GM-CSF-driven MO differentiation with an increased expression at the later stages of differentiation (day 7). As expected, siRNA-IRF4 treatment (si2931 or si2384) reduced IRF4 mRNA expression, which was more pronounced after 66h and 7d of culture.

Figure 5.14b depicts the IRF4 KD efficiency on protein level as determined by western blot (see section 4.6.3). Accordingly, IRF4 KD in MO prevented its upregulation during differentiation (18 h, 66 h and 7 d) compared with control samples (iDCs and siCTRL). Moreover, the results between siIRF4.a and siIRF4.b were consistent and reproducible.

A summarized view of IRF4 KD effects on iDC differentiation is given in Figure 5.15, including the general experimental setup to knock-down IRF4 during MO to iDC differentiation.

As indicated in Figure 5.15b, IRF4-deficient cells (siIRF4) failed to differentiate into iDCs, leading instead to the acquisition of a MAC-like phenotype with no significant impact on cell viability when compared with the siCTRL sample (Figure 5.15a, bottom panel).

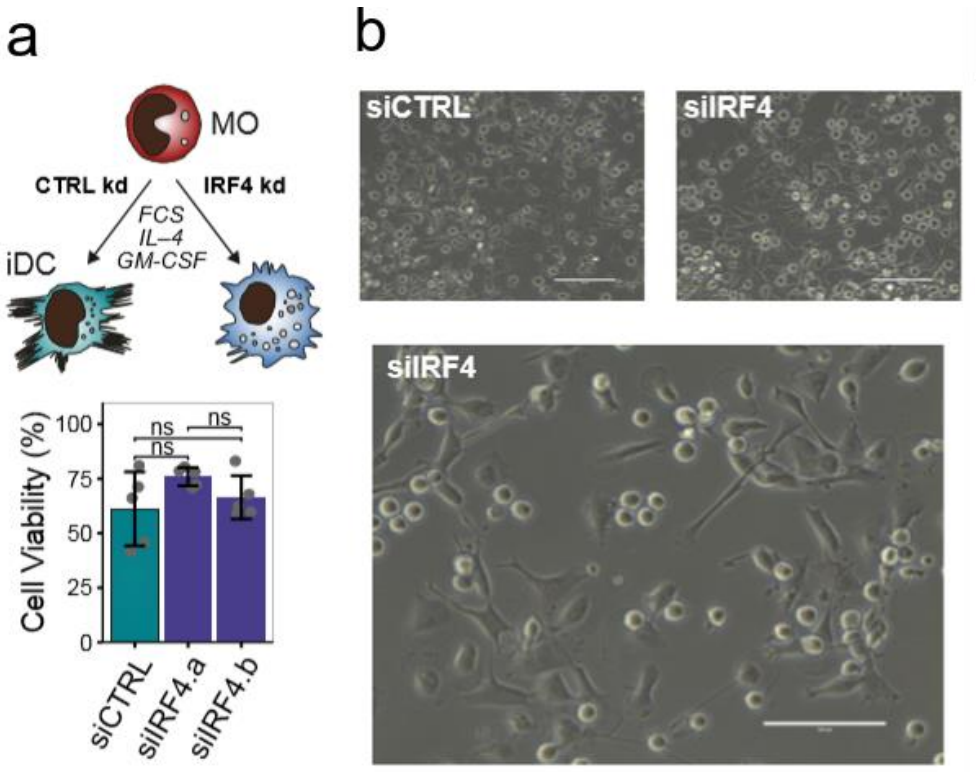

Figure 5.15 Effect of IRF4 KD on iDC differentiation

a Schematic of the experimental setup and effect of siIRF4 treatment on iDC viability (measured at day 7 of differentiation). **b** Representative light microscopy images of iDC (d7) treated with siCTRL or siIRF4.a (40x magnification, bar represents 100 µm).

In a next step, the effects of IRF4 depletion on gene expression were determined by RNA-seq. For this purpose, cells were harvested at day 7 of differentiation and RNA was isolated and purified as

previously described (see section 4.5.1). Then, RNA-seq was performed in siIRF4- and siCTRL-treated iDCs as described in section 4.7.3. Mock (electroporation without siRNA) and iDCs samples were used as additional controls. Due to the MAC-like phenotype observed upon IRF4 KD, RNA-seq data from MAC samples (data generated by Julia Minderjahn) were also included in our analyses to facilitate the interpretation of results. The gene expression data is presented in Figure 5.16.

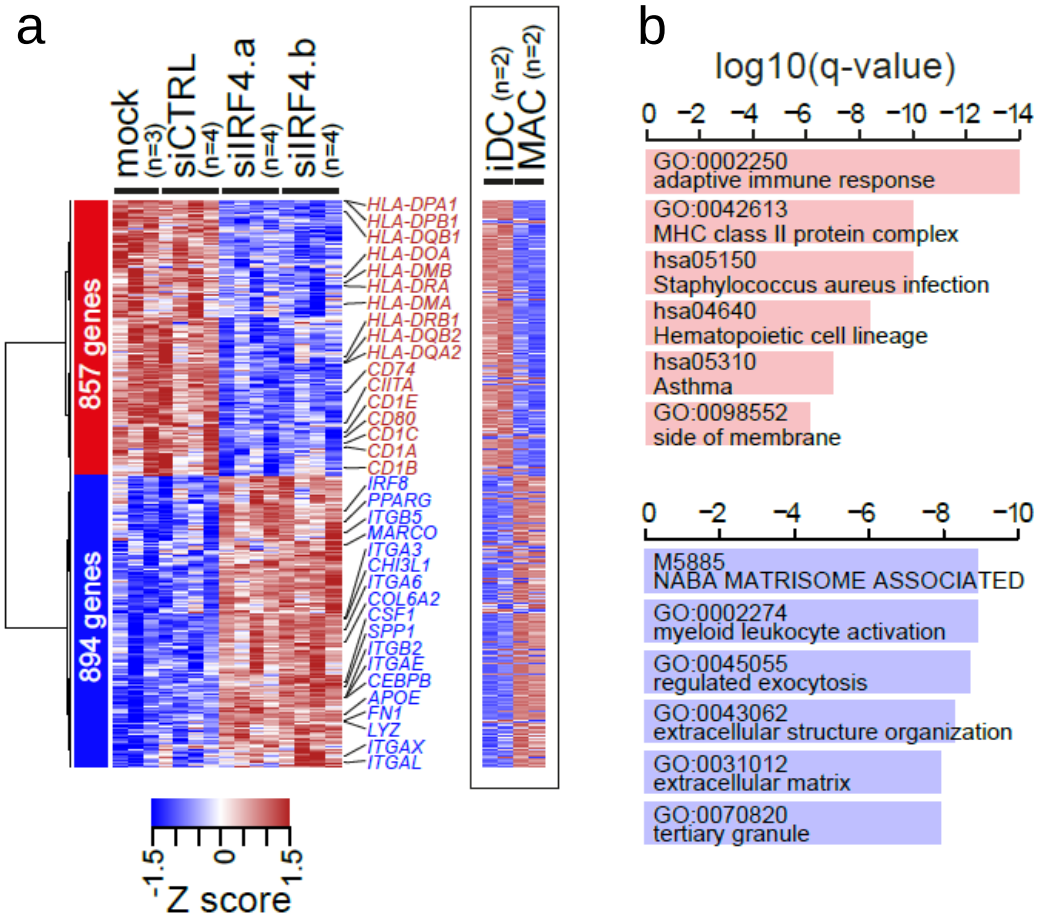

Figure 5.16 Effect of IRF4 depletion on gene expression (RNA-seq)

a Heatmap representing hierarchically clustered and scaled expression data of differentially expressed genes in control- (mock, siCTRL) versus siIRF4-treated cells (absolute logFC > 1, logCPM and logRPKM > 1 and a FDR < 0.05). Each column corresponds to an individual donor. Genes of interest in each cluster are highlighted. The boxed heat map represents scaled expression data of the same gene set (in the same order) of independent cultures of iDC and MAC. **b** Gene Ontology (GO) terms associated with genes upregulated (red bars) or downregulated (blue bars) in siCTRL– compared to siIRF4– treated cells, as analysed by Metascape. Bars represent corrected, log-transformed *P* values (q values) of the GO term enrichment.

As illustrated in Figure 5.16a, the results for siIRF4.a and siIRF4.b were consistent and reproducible among donors. In total, 1751 differentially expressed genes between siIRF4-treated cells and control samples (mock and siCTRL) were detected, which are distributed into two major clusters. As shown in Figure 5.16a, 857 genes were downregulated in siIRF4-treated samples compared to controls, including genes involved in the regulation of the immune system in humans (genes highlighted in red). In contrast, 894 genes were upregulated in siIRF4-treated samples, encompassing genes that are

primarily expressed in MAC (genes highlighted in blue). Supporting these observations, the heatmap for iDC and MAC samples (right heatmap), matched gene expression profile observed in control (mock and siCTRL) and siIRF4-treated cells, respectively.

The barplots (showing gene ontology (GO) analysis) in Figure 5.16b also indicate that genes downregulated upon siIRF4 KD are mainly associated with adaptive immune responses and MHC class II protein complex (top barplot). Conversely, genes upregulated in IRF4-deficient cells are matrisome associated and expressed in human MAC. Furthermore, these genes are also involved in myeloid leukocyte activation and regulation of exocytosis processes (bottom barplot).

5.4.2 Impact of IRF4 knock-down on chromatin accessibility

In addition to transcriptional profiling, we also intended to determine the effect of IRF4 depletion on chromatin accessibility (ATAC-seq). For this purpose, cells were harvested at day 7 of differentiation and ATAC-seq experiments were conducted as described in section 4.7.2. The ATAC-seq data results are given in Figure 5.17.

As shown in Figure 5.17a, 2766 differentially accessible sites between siCTRL and IRF4-depleted samples were detected, which are distributed into two major clusters. Upon IRF4 depletion, 1298 sites lost accessibility (top cluster) and 1468 sites gained accessibility (bottom cluster). *De novo* motif searches across the less accessible regions (Figure 5.17a, top motifs) revealed ETS:IRF composite motifs (EIRE) as well as AP1, PU.1 and IRF motifs. On opposite, *de novo* motif searches across the more accessible sites (Figure 5.17a, bottom motifs) revealed an enrichment for CEBP, AP1, EGR and E-box TFs, which are key motifs associated with MACs.

Focusing on DNA methylation changes during MO differentiation, results in Figure 5.17b (top panel) showed that most of the less accessible sites in IRF4-depleted cells were clearly demethylated during MO differentiation with a significant number of these sites already starting the demethylation process at MO stage (progressive demethylated regions). These findings suggest that sites that loose accessibility in IRF4-depleted cells are likely targets of DNA demethylation machinery.

In contrast, the most of more accessible sites in IRF4-depleted cells did not show significant changes in DNA demethylation over differentiation when compared to MO (Figure 5.17b, bottom panel). These data suggest that most of the sites that gain accessibility in IRF4-depleted cells remain stably demethylated during MO differentiation.

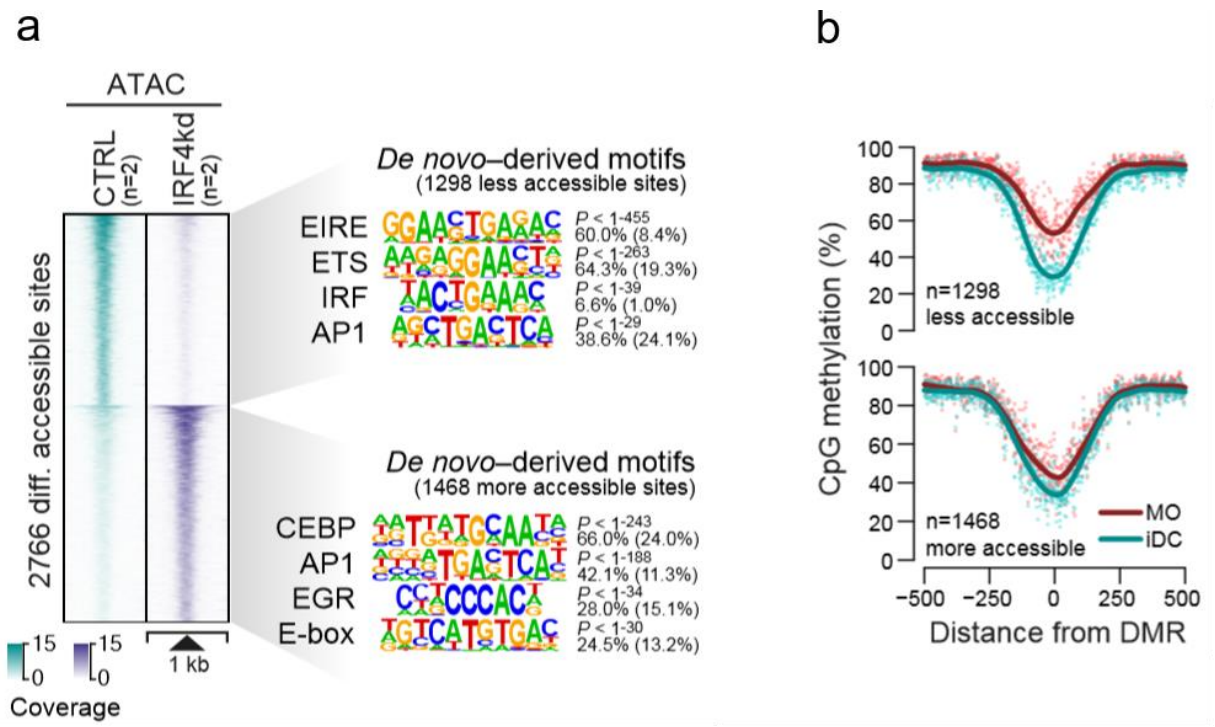

Figure 5.17 Effect of IRF4 depletion on chromatin accessibility (ATAC-seq)

a Distribution of ATAC-seq signals across the differentially accessible sites between siCTRL– and siIRF4– treated samples. *De novo*–derived motifs for each cluster are given along with the significance of motif enrichment (hypergeometric test) and the fraction of motifs in peaks (background values are in parenthesis). Top motifs corresponding to known factor families are shown for each cluster. **b** Genomic distance distribution of averaged DNA methylation ratios in MO and iDCs centered on differentially accessible sites (as introduced in a).

5.4.3 Impact of EGR2 depletion on iDC differentiation & gene expression

Like IRF4, EGR2 was also identified as a potential candidate for recruitment of DNA demethylation machinery during MO differentiation. Hence, in this chapter the effect of its depletion on iDC differentiation and cell survival was addressed.

Initially, a siRNA set (ten siRNAs) against EGR2 was tested (data not shown) by Corinna Kirschner and siRNAs with the best silencing efficiency (si1132 (a) and si2665 (b)) were used in further experiments.

The general experimental setup to knock-down EGR2 in MO and to assess its impact on iDC morphology and cell viability is shown in more detail in Figure 5.18.

As indicated in Figure 5.18b EGR2 depletion strongly affected iDC morphology (e.g., lack of cell dendrites) compared to siCTRL sample, and was associated with a statistically significant reduction in cell viability (Figure 5.18a, bottom panel).

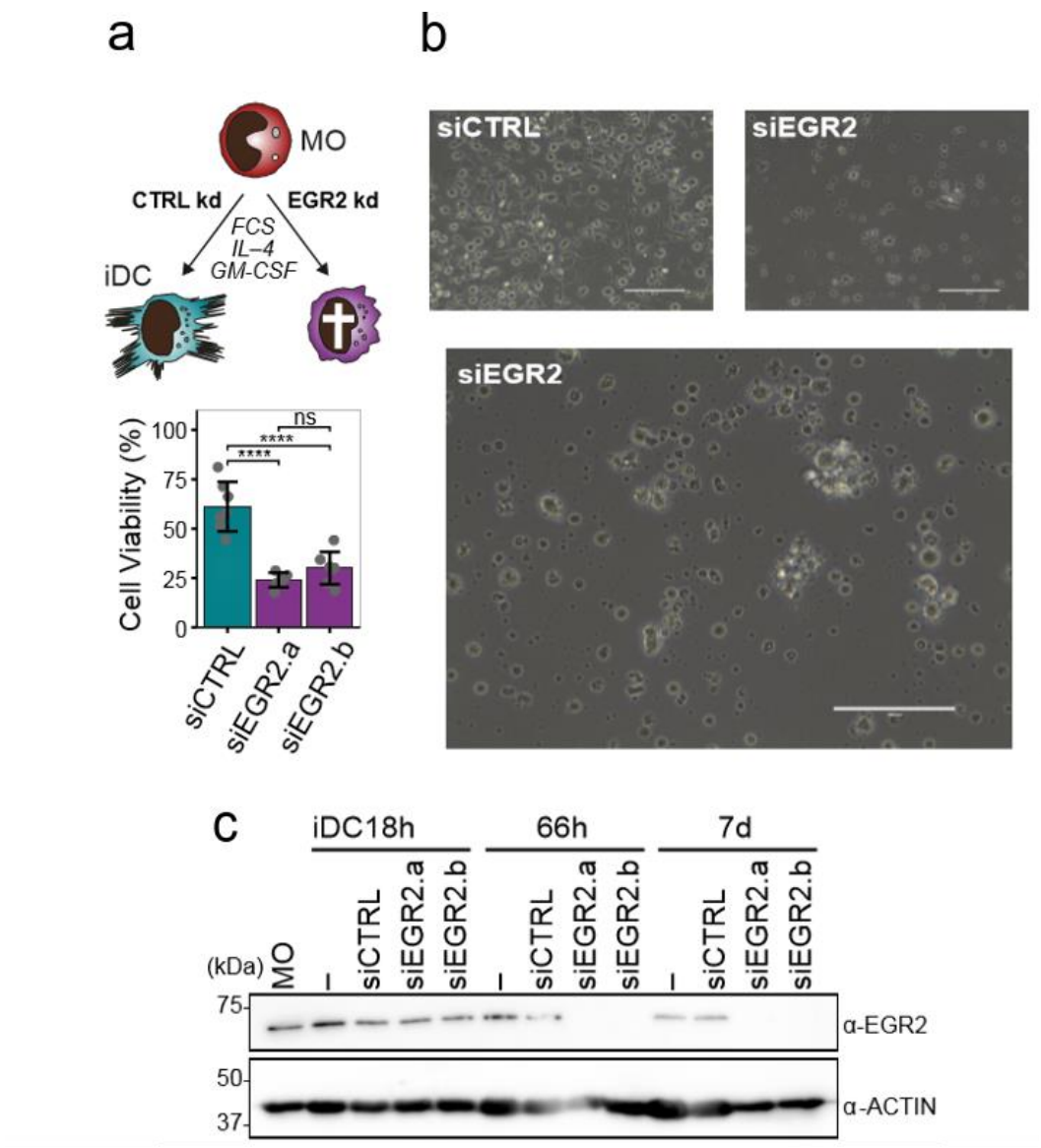

Figure 5.18 Effect of EGR2 depletion on iDC differentiation and survival

a Schematic of the experimental setup and effect of siEGR2 treatment on iDC viability (measured at day 7 of differentiation). **b** Representative light microscopy images of iDC (d7) treated with siCTRL or siEGR2.b (40x magnification, bar represents 100 µm). **c** EGR2 protein expression levels in MO, iDCs and siIRF4- (two different siRNAs) or siCTRL-treated iDCs at the indicated time points (18h, 66h and 7d). Blots were stained with α-EGR2 and α-actin antibodies (the latter represents the protein loading control).

Figure 5.18c depicts the EGR2 KD efficiency on protein level as determined by Western blot (see section 4.6.3). Since some EGR2 protein expression was already detected in MO, the impact of EGR2 KD was not immediate and only became more pronounced after 66 h and 7 d of culture as compared to control samples (iDCs and siCTRL). The results between siEGR2.a and siEGR2.b were consistent and reproducible.

In a next step, the effects of EGR2 depletion on genome-wide gene expression were determined by RNA-seq. For this purpose, cells were harvested at day 7 of differentiation and RNA was isolated and

purified as previously described (see section 4.5.1). Then, RNA-seq was performed as described in section 4.7.3 and the results are summarized in Figure 5.19.

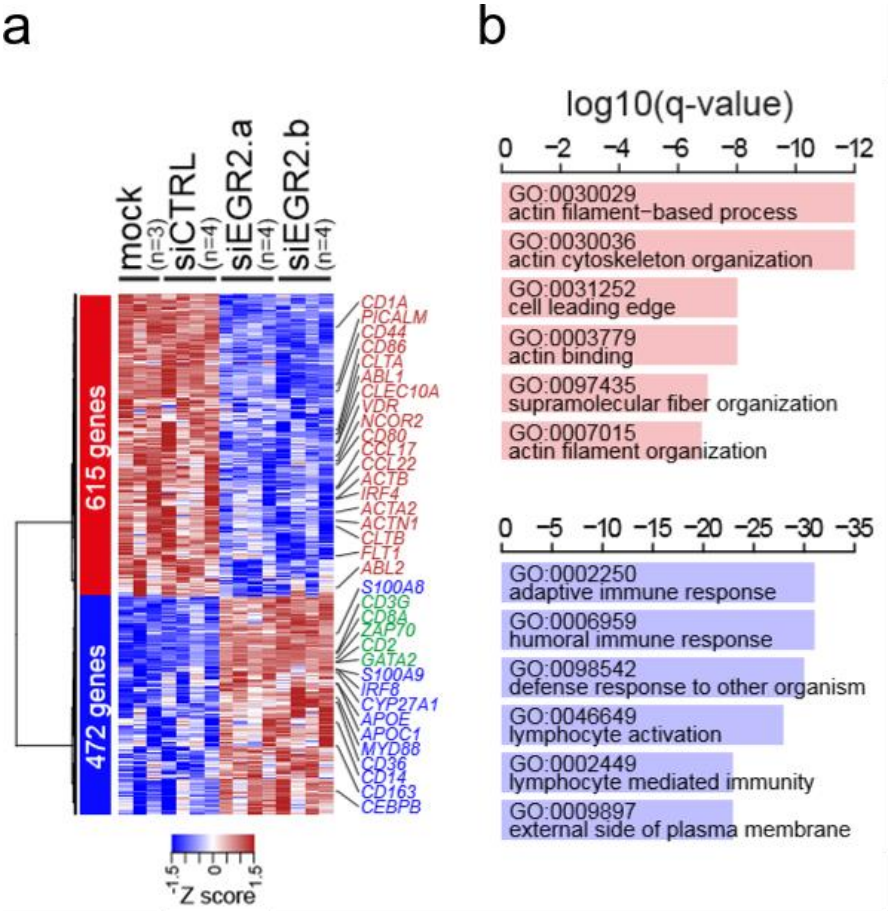

Figure 5.19 Effect of EGR2 depletion on gene expression (RNA-seq)

a Heat map presenting hierarchically clustered and scaled expression data of differentially expressed genes in control- (mock, siCTRL) versus siEGR2-treated cells (absolute logFC > 1, logCPM and logRPKM > 1 and a FDR < 0.05). Each column corresponds to an individual donor. Genes of interest in each cluster are highlighted. Genes typically expressed in contaminating lymphoid cells are marked in green. Each column corresponds to an individual donor and each row represents a single differential expressed gene. **b** Gene Ontology (GO) terms associated with genes upregulated (red bars) or downregulated (blue bars) in siCTRL– compared to siEGR2– treated cells, as analysed by Metascape. Bars represent corrected, log-transformed *P* values (q values) of the GO term enrichment.

As illustrated in Figure 5.19a, the results between siEGR2.a and siEGR2.b were consistent and reproducible among donors. In total, 1087 differentially expressed genes between siEGR2-treated cells and control samples (mock, siCTRL) were detected, which are distributed into two major clusters. As seen in Figure 5.19a, 615 genes were downregulated in EGR2-deficient samples compared to controls, including genes necessary for a normal iDC development and function, such as NCOR2, CCL22 and IRF4 (genes highlighted in red). Moreover, genes encoding DC cell surface markers, such as CD1A, CLEC10A and CD86 were also downregulated, emphasizing the strong impact of EGR2 depletion on iDC differentiation.

On the other hand, 472 genes were upregulated in EGR2-deficient samples compared to controls, containing some important genes for the regulation of inflammatory processes and immune response (genes highlighted in blue) as reinforced by the functional annotation in Figure 5.19b, bottom panel. In addition, genes upregulated in MO cultures upon EGR2 KD also included a number of T-cell specific markers (genes highlighted in green), suggesting that few T cells contaminated the MO preparation.

Functional annotation in Figure 5.19b unveiled a significant impact of EGR2 depletion on genes involved in the organization of actin filaments during MO differentiation (top panel), which in turn, fits well with the observed cell morphology (absence of cell dendrites in siEGR2-treated cells).

In general, the gene expression data revealed a pronounced impact of EGR2 depletion on transcriptional programs, which supports the observed morphological effects.

5.4.4 Impact of EGR2 depletion on chromatin accessibility

Since EGR2 is essential for iDC differentiation, function and survival, we also checked whether the EGR2 depletion impairs chromatin accessibility. For this purpose, cells were harvested at an earlier time point of differentiation (day 3) since the assay failed to provide good quality data at later time points and ATAC-seq experiments were conducted as described in section 4.7.2. The ATAC-seq results are summarized in Figure 5.20.

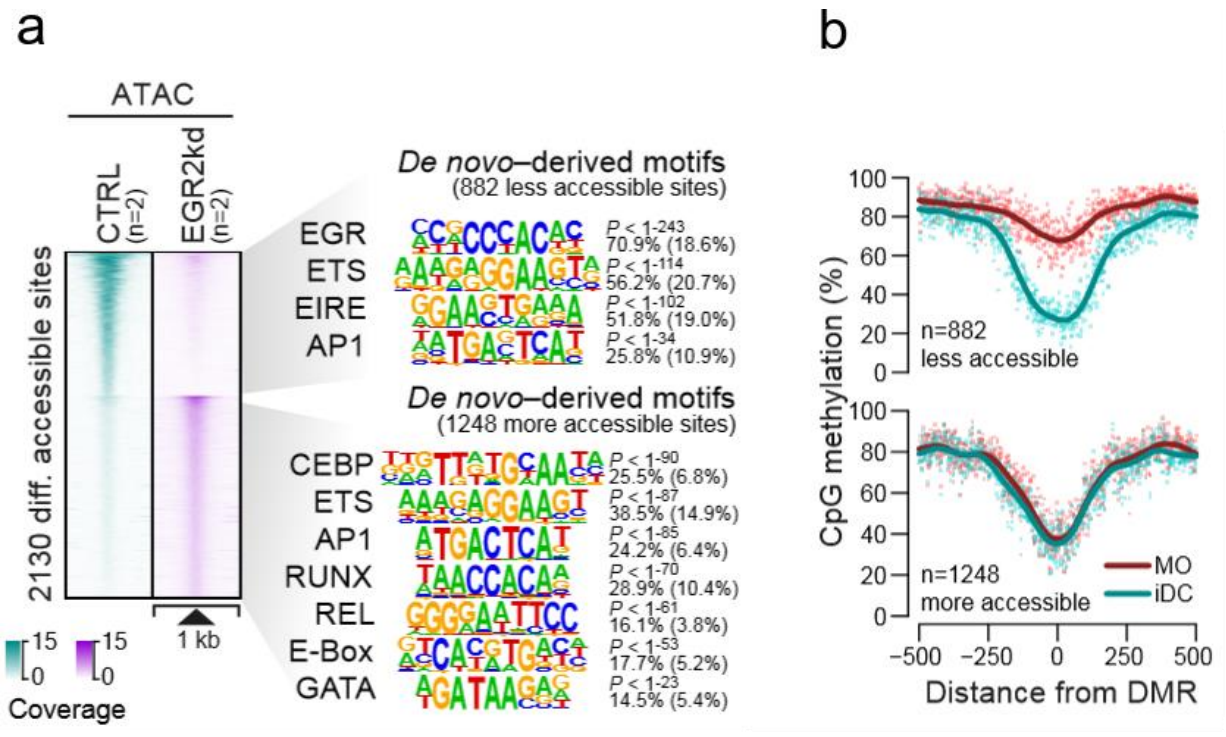

Figure 5.20 Effect of EGR2 depletion on chromatin accessibility (ATAC-seq)

a Distribution of ATAC-seq signals across the differentially accessible sites between siCTRL– and siEGR2– treated samples. *De novo*–derived motifs for each cluster are given along with the significance of motif enrichment (hypergeometric test) and the fraction of motifs in peaks (background values are in parenthesis). Top motifs corresponding to known factor families are shown for each cluster. **b** Genomic distance distribution of averaged DNA methylation ratios in MO and iDCs centered on differentially accessible sites (as introduced in a).

As shown in Figure 5.20a, 2130 differentially accessible sites between siCTRL and EGR2-depleted samples were detected, which are distributed into two major clusters. Upon EGR2 depletion, 882 sites lost accessibility (top cluster) and 1248 sites gained accessibility (bottom cluster).

De novo motif searches across the less accessible sites revealed a strong overlap with EGR motifs and also contained the co-associated motifs for PU.1, ETS:IRF (EIRE) or AP-1 (Figure 5.20a, top motifs). Supporting previous observations on transcriptional level (Figure 5.19a), *de novo* motif searches across more accessible sites (Figure 5.20a, bottom motifs) revealed a mixed motif signature, including myeloid (ETS, CEBP) and T cell motifs (e.g. RUNX, GATA).

Focusing on DNA methylation changes during MO differentiation, results in Figure 5.20b (top panel), showed that most of the less accessible sites in EGR2-depleted cells were markedly demethylated during MO differentiation. These findings suggest that sites that loose accessibility in EGR2-depleted cells were targets of *de novo* DNA demethylation, thus proposing an epigenetic pioneering role of EGR2 in the DNA demethylation process.

In contrast, the most of more accessible sites in EGR2-depleted cells did not show significant changes in DNA demethylation over differentiation when compared to MO (Figure 5.20b, bottom panel). These data suggest that most of the sites that gain accessibility in EGR2-depleted cells remained stably demethylated during MO differentiation.

5.4.5 Impact of IRF4- & EGR2-depletion on DNA methylation

Considering the impact of IRF4- and EGR2-depletion on iDC differentiation, gene expression and chromatin accessibility, their effect on DNA methylation was evaluated. Beforehand, cells were harvested at day 7 of differentiation and DNA (including MO) was purified and quantified as described in section 4.4.1. Then, we selected iDC DMRs normally bound by IRF4, EGR2 or PU.1 and the DNA methylation ratios at these regions in siIRF4- and siEGR2-treated cells as well as in control samples were determined by MassARRAY EpiTYPER.

The EpiTYPER experiments were performed as described in section 4.4.3, using the aforementioned DNA samples and appropriated primers (primer sequences can be found in section 3.10.3) to amplify bisulfite-converted DNA. Eight example regions are shown in Figure 5.21.

Figure 5.21 provides IGV genome browser tracks for the indicated epigenomic data sets at example regions around binding sites of EGR2, IRF4 or PU.1. As shown in Figure 5.21a, this region was bound by IRF4 (ChIP) and failed to gain accessibility (ATAC) upon IRF4 KD. However, EpiTYPER data (bottom panel) showed no significant changes on DNA methylation ratios (upon IRF4 KD) compared with siCTRL or even with EGR2 KD samples. Similar results were observed in a region bound not only by IRF4 but also by EGR2 and PU.1 (co-factor of IRF4) (Figure 5.21b). On the opposite, regions bound by EGR2 (Figure 5.21b-f) both failed to demethylate (high DNA methylation ratios in siEGR2-treated iDCs) or gain accessibility (ATAC) upon EGR2 KD.

Interestingly, at some genomic positions bound by EGR2, its depletion delayed DNA demethylation even at nearly inaccessible iDC-DMRs (Figure 5.21g,h).

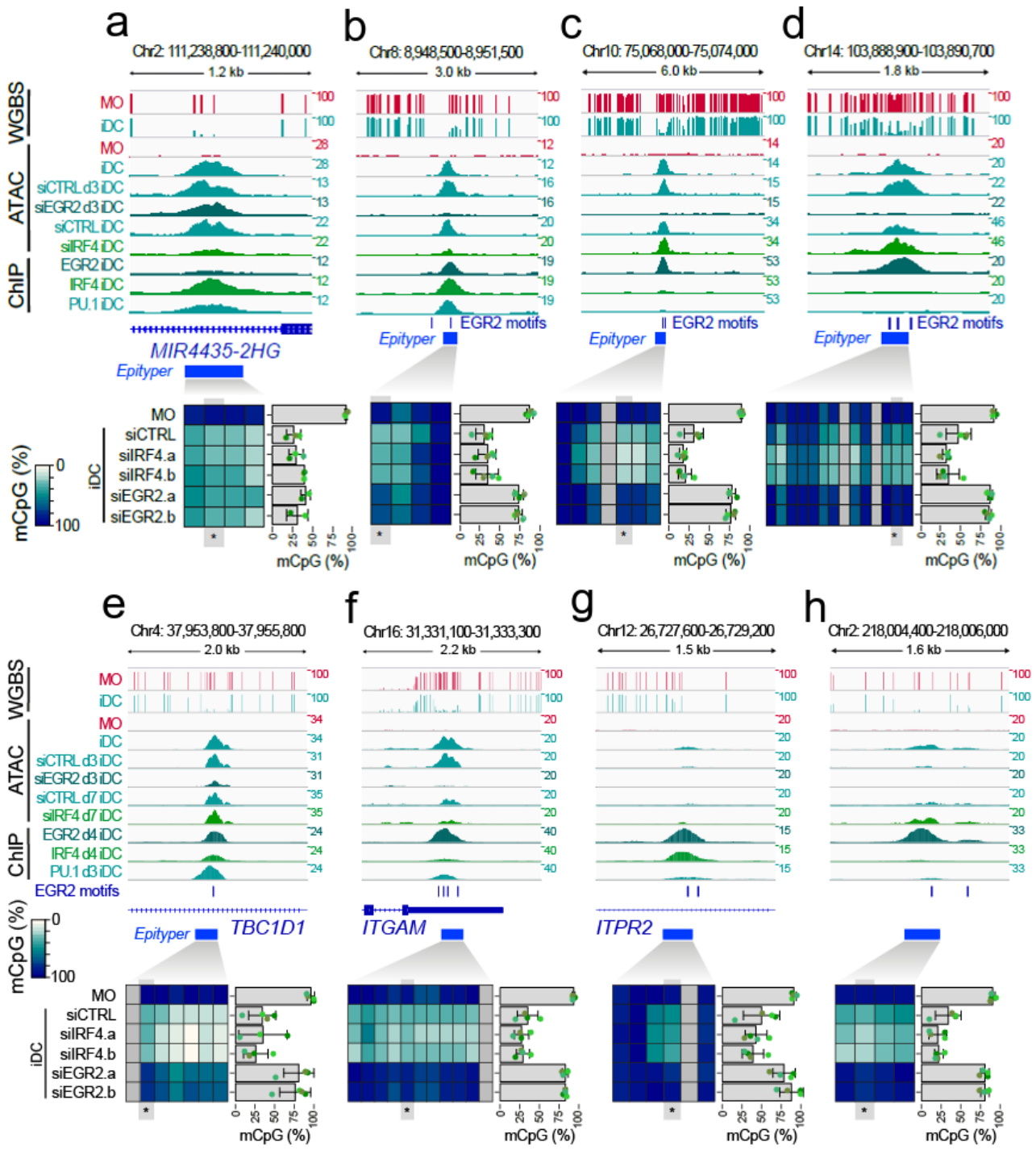

Figure 5.21 Effects of IRF4- and EGR2-depletion on DNA methylation (EpiTYPER)

a–h IGV genome browser tracks for the indicated WGBS, ATAC- and ChIP-seq data sets at example regions around binding sites of EGR2, IRF4 or PU.1. Corresponding DNA methylation ratios for the indicated regions are given below. For heatmaps, methylation ratios (representing means of n ≥ 4) are indicated by coloring (white: no methylation, dark blue: 100% methylation) with each column representing a single CpG. For each region the data of a single CpG (highlighted and marked by asterisks) is shown. Bars represent means ± SD, individual data points are shown as colored dots (each color representing a different donor).

These data suggest a critical role of EGR2 (and not of IRF4) in regulating *de novo* DNA demethylation processes during MO towards iDC differentiation.

5.4.6 Overview of IRF4- & EGR2-depletion effects

Overall, the effects of EGR2 depletion were distinct from those observed upon IRF4 KD. Despite their impact on iDC morphology (Figure 5.15b & Figure 5.18b), EGR2 depletion negatively affected cell survival (Figure 5.18a), while IRF4 KD did not impair cell viability compared to control sample (Figure 5.15a). Similarly, both TFs-depletion showed marked effects on genome-wide gene expression (see Figure 5.16 & Figure 5.19), although the transcriptional programs were clearly distinct as summarized in Figure 5.22.

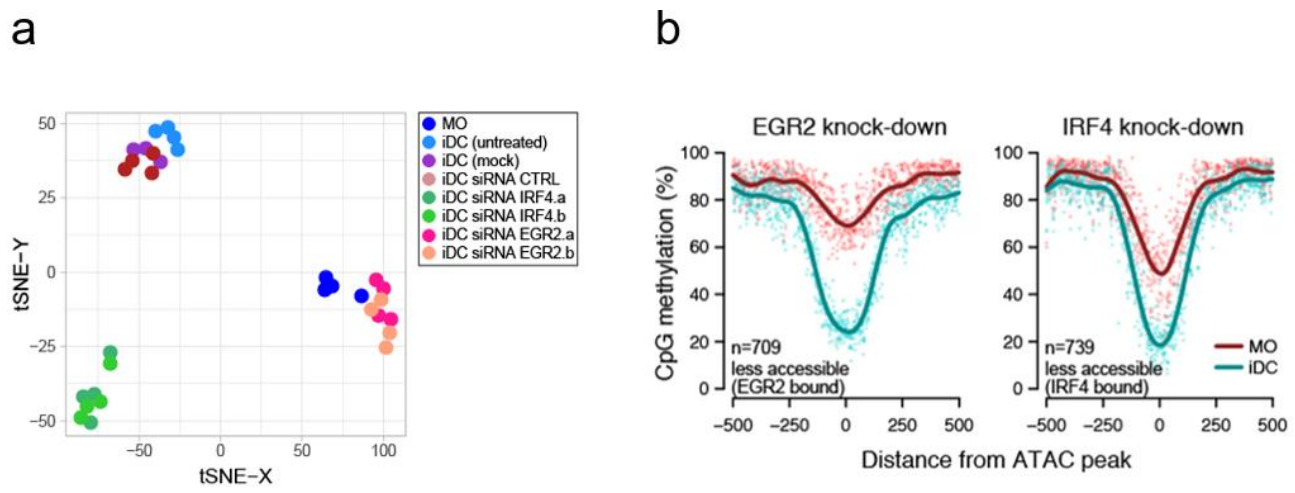

Figure 5.22 Effects of IRF4- and EGR2- depletion

a Two-dimensional visualization of the RNA-seq sample distribution using tSNE embedding. Replicates of the same treatment are indicated by coloring. **b** Side-by-side comparison of genomic distance distributions of averaged DNA methylation ratios in MO and iDCs centered on differentially accessible and TF bound sites (as introduced in Figure 5.17a and Figure 5.20a).

The tSNE (t-Distributed Stochastic Neighbor Embedding) plot in Figure 5.22a illustrates the RNA-seq samples distribution, including all treatments and respective replicates. As shown, the siIRF4-treated iDCs (green coloring) cluster together and separate from siEGR2-treated iDCs (pink coloring). These data indicate distinct and pronounced (both treatments cluster far apart from iDC control samples) effects of IRF4- and EGR2-depletion on gene expression.

Nevertheless, the effects of EGR2 depletion on DNA methylation (Figure 5.21) were clearly stronger compared to IRF4 depletion with many DMR failing to demethylate. Like EGR2, IRF4 KD also impaired chromatin accessibility, however, regardless of methylation changes (Figure 5.21a,b).

Focusing on less accessible sites (upon KD, as shown in Figure 5.17a & Figure 5.20a) that were normally bound by either factor, we verified that less accessible sites in EGR2-depleted cells were

strongly demethylated during normal MO differentiation, whereas less accessible sites in IRF4-depleted cells already initiated demethylation at an earlier stage of MO development (Figure 5.22b). These data suggest an epigenetic pioneering role of EGR2 in mediating *de novo* demethylation events.

5.5 EGR2 function in *de novo* DNA demethylation processes

As demonstrated in section 5.4.6, EGR2 is presumably involved in the initiation of *de novo* DNA demethylation processes during MO to iDC differentiation. Supporting these findings, we observed that CpGs inside of the EGR2 motif could be protected from demethylation (EGR2 motif overlaps with a highly methylated CpG while the surrounding CpGs appear demethylated). A more comprehensive and detailed analysis about this phenomenon can be found in Figure 5.23 & Figure 5.24.

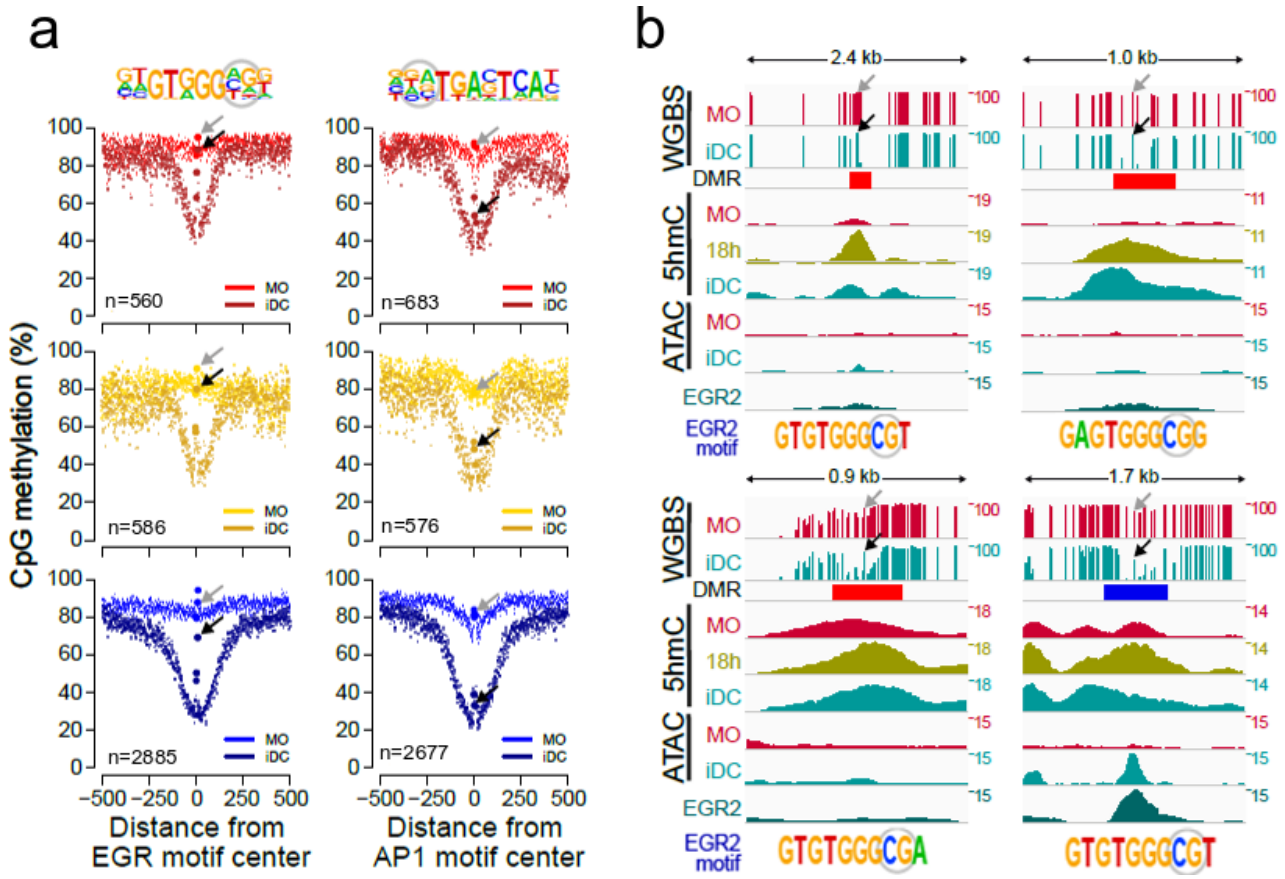

Figure 5.23 Methylation footprint of EGR2

a Genomic distance distribution of averaged DNA methylation ratios in MO and iDCs in iDC-specific DMR in “TF peak”, “open” and “no peak” groups (in blue, yellow and red, respectively, as defined in Fig. 3c) centered on the consensus EGR (left panel) or AP1 motifs (right panel), which are shown on top of each panel (CpG positions that are highlighted by arrows in grey (MO) and black (iDCs) in the lower panels are circled). The number of motif-containing DMR is given in the lower-left corner of each panel. **b** IGV genome browser tracks of example loci showing single CpG methylation ratios (WGBS), 5hmC and ATAC-seq coverage in MO and iDCs and EGR2 ChIP-seq coverage in iDC. Genomic locations of DMR are shown (coloring according to the classification into DMR groups “no peak” or “TF peak” in red or blue, respectively). EGR motif sequences are given below each track and arrows in the top tracks mark the positions of the circled CpG.

As shown in Figure 5.23a, CpGs within the EGR motif remained methylated (left panel) during MO differentiation (black arrows) across all DMR groups (“no peak”, red; “open”, yellow and “TF peak”, blue). In contrast, CpGs within the AP1 motif (right panel) were markedly demethylated upon MO differentiation into iDC (black arrows) across all DMR groups.

In Figure 5.23b some DMR examples are given that show protected CpGs (which remain methylated as indicated by the black arrows) at EGR2 motifs. The top panels and the left bottom panel (“no peak” DMR group, red coloring) are examples of actively demethylated regions (5hmC enrichment) that do not show any signs of open chromatin (ATAC coverage), nor EGR2 binding (track above EGR2 motif). Still, CpGs in EGR2 motif are protected from demethylation.

The same methylation footprint in EGR2 motif is also observed on the right bottom panel, where the demethylated DMR is accessible and bound by EGR2. This DMR is part of the “TF peak” group as indicated by the DMR blue coloring.

Additional examples, including methylation data (EpiTYPER) are provided in Figure 5.24.

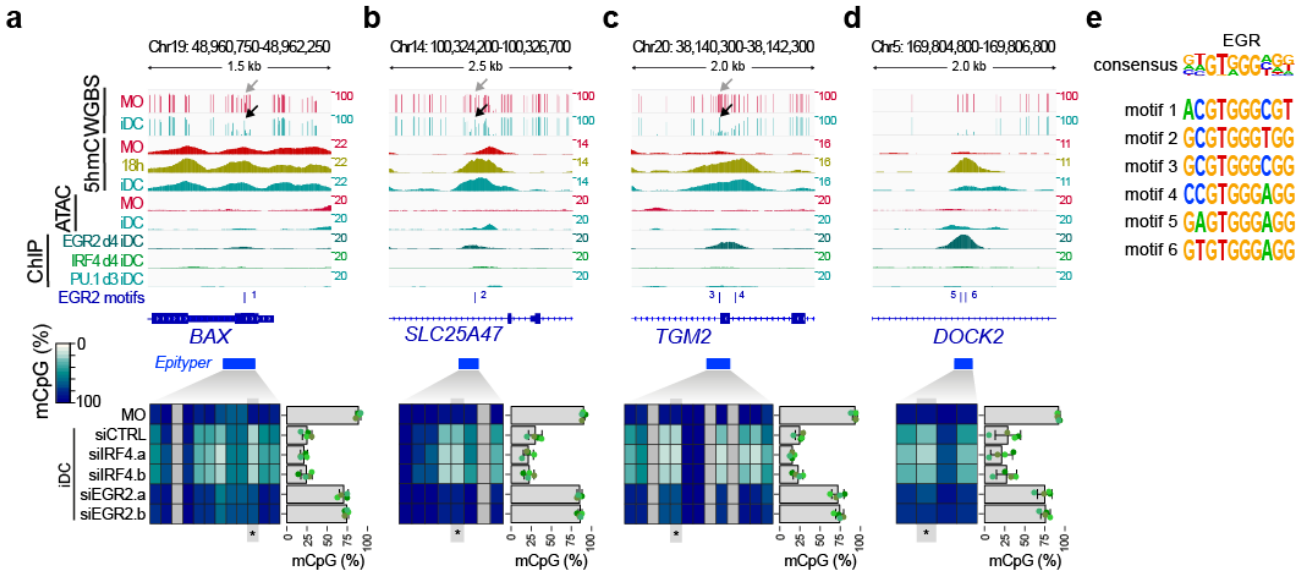

Figure 5.24 DNA methylation at largely inaccessible DMR

a-d IGV genome browser tracks for the indicated WGBS, 5hmC and ATAC data sets at example DMR with EGR2 consensus motifs. CpGs included in the EGR2 motif that resist demethylation are indicated above each WGBS track with arrows. Corresponding DNA methylation ratios (as measured by EpiTYPER) for the indicated regions are given below. For heatmaps, methylation ratios (representing means of n ≥ 4) are indicated by coloring (white: no methylation, dark blue: 100% methylation, gray: not detected) with each column representing a single CpG. For each region the data of a single CpG (highlighted and marked by asterisks) is shown. Bars represent means ± SD, individual data points are shown as colored dots (each color representing a different donor). **e** Motif sequences for the six EGR2 motifs shown in **a-d**. The consensus EGR2 motif derived from DMR regions is shown on top.

Nearly all DMR regions displayed in Figure 5.24 are actively demethylated (5hmC enrichment), inaccessible (ATAC) and unbound (ChIP) by IRF4, PU.1 and EGR2. In some cases, however, we detected a slight binding by EGR2 protein (Figure 5.24d).

In Figure 5.24 are depicted few examples of inaccessible DMR (lack of ATAC signal in iDC) that failed to demethylate in EGR2-depleted cells (compared to siCTRL and siIRF4 samples) and that are unbound by EGR2 (except region in Figure 5.24d). In addition, these regions contain CpGs in the EGR2 motif (Figure 5.24e, motif sequences (1-4)) that are protected from DNA demethylation (as indicated by the black arrows). These data suggest that at these regions, EGR2 may transiently interact with methylated CpGs at its binding sites and initiate the DNA demethylation.

As illustrated in Figure 5.24d, EGR2 may also be able to initiate DNA demethylation at largely inaccessible DMR even in the absence of a methylation footprint in its binding motif sequence.

Taken together, our findings from this section suggest that EGR2 may initiate the DNA demethylation at both transient and stable binding sites.

5.5.1 EGR2 & DNA demethylation machinery recruitment

5.5.1.1 Proximity-dependent Biotin Identification (BioID)

To demonstrate the presence of an interaction between EGR2 and TET2 as well as to uncover other potential *in vivo* interaction partners of EGR2 in iDCs (obtained from *in vitro* MO differentiation) the BioID approach was used (see section 4.6.4).

As a first step of this method, we designed a gBlock gene fragment including a 3xFLAG sequence and the EGR2 protein fused along with a flexible linker to the promiscuous biotin ligase to biotinylate the proteins located in close proximity. In parallel, a negative control gene fragment, containing a 3xFLAG sequence and the biotin ligase fused to a nuclear localization sequence (NLS) was also designed. The respective gBlock gene fragments can be found in section 3.13.

Afterwards, gBlock gene fragments were cloned using the Gibson assembly method (section 4.3.5) and after isolation and linearization of plasmid DNA, the 5'-capped, poly-adenylated mRNAs were obtained by *in vitro* transcription as described in section 4.5.3. Then, we transfected the synthetic mRNAs by electroporation into iDCs-d3 (section 4.1.5) and harvested cells at different time points to measure the fusion proteins kinetic by Western blot (see section 4.6.3 and figure Figure 5.25).

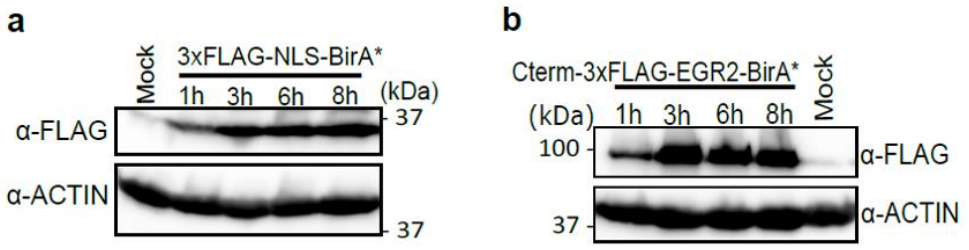

Figure 5.25 Protein kinetic for the BioID fusion proteins at different time points after transfection

Immunoblots depict the protein expression levels for the following BioID fusion proteins at the indicated time points (1 h, 3 h, 6 h and 8 h): **a** 3xFLAG-NLS-BirA* (control) and **b** Cterm-3xFLAG-EGR2-BirA*. Blots were stained with an anti-FLAG antibody and anti-ACTIN antibody was used as protein loading control. A mock sample was also included as negative control.

As depicted in Figure 5.25, both fusion proteins were overexpressed in iDC over the time course (1 h, 3 h, 6 h and 8 h) after transfection as detected by the anti-FLAG (α-FLAG) antibody. As protein expression was most abundant between three and six hours after transfection, we added biotin to culture medium (biotinylation of vicinal proteins) three hours after transfection and harvested the respective lysates eight hours after transfection.

To check whether EGR2 fusion protein would interfere with normal EGR2 binding patterns, we also performed ChIP-seq experiments (cells were harvested six hours after transfection) and an IGV genome browser representation of this data is given in Figure 5.26.

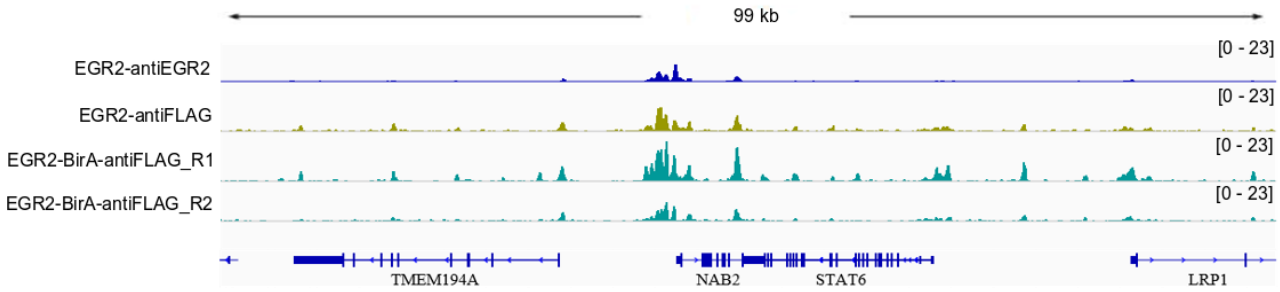

Figure 5.26 IGV genome browser track of EGR2-fusion proteins in iDCs-d3

ChIP-seq signals of EGR2 (obtained with an anti-EGR2 (blue) or anti-FLAG (green) antibody) and EGR2-BirA-antiFLAG replicate one and two (obtained with an anti-FLAG antibody (light blue)) across the indicated genomic loci.

As shown in Figure 5.26, EGR2-BirA fusion protein (replicate one and two) showed a similar binding pattern compared to both endogenous EGR2 (ChIP-seq performed by David Dittmar) and EGR2-antiFLAG. Since the fusion of EGR2 to the biotin ligase did not prevent its binding to DNA, we proceeded with the BioID approach.

For this purpose, iDCs-d3 were individually transfected with EGR2-BirA and NLS-BirA (negative control) fusion proteins and the detailed BioID protocol can be found in section 4.6.4. Biotinylated proteins were bound on Streptavidin-coated magnetic beads and subsequent trypsin digestion and

mass spectrometry analysis was carried out in cooperation with Prof. Dr. Axel Imhof and Dr. Andreas Schmidt at the Zentrallabor fur Proteinanalytik of the University of Munich (ZfP, LMU).

Preliminary mass spectrometry raw data analysis was performed at the ZfP and the subsequent data analyses were performed based on the reported protein iBAQs (intensity-based absolute quantification), a measure of protein abundance. To include peptides where the iBAQ value was equal to zero (and to avoid dividing by zero) a pseudo-count of two was used to replace these iBAQ values. Afterwards, iBAQ values were log2 transformed and all non-reversed, non-contaminated entries were summed for each replicate (two replicates per treatment). Then, iBAQ values were normalized by dividing the individual iBAQs from a given replicate by the respective summed iBAQ, yielding a relative iBAQ (riBAQ). For each protein, riBAQ values were averaged from both replicates to obtain the mean riBAQ for that protein under the respective treatment (NLS-BirA or EGR2-BirA).

To get a list of enriched proteins upon EGR2-BirA treatment, each protein mean riBAQ value for EGR2-BirA was divided by the mean riBAQ value for the NLS-BirA (negative control). Proteins that exhibited a ratio inferior at four were considered as not enriched. In total, we detected 169 proteins, which were enriched in the EGR2-BirA compared to the NLS-BirA. An overview of the GO term enrichment analysis (100 enriched clusters) is given in Figure 5.27.

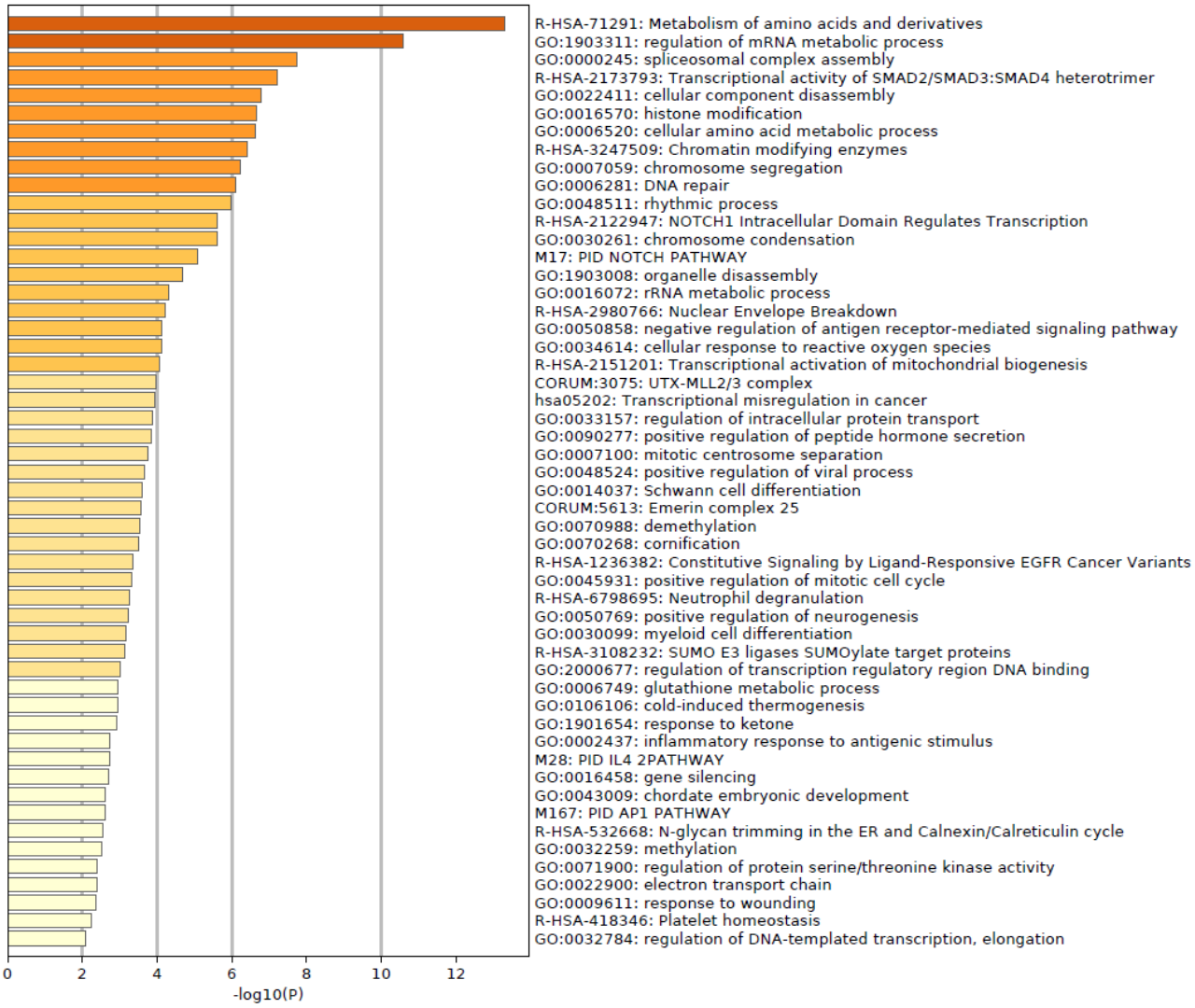

Figure 5.27 GO-terms for enriched genes in EGR2-BirA samples

Gene ontology (GO) analysis of EGR2-BirA enriched genes in iDCs-d3 analysed by Metascape. The significance of the enrichment of a particular term is depicted with the log10 of the p-value and bars are colored by p-values.

Functional annotation in Figure 5.27 revealed that enriched proteins in EGR2-BirA were involved in distinct cellular processes including demethylation, which is consistent with a potential association between EGR2 and DNA demethylation machinery.

Due to the extensive list of enriched EGR2-BirA-specific proteins, we further filtered some genes based on the information obtained from Metascape, thus allowing a better visualization of functional protein association networks (using the STRING platform). Only genes associated with chromatin remodeling, demethylation, immune system as well as known co-factors of EGR2 (e.g. NAB2 and HCFC1) are highlighted in Figure 5.28.

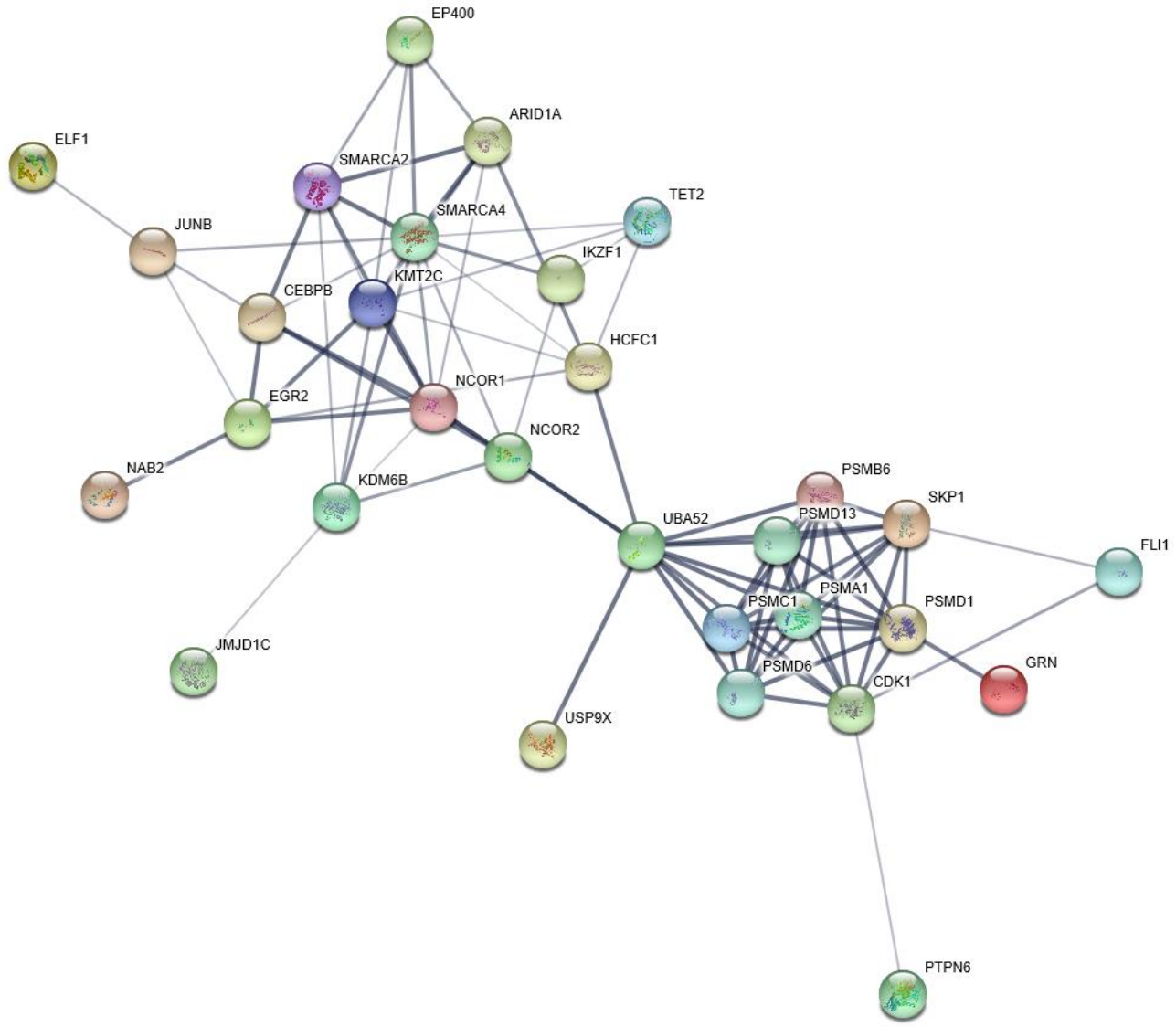

Figure 5.28 EGR2 protein-protein interactions in iDCs-d3

STRING analysis illustrating the functional protein association network. The given network summarizes predicted associations for selected proteins enriched in the EGR2-BirA samples. The network nodes are the proteins, the edges represent the protein-protein associations and the line thickness indicates the strenght of data support. Seven types of evidence are used in predicting these associations (textmining, experiments, databases, co-expression, neighborhood, gene fusion and co-occurrence).

As illustrated in the functional protein association network (Figure 5.28), chromatin remodeling enzymes (e.g EP400, SMARCA2/4 and ARID1A), histone demethylases (e.g JMJD1C and KDM6B) and also TET2 are likely associated with EGR2 in iDCs.

Taken together, identification of TET2 protein in the EGR2-BioID experiments reinforces our hypothesis about the EGR2 critical role in TET2 recruitment during MO differentiation. Nevertheless, to facilitate further statistical analyses and to increase the reproducibility of our data, additional donors must be included in future BioID experiments.

5.5.1.2 Co-Immunoprecipitation (CoIP)

Since the BioID approach only allows to study the vicinal proteins within close proximity (estimated to be within 10 to 20 nm)[178] to the fusion protein, the actual direct/physical protein interactions might be poorly predicted using this method. Therefore, we performed co-immunoprecipitations (CoIPs) in iDC as well as in a myeloid cell line (THP-1) to confirm the presence of a direct interaction between EGR2 and TET2. As NAB2 TF is a known co-factor of EGR2[193], it was also included in the CoIP experiments. Due to the absence of proper antibodies against EGR2 or NAB2, CoIPs were performed using cells transfected with mRNAs encoding the FLAG-tagged proteins. Briefly, these *in vitro* transcribed (ivt) mRNAs were generated as described in sections 4.3.5-4.3.7 and 4.5.3, and transiently transfected by electroporation into iDCs (day 3) or THP1 cell line as described before (section 4.1.5). To check protein expression levels after transfection with the respective constructs, iDCs were harvested at different time points (1h, 3 h, 6 h and 8 h) and the whole cell lysates were used for western blot (see section 4.6.3). The results are given in Figure 5.29.

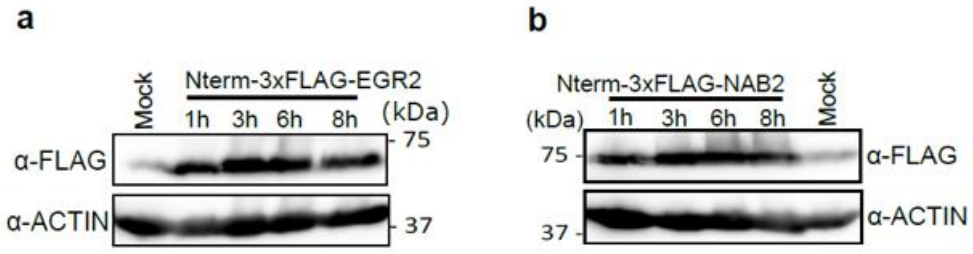

Figure 5.29 Protein kinetic for the EGR2 and NAB2 fusion proteins in iDC-d3

The immunoblots depict the protein expression levels detected for **a)** Nterm-3xFLAG-EGR2 and **b)** Nterm-3xFLAG-NAB2 fusion proteins at the indicated time points (1 h, 3 h, 6 h and 8 h) after cell transfection. Blots were stained with a α-FLAG antibody and a α-ACTIN antibody was used as the protein loading control. A mock sample was also included as control.

As presented in Figure 5.29, the EGR2 and NAB2 constructs were overexpressed in iDCs over the time course (1 h, 3 h, 6 h and 8 h) after transfection as detected by the anti-FLAG (α-FLAG) antibody. Based on these observations, cells were harvested three hours after transfection to perform the CoIP experiments.

CoIPs and reciprocal CoIPs were performed using the Dynabeads® Antibody Coupling Kit to crosslink the α-TET2, α-FLAG and α-IgG antibodies to the magnetic beads before co-immunoprecipitation of the protein complexes. The detailed protocol can be found in section 4.6.5 and the results are depicted in Figure 5.30.

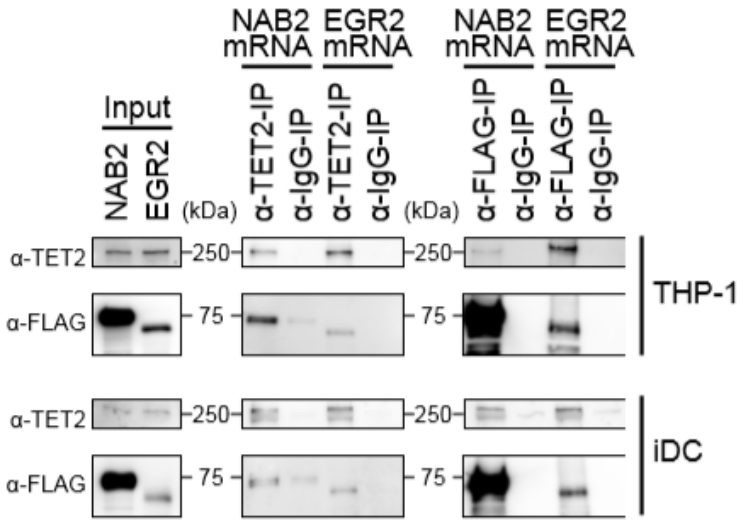

Figure 5.30 TET2 interacts with EGR2 and NAB2

CoIPs and reciprocal CoIPs of EGR2 or NAB2 (expressed as FLAG-tagged proteins via mRNA electroporation) and TET2 in THP-1 and iDC. Interactions were assessed by IP using a α-FLAG antibody followed by immunoblot analysis with a α-TET2 antibody (FLAG-IP) and conversely for the reciprocal CoIP. Moreover, input samples (cell lysate before CoIP) and α-IgG-IP were used as controls.

As shown in Figure 5.30, in both CoIP and reciprocal CoIP, TET2 -EGR2 and TET2-NAB2 interactions were detected, emphasizing the involvement of EGR2 as well as NAB2 in the DNA demethylation process in iDC. Of note, the weak signals (low band intensity) on the blots are primarily due to the transient nature of these interactions.

5.6 DNA methylation-spikes across different cell types

In line with the results in Figure 5.23 & Figure 5.24, we observed the same methylation footprint of EGR2 at bound regions (“TF peak” DMR group) as well as regions that show no detectable EGR2 binding (“open” DMR group) nor accessible chromatin (ATAC) (“no peak” DMR group).

As exemplified in Figure 5.24, the methylation footprint at EGR2 binding sites was associated with a delay on DNA demethylation upon EGR2 depletion. Therefore, our data suggest that at certain loci, methylated CpGs can be protected from demethylation by transiently bound EGR2.

Considering the importance of this methylation footprint of EGR2 for DNA demethylation processes around its binding sites, in a next step we checked if this phenomenon could also be found in other cell types and represent a feature of other factors. For this purpose, we extended our analyses and compared 5mC- spikes (which may indicate demethylation protection) in MO and iDC (Figure 5.31) as well as in other cell types like MAC, neutrophils, B and T cells. As an example of factor to support our observations, we focused on the pioneer CEBP (using published data).

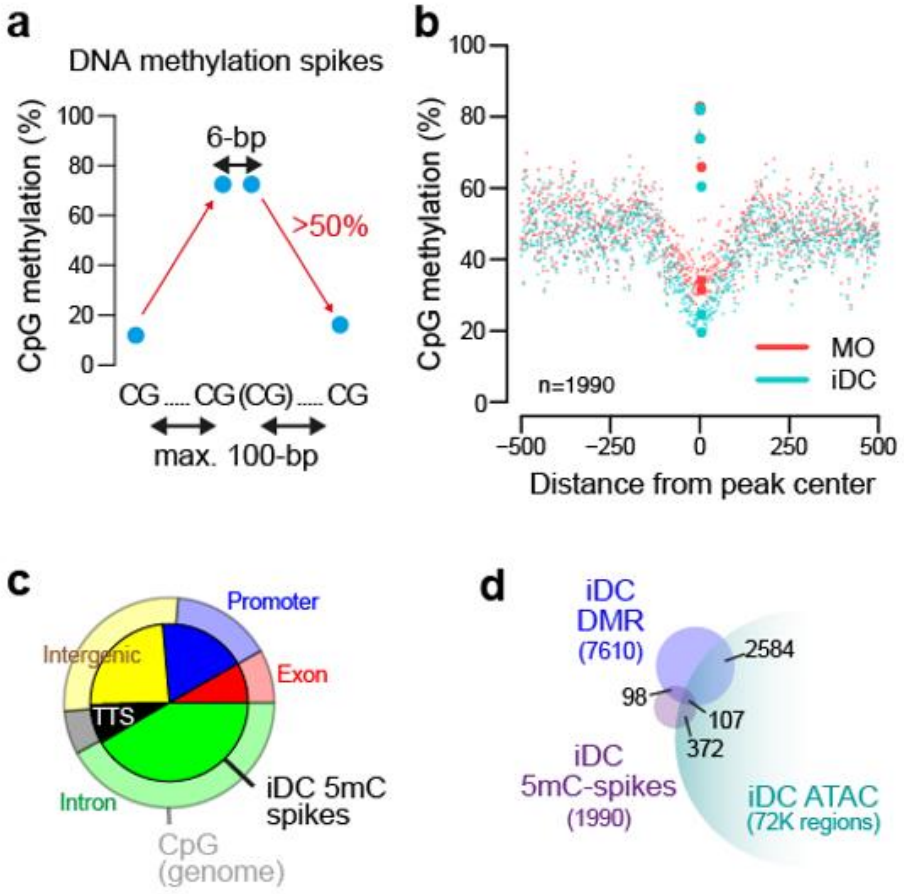

Figure 5.31 DNA methylation-spikes in MO and iDCs

a Strategy for identifying small (max. 6-bp wide) spikes of DNA methylation. **b** Genomic distance distribution of averaged DNA methylation ratios in MO and iDCs across iDC 5mC- spikes. Values of central CpGs are indicated by larger dots and the number of 5mC spikes is given in the lower-left corner. **c** Pie chart illustrating the genomic location distribution of iDC 5mC spikes (inner circle) relative to CpGs across the entire genome. **d** Venn diagram depicting the overlap between 5mC spikes, DMR and accessible regions (ATAC peaks) in iDC.

Figure 5.31a shows a strategy to identify small spikes of DNA methylation in MO and iDCs as indicated by a sharp increase in DNA methylation ratios greater than 50% followed by a similar reduction in CpG DNA methylation. As shown in Figure 5.31b, regions around iDC-spikes (larger dots) were less methylated in iDC when compared to MO. These data are consistent with our findings for EGR2, which assume that at certain positions, EGR2 binding sites overlap with methylated CpGs while the surrounding CpGs are demethylated.

As illustrated in Figure 5.31c, iDC 5mC- spikes (DNA methylation-spikes) were enriched in promoter and TTS regions. Like DMR, only a limited number (372) of 5mC- spikes in iDC overlapped with iDC ATAC peaks. Interestingly, only 98 DMR overlapped with iDC 5mC- spikes and 107 DMR overlapped with both 5mC- spikes and ATAC peaks in iDCs.

The Venn diagram in Figure 5.32a (left panel) illustrates the overlap between MO and iDC 5mC-spikes.

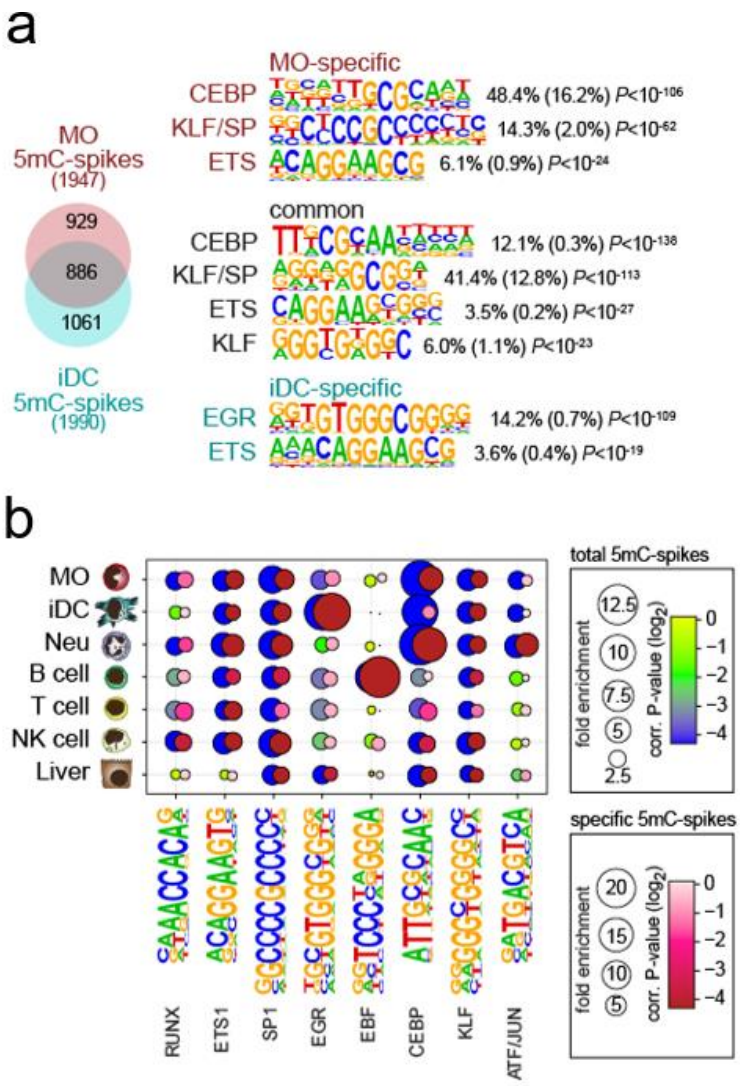

Figure 5.32 DNA methylation-spikes are cell type-specific and enriched for TF motifs

a Venn diagram indicating the number of common and cell type-specific 5mC- spikes between MO and iDC. *De novo*–derived motifs for each 5mC-spike set are given along with the significance of motif enrichment (hypergeometric test) and the fraction of motifs in peaks (background values are in parenthesis). Top motifs corresponding to known factor families are shown for each set. **b** Balloon plot depicting the motif enrichment of selected motifs across all 5mC- spikes identified in each cell type (blue-green-yellow coloring according to corrected enrichment p-value) or cell type-specific 5mC- spikes (red-pink-white coloring according to corrected enrichment p-value). The balloon size represents the fold-enrichment and the coloring indicates the corrected P-value (Hypergeometric test, Benjamini-Hochberg multiple testing correction) of the motif occurrence in 5mC- spikes.

As expected, the motif signature across MO-, iDC-specific or common 5mC- spikes (Figure 5.32, right panel) revealed that iDC-specific 5mC- spikes were enriched for the EGR consensus motif whereas common and MO-specific 5mC- spikes were highly enriched for CEBP and KLF consensus motifs. These data suggested that DNA methylation-spikes are cell type-specific and not restricted to EGR2. In an effort to generalize our findings, published WGBS data sets from additional human cell types (including neutrophils, B cells, T cells, NK cells and hepatocytes) were used and the main results obtained from computational analyses are shown in Figure 5.32b.

The balloon plot in Figure 5.32b shows the TF motif enrichment across the total and specific 5mC-spikes in different cell types. The correlation between the total *(*blue-green-yellow gradient color) or specific (red-pink-white gradient color) 5mC- spikes in different cell types with their motif enrichment is also provided. Accordingly, we detected a strong cell type-specific enrichment of particular TF motifs (larger filled red balloons), including CEBP in MO and neutrophils, EGR2 in iDC, and EBF in B cells.
To correlate the presence of 5mC- spikes in CEBP consensus motif with binding (ChIP-seq, published data) of CEBP TFs (CEBPα or β) across different human cell types (MO, MAC and liver) pie charts were generated as shown in Figure 5.33a.

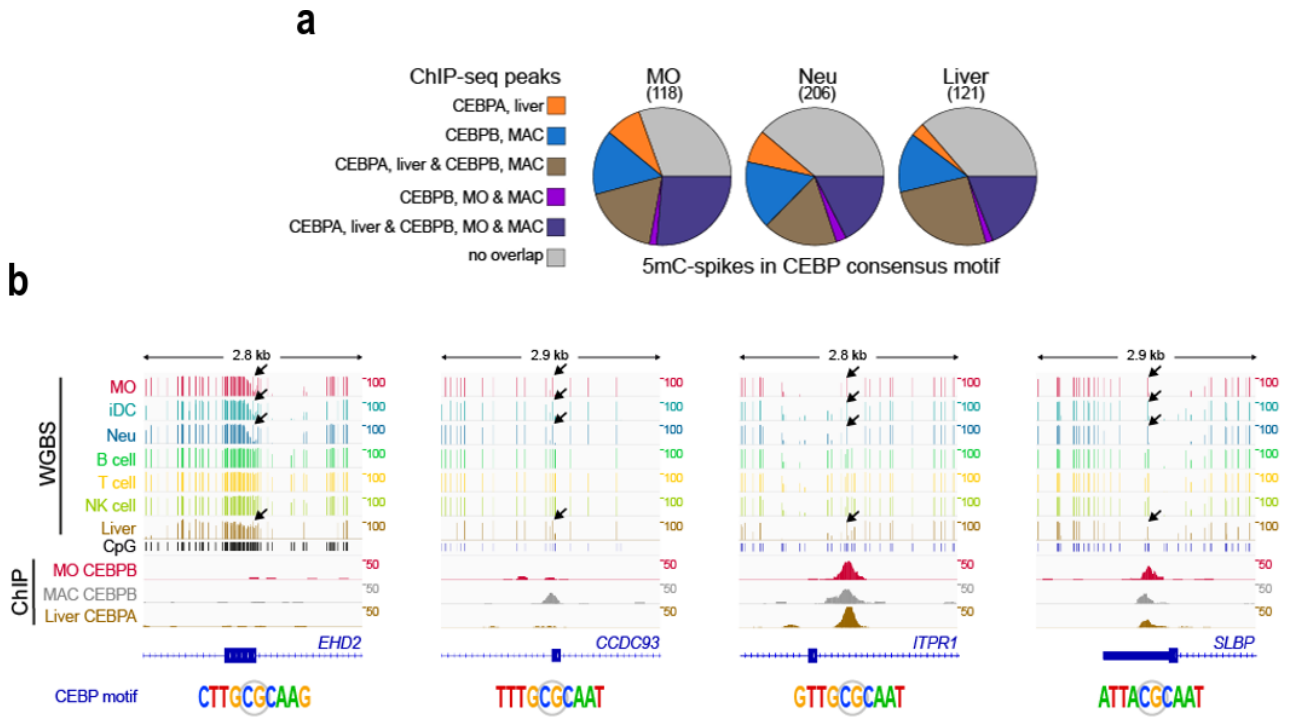

Figure 5.33 Typical 5mC- spikes in CEBP consensus motif and TF binding

a Pie charts illustrating the overlap of 5mC- spikes with the indicated CEBP factor ChIP-seq peaks in MO, MAC and liver. **b** IGV genome browser tracks for the indicated WGBS and ChIP data sets at example regions around 5mC- spikes overlapping the CEBP consensus sequence.

As illustrated in these pie charts, a significant overlap of ChIP-seq peaks for CEBPα or β in human liver, MO or MAC was observed, proposing that binding of CEBP TF may protect 5mC- spikes from DNA demethylation.
Figure 5.33b displays IGV genome browser tracks for the specified WGBS and ChIP data sets at four loci overlapping CEBP consensus motifs. As indicated by the black arrows, MO, iDC, neutrophil and liver cell types exhibited a typical 5mC-spike similar to the methylation footprint detected at EGR2 binding sites (see Figure 5.23 & Figure 5.24).

In a similar way to EGR2, we also observed 5mC- spikes at CEBP motifs with no detectable CEBP binding (ChIP), suggesting that a transient interaction of CEBP with its binding sites may also be sufficient to mediate the recruitment of DNA demethylation machinery to its binding sites.

6 Discussion & Perspectives

DNA methylation turnover is a critical process during normal hematopoiesis and its dysregulation is associated with numerous malignancies[194,195]. In fact, TET2, the most relevant hydroxylase initiating this process in hematopoietic cells[196–198] is often mutated in myeloid[60,199–201] and lymphoid[202,203] malignancies, although the latter have been reported to a lesser extent.

To better understand, how mutations in the DNA demethylation machinery might affect normal hematopoiesis leading to malignant transformation, it is fundamental to gain deeper insight into the mechanisms involved in 5mC removal. This is why this thesis was mainly focused on the DNA demethylation process during post-mitotic MO differentiation. It combines high-throughput-based methods, biochemical, computational and functional analysis to identify mechanisms and TFs involved in DNA demethylation recruitment, which are currently poorly understood in these cells.

Previous studies described DNA demethylation during hematopoietic cell differentiation as an active process[13], thereby involving the enzymatic activity of TET enzymes to initiate DNA demethylation. To study the active DNA demethylation process, an *in vitro* system to differentiate peripheral human MO into iDC upon stimulation with IL-4 and GM-CSF cytokines has been repeatedly used by our group and others[31,32,204–206]. One of the greatest advantages of this model is that during differentiation[32] or even in response to infection, cells do not proliferate[207], which allows to investigate molecular mechanisms occurring in the absence of DNA replication, including active demethylation. Therefore, in this thesis this *in vitro* system was used to address our key research questions.

In vitro studies using this model have contributed to identify dynamic changes in DNA methylation during post-mitotic MO differentiation, which are clearly associated with extensive losses of DNA methylation and only very few *de novo* methylation events[31,32,206,208], thus reflecting our observations at iDC- and MO-specific DMRs, respectively (Figure 5.2). Although some studies were also focused on transcriptional changes during differentiation and how they correlate with DNA methylation events[206,209], they failed to provide additional epigenetic landscapes analyses and also to compare genome-wide DNA methylation changes between MO and iDCs. To bypass this, we analysed 5hmC and ATAC-seq peaks distribution in MO and iDCs to capture active demethylation and chromatin accessibility changes, respectively. In addition, genome-wide binding patterns for certain TFs were also assessed by ChIP-seq, improving our understanding of how epigenetic changes correlate between each other to drive the normal differentiation of MO into iDC.

6.1 Genome-wide active DNA demethylation during MO differentiation

Dynamic changes in chromatin structure and DNA methylation are critical during cell differentiation and lineage specification[210]. The switch/sucrose non-fermentable (SWI/SNF) chromatin remodeling complex is reported to be fundamental in these processes[211–213], including its interplay with TET enzymes. For instance, Sepulveda and colleagues[214] stated that Tet-mediated DNA demethylation is necessary for the activity of SWI/SNF-containing complexes that promotes the expression of Sp7 master gene during osteoblast differentiation.

In the present work, we verified a progressive enrichment in 5hmC (active demethylation) and chromatin accessibility (ATAC) (Figure 5.3) across our iDC-specific DMRs with some DMRs being demethylated even before open chromatin or TF binding detection (Figure 5.11a). Although these data indicate that at particular DMRs active demethylation can precede other epigenetic processes, the 5hmC distribution across DMRs was more or less pronounced according to the time point of iDC differentiation. Indeed, the distribution of 5hmC was higher upon 18h of culture compared to iDCs-d7 that exhibited a decreased but broader distribution of 5hmC (Figure 5.3). These data may suggest that iDCs at day 7 of differentiation might represent a mixture of regions where most of the demethylation is almost or even complete (5mC → 5C) and regions that are later demethylated (5hmC).

It was previously reported that demethylation might be stalled at 5hmC[215]. In compliance with this study, Xiong and colleagues[216] demonstrated that in mouse ESCs, SALL4A is recruited by TET1 to stabilize the chromatin association of TET2, and thus facilitate the further oxidation of 5hmC. This model, however, cannot be extrapolated to our system because TET1 expression is undetectable in MO or MO-derived cells[32] and SALL4 is not associated with our iDC-specific DMRs. Although TET2 is capable to successively oxidize 5hmC to 5fC and 5caC, it is still unclear whether another enzyme or TF is required to promote TET2-mediated oxidation of 5hmC in MO-derived cells.

A comprehensive analysis of DNA methylation ratios across all three DMR groups (“no peak”, “open” and “TF peak”) during iDC differentiation, revealed distinct DNA methylation turnover rates in these groups (Figure 5.11b). It seemed like the DNA demethylation was delayed at “no peak” and “open” groups compared to “TF peak” group. These data may suggest a low TET2 processivity at these regions, which in turn, might be associated with a repressed (“no peak” group, Figure 5.11b) or reduced (“open” group, Figure 5.11b) chromatin accessibility[217]. In addition, the “TF peak” group also showed more motif co-occurrences compared to “no peak” and “open” DMR groups (Figure 5.12). These findings might associate with a broader demethylation and an increased chromatin accessibility at the “TF peak” DMR group that create more opportunities for TF cooperativity in the course of iDC differentiation.

In line with previous observations[31] only a fraction of iDC DMR associated genes overlapped with genes that are significantly regulated during MO differentiation, indicating that for many DMRs active demethylation processes occur independently of transcriptional changes. In addition, we also observed that at some DMRs, the active demethylation takes place even before chromatin remodeling and/or TF binding detection. These regions are presumably required for later processes like cell maturation to render them functional. These findings were partially observed by Vento-Tormo and colleagues[206] work, who verified that for a subset of genes, the DNA demethylation during MO differentiation (using the same *in vitro* system) precedes their upregulation upon cell maturation.

As observed for active demethylation (5hmC), the post-mitotic MO differentiation is also characterized by an increase in chromatin accessibility with different open chromatin dynamics across iDC-DMRs (Figure 5.11). Some DMRs can get early, later, transient or even not accessible during differentiation, suggesting that chromatin structure is not static during MO differentiation but a well-controlled process that directly correlates with stable TF binding detection (Figure 5.11). As previously described in other systems[211–213], our findings indicate that chromatin remodeling at DMRs is triggered during MO differentiation and thereby its dysregulation may impair the normal differentiation process.

Another important unsolved question was whether chromatin remodeling and active DNA demethylation are interdependent processes, or instead, both processes can occur independently of each other. To address that, we combined 5hmC, ATAC-seq and ChIP-seq data across iDC-DMRs and verified (as also mentioned above) that some DMRs were actively demethylated even before chromatin remodeling or TF binding detection. These observations confirmed that at particular DMRs chromatin remodeling and active demethylation can be independent processes. Although it was recently reported that Klf4-mediated enhancer active demethylation precedes open chromatin detection during reprogramming in mouse[98], our work is the first to describe an independent relationship between chromatin remodeling and active DNA demethylation (at some DMRs) during *in vitro* differentiation of human MO. These findings unveiled a relevant hallmark of epigenetic gene regulation during post-mitotic MO differentiation.

Previous studies confirmed the role of TET2 in active demethylation during MO differentiation into iDC (using the same *in vitro* system), however, they only focused on DNA methylation changes, which is insufficient to understand other potential functions of TET2 during MO differentiation. To overcome this, we used high-throughput sequencing technologies like RNA-seq and ATAC-seq to also study the role of TET2 in gene expression and chromatin accessibility, respectively.

As mentioned before, TET2 is the main responsible enzyme for the iterative oxidation of 5mC to 5hmC in our system, while TET1 and TET3 are primarily involved in DNA demethylation in embryonic stem cells/adult brain[29,218] and in the zygote[219], respectively. Besides its essential role in MO-derived cells[13,32,206], TET2 is also required for active demethylation processes in proliferating hematopoietic progenitor cells[98,220,221].

The marked effects of TET2 depletion on different hematopoietic cell types have been extensively reported[221–223]. Together, TET2 KD was associated with DNA hypermethylation (gain of 5mC) and reduced hydroxymethylation (loss of 5hmC) as well as with aberrant HSC self-renewal (in mice)[224]. In addition, silencing of TET2 in human hematopoietic progenitor cells was found to disturb the granulocytic and erythroid differentiation favoring the monocytic development[222].

In the present work, silencing of TET2 was associated with very mild effects on gene expression (only few genes were differentially expressed between control and siTET2 samples), cell morphology and chromatin structure (slight effects upon TET2 depletion) in iDC. These effects are presumably due to the expression of TET2 already in MO that resulted in a delayed effect of siRNA treatment on protein expression (Figure 5.4c) with many sites initiating demethylation at earlier stages of MO development (Figure 5.5). Similarly, the presence of low levels of TET3[32] might compensate for the reduced levels of TET2 and likewise attenuate the effect of TET2 depletion in this system. Actually, Montagner and colleagues[225] proposed that absence of TET2 during mast cell differentiation (cells obtained from bone marrow of *Tet2*$^{-/-}$ mice) could be compensated by the activity of other TET proteins.

Nevertheless, our results were reproducible among donors and genes downregulated upon TET2-silencing were normally induced during iDC differentiation (Figure 5.7). Furthermore, the impact of TET2-silencing on DNA methylation was consistent with its critical role in DNA demethylation during *in vitro* MO differentiation (as also reported by Vento-Tormo and colleagues[206]). DNA demethylation was impaired at many regions (Figure 5.5) and in compliance with previous findings[31], demethylation did not occur simultaneously at all CpGs positions during differentiation (while some CpGs were early others were late demethylated).

Interestingly, some iDC DMRs lost chromatin accessibility upon TET2 depletion, suggesting that TET2 may be involved in both DNA demethylation and chromatin remodeling processes. In line with this, intragenic DMRs (upon TET2 KD) showed a delay in DNA demethylation and reduction in open chromatin (amplicons CABP4, TMIGD3 and MAN1C1 in Figure 5.5), which were correlated with downregulation of their associated genes (Figure 5.7b). These data indicate that at particular iDC-DMRs, TET2-mediated DNA demethylation and chromatin remodeling processes are coupled with changes in the expression of genes[206,214] typically induced during iDC differentiation.

Overall, our findings suggested that active DNA demethylation (in cooperation with other epigenetic processes like chromatin remodeling) plays a critical role in the activation of previously DNA methylated cis-regulatory modules. For instance, we verified a local delay in DNA demethylation (upon TET2 depletion) at some DMRs overlapping promoter regions (e.g. EpiTYPER amplicon CCL13, Figure 5.5a). In line with previous work performed in the scope of hematopoietic differentiation[226], our data indicate that promoter DNA demethylation is an essential process during post-mitotic MO differentiation.

Since the discovery of TET proteins as well as their involvement in DNA demethylation processes[19,36,38], TET enzymes as well as their intermediate oxidative products have been largely studied and characterized[19,42,66]. However, the mechanisms controlling TET2 recruitment to individual target loci in hematopoietic cells is poorly understood.

Unlike TET1 and TET3 full-length isoforms, TET2 lacks the CXXC DNA-binding domain and thus is not capable to bind methylated CpGs[24] and to initiate the DNA demethylation process. To overcome this limitation, we proposed (as also suggested by several groups[97,98,104–106]) that TET2 might interact with other TF to induce local DNA demethylation. To address this possibility, we initially divided iDC-DMRs into progressive and *de novo* demethylated regions according to the DNA methylation ratios in MO and investigated TF motif signatures across these regions to identify potential candidates for TET2 recruitment (see section 5.3 & Figure 5.9).

As expected, both region sets (progressive and *de novo* demethylated) displayed motif signatures for TFs previously associated with iDC differentiation like PU.1, IRF4 and STAT-family factors[122,167,172]. However, certain TFs were more enriched in regions where DNA demethylation already started at earlier stages of MO development (progressive demethylated regions), such as PU.1, EIRE composite motif and AP1-family factors. These data are in agreement with the expression of these TFs in MO[227–229]. On the other hand, *de novo* demethylated regions (highly methylated regions in MO) were more enriched for TF induced upon culture, such as STAT and EGR[206,230].

Considering that, EGR2 and IRF4 are induced and transcriptionally activated in iDCs (Figure 5.10) they represented valid candidates to trigger *de novo* demethylation processes and thereby were used as target genes for genetic functional analyses.

6.2 Effects of key regulator depletion on MO differentiation

IRF4 and EGR2 represented valid candidates to mediate *de novo* DNA demethylation events during *in vitro* MO differentiation into iDC. Therefore, we assessed the consequences of TF knock-down on MO differentiation, including associated epigenetic processes like DNA demethylation and chromatin remodeling.

IRF4 (also known as multiple myeloma oncogene-1 (MUM1)) belongs to the IRF family of TFs[231,232]. The IRF family is composed for nine members (IRF1–IRF9) with regulatory functions in virus-related defense, innate and adaptive immunity, cell development and oncogenesis[233,234].

Like in other IRF family members, the protein structure of IRF4 contains three domains (a DNA-binding, an IRF-association domain 1 (IAD1) and an auto-inhibitory domain) and a nuclear-localization signal[235]. The IAD1 domain may facilitates homo- and heterodimerization between IRF4 and other TFs, including PU.1[236].

IRF4 is predominantly expressed in T cells, B cells, MAC and DCs and is critical for lymphoid and myeloid lineage development and function[235]. For instance, studies performed in murine models

showed that IRF4 is essential for DC[190], myeloid-derived suppressor cells (MDSC)[237] and MAC differentiation[238]. Furthermore, IRF4 is necessary during immune responses for lymphocyte activation and production of immunoglobulin-secreting plasma cells. Its dysregulation is also associated with many lymphoid malignancies[239,240].

Despite previous reports of the pivotal role of IRF4 in DC development and function[190–192], its functions in iDC-associated epigenetic processes like DNA (de)methylation and chromatin remodeling were largely unknown. In this work we addressed the functional importance of IRF4 in these processes during *in vitro* human MO differentiation into iDC.

It is well-established that IRF4 is induced by LPS, IL-4, and CD40[241–243], which is consistent with the observed IRF4 upregulation upon IL4/GM-CSF driven MO differentiation (Figure 5.15c). In turn, its depletion resulted in a marked reduction on IRF4 protein levels at earlier stages of differentiation (Figure 5.15c), that was accompanied by pronounced effects on cell morphology (Figure 5.15b) without affecting cell survival (Figure 5.15a). IRF4-depleted human MO acquired a MAC-like morphology (Figure 5.15b) despite culture in the presence of IL4/GM-CSF, which was equivalent to observations made in mice[244].

Supporting these findings, IRF4-deficient cells exhibited profound changes on transcriptional level (Figure 5.16), including loss of DC and gain of MAC gene expression signatures. Actually, silencing of IRF4 prevented the upregulation of important DC markers like CD1a, CD1b, CD1c and CD80[245,246] as well as many HLA genes with important functions in the regulation of the immune system[247,248]. In contrast, genes encoding integrins (e.g. ITGB2, ITGB5) and lysozymes (e.g. LYZ), essential for phagocytosis and inflammatory responses in MAC[249–251] were upregulated in siIRF4-deficient cells.

Interestingly, it has been stated that interactions of IRF members with transcriptional partners like IRF4 (or IRF8) and the co-acting transcriptional regulator PU.1, allow IRFs to control distinct transcriptional programs[252]. In fact, the majority of IRF4 binding sites in DC are composite ETS:IRF (EIRE) elements that require PU.1 for IRF4 to bind DNA[253].

The transcription factor PU.1 (also known as Spi-1) belongs to the erythroblast transformation specific (ETS) family and in the same manner as most ETS proteins, PU.1 can bind alone to a specific core motif (A/GGAA) [254] or in a complex with IRF4 or IRF8. Hence, PU.1, Spi-B and IRF4, IRF8 can cooperatively assemble on composite DNA elements (ETS:IRF) and regulate transcription in the immune system[253]. For instance, the heterodimer composed by PU.1 and IRF4 has been implicated in dendritic cell gene expression[255] as well as in leukemia suppression[256].

Therefore, during MO to iDC differentiation the transcription regulation of genes containing ETS:IRF (EIRE) composite elements (that require IRF4 binding) was likely disrupted by silencing of IRF4, contributing to the marked changes on DC-specific transcriptional programs.

Taken together, our observations are in line with previous reports that highlight the IRF4 pivotal role in DC development and function[138,190–192,244].

IRF4 along with other TFs (BATF, a FOS-like and AP-1 family TF) was suggested to mediate chromatin accessibility during T cell differentiation[257]. However, a function of IRF4 in chromatin structure and how this event may correlate with active DNA demethylation during MO differentiation remain unclear.

To better understand the role of IRF4 not only in chromatin remodeling but also in DNA methylation, we investigated the functional consequences of IRF4 depletion on these processes using ATAC-seq and EpiTYPER techniques, respectively. Data analysis revealed a clear impact of IRF4 depletion on chromatin accessibility (ATAC-seq) (Figure 5.17) in IL4/GM-CSF-MO-derived iDC with loss of IRF4 containing motifs (see Figure 5.17a). In addition to its critical function during iDC differentiation, our data highlight a previously unreported potential role of IRF4 in chromatin remodeling processes, which may contribute to get better profiles of epigenetic landscapes in the course of MO differentiation.

As a secondary impact of IRF4 depletion on chromatin accessibility, we found that more accessible sites were strongly enriched for CEBP, AP1, EGR2 and E-box motif, which are not related to IRF4 binding but were previously identified as being part of a MAC-enhancer signature[227]. These data are then consistent with the MAC-like morphology and gene expression signatures observed upon IRF4-silencing.

Focusing on DNA methylation changes during MO differentiation, sites that loose accessibility in IRF4-depleted cells (and bound by IRF4) are targets of the demethylation machinery during normal MO differentiation (Figure 5.22b, right panel). However, most sites affected by the IRF4 KD already initiated DNA demethylation during earlier stages of MO development (Figure 5.17b), suggesting that IRF4 could be involved in active DNA demethylation processes at progressive and not at *de novo* demethylated regions.

Notably, regions normally bound by IRF4 and even by PU.1 (resembling their binding at ETS:IRF (EIRE) composite elements) were still demethylated despite their failure to gain accessibility (upon IRF4 KD) during MO differentiation (Figure 5.21). These data suggest that IRF4 mediates accessibility changes regardless of active demethylation, and that IRF4 is likely not involved in the recruitment of TET2 to its individual or composite binding sites (ETS:IRF (EIRE)).

As previously mentioned, another valid candidate for TET2 recruitment during IL-4/GM-CSF-driven MO differentiation was EGR2.

EGR2 along with the other three members, EGR 1, 3 and 4 (EGR1/Krox24, EGR3 and EGR4) is part of the family of EGR genes. Their expression is induced in response to mitogens[258], differentiation, apoptotic signals and tissue injury, implicating them in multiple biological processes[259].

EGR proteins are TFs composed of three zinc fingers that recognize a G:C-rich DNA motifs in the promoters of target genes[258,260,261]. Their transactivation activity is modulated by coactivators like the host cell factor C1 (HCF-1)[262] and corepressors such as NGFI-A Binding Protein 1 (NAB1)[263] and NGFI-A-binding protein 2 (NAB2)[264].

EGR genes are expressed in many cell types, including for example cells within the nervous system[265,266] and immune system[267–269]. In particular, EGR2, is described as an essential transcriptional regulator of immune cells and can act either as repressor or activator of gene expression[259,270,271]. During monocytic maturation (in a murine model for MAC differentiation) EGR2 was found to repress neutrophil-specific genes[229,272] through its association with the NAB2 corepressor[264]. NAB2 was also reported as an important co-factor for the recruitment of the mediator subunit INST13 to EGR1/2-bound enhancer elements in human MO/MAC[273]. EGR2 is also involved in upregulation of genes (Tbx21 and Notch1) relevant in T-cell activation[274] as well as in upregulation of major factors driving T cell anergy[275]. Therefore, its function in hematopoietic cells is complex and requires further elucidation.

EGR2 has been previously involved in many hematopoietic differentiation processes, such as T cell[274,276], B cell[276], neutrophil[272], natural killer cell[277] and human MO-derived macrophages[227] differentiation. Yet, a role in IL4/GM-CSF-driven MO differentiation has not been demonstrated.

This is why this thesis combines different high-throughput-based methods to understand the role of EGR2 not only in MO differentiation but also in associated epigenetic processes, including active DNA demethylation.

In line with a role of EGR2 in inducing active DNA demethylation events, the low levels of EGR1 observed in MO (Figure 5.10) could indicate that this factor initiates 5mC oxidation in MO and its function is later replaced by EGR2. At present, however, we are unable to confirm this theory and thereby further analyses were only focused on EGR2.

In this thesis, we used our well-established approach to deplete EGR2 in MO and to investigate the functional consequences of EGR2 depletion on iDC differentiation and associated epigenetic processes. Unexpectedly, in some of our experiments EGR2 protein expression was already detected in MO (Figure 5.18c), which may explain the later effects of EGR2 depletion on protein level (only after 66h and 7d of culture). Supporting these observations, RNA-seq data from the BLUEPRINT epigenome project also revealed medium expression levels of EGR2 in classical MO ($CD14^+CD16^-$). In contrast, another group showed no signs of EGR2 protein expression in MO[227]. These inconsistent data are presumably related with different cell purification methods, culture or individual donors, which might induce the premature expression of EGR2 in MO.

EGR2-deficient cells showed marked morphological changes and a significant reduction in cell viability compared with cells treated with a siRNA control (Figure 5.18a,b), suggesting that EGR2 has a critical role in the IL4/GM-CSF-driven MO differentiation, which has not been previously described. Therefore, our findings contribute to a better characterization of *in vitro* iDC (obtained from human MO).

Corroborating the negative effects of EGR2 depletion on iDC differentiation, profound changes on DC transcriptional programs were detected (Figure 5.19). Genes downregulated upon EGR2 KD were largely enriched for functional categories involving actin cytoskeleton organization which is known to

unfavourably affect DC functions like T cell adhesion and activation[278–280], phagocytosis[281], or migration behaviour[282,283]. Furthermore, EGR2-silencing resulted in downregulation of important DC markers like CD1a, CD86 and CD80[162,174] as well as other genes reported as essential for IL4/GM-CSF-driven MO differentiation, such as IRF4[244] and NCOR2[205] (also identified in our EGR2 BioID experiments in iDC, Figure 5.28).

Conversely, genes upregulated upon EGR2 depletion, such as APOE[284] and MYD88[285] were primarily enriched for functional categories involving adaptive immune responses. Another functional category was related to processes mediated by lymphocytes, which is in compliance with detection of a T-cell gene signature (Figure 5.19) in the differential gene expression analysis. Despite some potential contamination of our cultures by T cells, we identified different myeloid specific markers in the transcriptional analysis like S100A8/S100A9[286].

Besides the pronounced effects of EGR2-silencing on iDC differentiation and gene expression, its silencing also impaired chromatin remodeling (ATAC-seq) (Figure 5.20) and DNA methylation (Figure 5.21), suggesting that EGR2 may be involved in both chromatin remodeling and active DNA demethylation processes.

To better understand the role of EGR2 in both processes, we focused our analyses on DNA methylation changes during MO differentiation and verified that the most of less accessible sites affected by EGR2 depletion (and bound by EGR2) were *de novo* demethylated during MO differentiation (sites did not show marked signs of previous demethylation in MO) (Figure 5.20b). Similarly, our local DNA methylation analyses as determined by EpiTYPER, showed that EGR2 depletion impaired DNA demethylation at many DMRs independently of their chromatin accessibility state and EGR2 binding detection. These observations suggest that EGR2 might be able to initiate the DNA demethylation at both accessible and inaccessible DMRs. Together, these results support the novelty and importance of our findings, suggesting an epigenetic pioneering role of EGR2 in triggering *de novo* demethylation events during post-mitotic MO differentiation.

6.3 Epigenetic pioneering role of EGR2 in triggering de novo demethylation processes

As previously mentioned, IRF4 is presumably associated with progressive demethylation events, while EGR2 might be involved in *de novo* demethylation events. In contrast to the very little effects of IRF4-silencing on DNA methylation, EGR2-silencing clearly impaired the DNA demethylation at its binding sites. In particular, EGR2 was also postulated to initiate the DNA demethylation at inaccessible DMRs, suggesting an epigenetic pioneer role of EGR2 in mediating *de novo* demethylation processes.

Summarizing the relevant properties of pioneer factors, they should be able to recognize their target DNA sequences within “closed” chromatin, to initiate chromatin remodeling, to allow binding of non-pioneer TFs and to establish epigenetic stability of the accessible DNA state[287]. These features allow

pioneer factors to play a critical role in different biological processes, such as during cell reprogramming (e.g. Oct4[288]) and cell-lineage specification (e.g. EBF1[89]). In fact, some of these characteristics are partially shared by the epigenetic pioneer EGR2. For instance, EGR2 has the ability to induce DNA demethylation in closed chromatin regions and also might be able to participate in chromatin remodeling processes as reflected by the ATAC-seq data (Figure 5.20).

Despite our knowledge about the potential role of EGR2 in mediating *de novo* demethylation, the mechanisms involved in TET2 recruitment to its binding sites were not fully understood until we focused our analyses on all EGR2 consensus motifs containing DMRs. Computational analyses revealed that CpGs within the EGR2 core consensus are usually protected from demethylation during differentiation and that this phenomenon occurs across all DMRs groups regardless of EGR2 binding detection (Figure 5.23). Similarly, we verified that this methylation footprint of EGR2 at iDC DMRs is likely lineage-specific as well as a feature shared by certain epigenetic pioneer factors, such as CEBP and KLF.

Members of CEBP and KLF family factors (CEBPα and Klf4) were previously reported to bind methylated DNA[289,290] and recently shown to interact and recruit Tet2 to induce DNA demethylation during reprograming in mouse[98]. In compliance with our observations for EGR2 (also previously reported to bind methylated DNA[101,291]), these data suggest that DNA methylation-spikes overlap with TF binding sites, which may have a pivotal role in cellular processes like active DNA demethylation.

Previous studies (in other systems) that described the ability of some TFs to bind methylated DNA and induce TET2-mediated DNA demethylation at its binding sites, showed a particular interaction between the candidate factor and TET2. To check the presence of an interaction between EGR2 and TET2 in our system as well as to identify candidate interaction partners of EGR2, we used biochemical approaches, such as BioID and CoIP.

Preliminary EGR2 BioID data in iDC (Figure 5.28) unveiled TET2 as a candidate interaction partner of EGR2 (proteins in close proximity) and further CoIP experiments confirmed the TET2-EGR2 interaction (Figure 5.30). These data reinforced our findings, which suggested that EGR2 is essential for TET2 recruitment to its binding sites during *in vitro* MO differentiation.

Complementarily, NAB2, a known co-factor of EGR2[193], was also identified in our EGR2-BioID experiments (Figure 5.28). In a similar way to EGR2, it also co-immunoprecipitated with TET2 (Figure 5.30), indicating a potential involvement of the EGR2:NAB2 complexes in demethylation processes.

The EGR2:NAB complexes have been demonstrated to regulate gene expression during peripheral nerve myelination[292–294] as well as in the hematopoietic system[229,272], where the association between EGR2 and the NAB corepressors represses specific target genes. It was also established that NAB2 represses transcription by interaction with the chromodomain helicase DNA-binding protein 4 (CHD4) subunit of the nucleosome remodeling and deacetylase (NuRD) chromatin remodelling complex[193]. Taking into consideration that work, we may hypothesize (in our system) that EGR2:NAB2 complexes

through NAB2-mediated recruitment of the NuRD complex, might mediate changes in chromatin structure around EGR2 binding sites facilitating stable TF binding. This process would be coupled with changes in DNA methylation triggered by EGR2-mediated recruitment of TET2 to its stable binding sites as shown in Figure 6.1.

Holding our hypothesis, preliminary mass spectrometry data (EGR2-BioID, Figure 5.28) identified chromatin remodeling enzymes (e.g EP400[295], SMARCA2/4[296] and ARID1A[297]) and proteins involved in the epigenetic control of gene transcription (e.g. histone demethylases like JMJD1C and KDM6B[298]) within close proximity to EGR2.

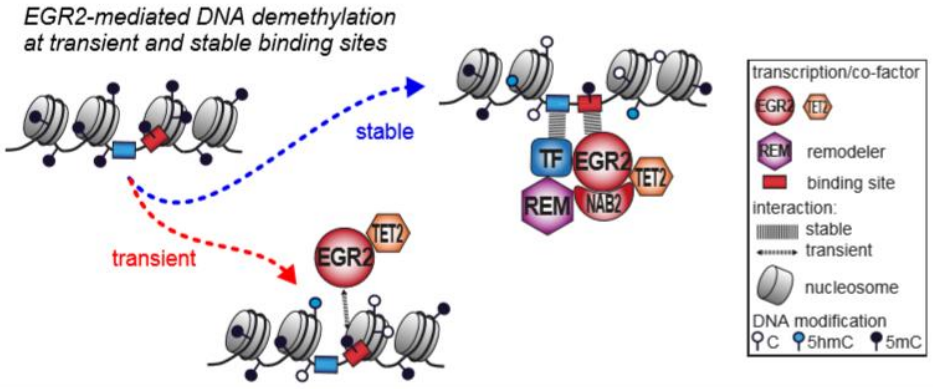

Figure 6.1 Model of EGR2-mediated DNA demethylation at transient and stable binding sites

In addition to EGR2-mediated recruitment of TET2 to its stable binding sites, in a fraction of genomic regions that are actively demethylated regardless of open chromatin or TF binding detection ("no peak" and "open" DMR groups), we assume that EGR2 transiently binds its methylated binding sites and recruits TET2 to induce local demethylation (as represented in the model from Figure 6.1). Generalizing our findings for EGR2, we also hypothesize that pioneer factors like CEBP (showing methylation-spikes at their motifs and no detectable protein binding (Figure 5.33b)) may also be able to transiently bind to its binding sites and induce DNA demethylation at those regions.

In conclusion, the presence of EGR2 consensus motif in iDC-specific DMR, as well as its ability to recognize and bind methylated DNA make EGR2 an epigenetic pioneer in targeting *de novo* demethylation processes at its transient and stable binding sites. Importantly, our data suggest that EGR2 recruits TET2 even in the absence of open chromatin (ATAC) or TF binding (ChIP) detection, and that at some transient sites ("no peak" DMR group) active DNA demethylation and chromatin accessibility changes are uncoupled. In addition, we assume that 5mC- spikes (perhaps indicating DNA demethylation protection) are cell type-specific and usually associated with consensus motifs for known epigenetic pioneer factors like CEBP, KLF of EBF family members.

Therefore, these findings provide novel evidences to better understand how the specificity of demethylation is controlled during post-mitotic MO differentiation. Similarly, the work presented in this thesis represents an important advancement in our understanding about MO biology and in our understanding of how mutations in DNA demethylation machinery might affect the normal hematopoiesis driving malignant transformation.

6.4 Perspectives

This thesis improved and expanded our understanding about MO biology as well as its differentiation-associated epigenetic processes.

In the present work, the investigation of two key regulators in iDC differentiation (IRF4 and EGR2) allowed us to confirm that both TFs are necessary to control iDC-specific transcriptional programs and chromatin remodeling processes. In contrast to IRF4, EGR2 also revealed to be an epigenetic pioneer TF driving *de novo* DNA demethylation processes during MO differentiation.

Although our knowledge of active DNA demethylation processes during MO differentiation, the mechanisms and TF(s) that might be involved in the induction of DNA demethylation at earlier stages of MO development are largely unknown. Therefore, to investigate these processes at earlier MO progenitor cells, we could use $CD34^+$ HSCs from cord blood (CB) in future experiments. Actually, $CD34^+$ HSCs from CB can be used to expand DC precursors cells ($CD14^+$) and to differentiate them into DC in the presence of GM-CSF and IL4 as recently reported by Bedke and colleagues[299]. Hence, using this system in future TET2 knock-down experiments, we could perhaps bypass problems associated with attenuation of TET2 knock-down effects in the current system and may interfere with the DNA demethylation process from an earlier step.

An alternative strategy to counteract attenuation of TET2 knock-down effects in our system would involve the specific inhibition of TET2 protein in the cell before its siRNA-mediated knock-down. However, an effective TET2 inhibitor is currently unavailable.

As previously mentioned, preliminary BioID-EGR2 results revealed potential candidate proteins interacting with EGR2, which may have a pivotal role not only in DNA demethylation but also in other cellular processes like chromatin remodeling. Nevertheless, further BioID experiments are required to increase the reproducibility of our data among different donors and to allow robust statistical analyses. Then, some candidate proteins may be targeted and further investigated in our lab.

Despite the efforts to understand the target specificity of the DNA demethylation machinery during MO differentiation, the processing of 5hmC to 5C in iDCs remains unclear.

While in mouse work[216] it was previously demonstrated that TET2 requires association with another protein to further oxidize 5hmC, in iDCs the presence of such protein or TF is unknown. Similarly, it is

still unclear whether uracil DNA glycosylases like TDG are involved in the 5fC/5caC excision repair in iDCs as reported in other systems[36,300,301].

Although initial work in our group[32] suggests that neither MBD4, nor TDG knock-down lead to an accumulation of 5fC, problems associated with transfection procedures might have masked any specific knock-down effect. Therefore, our well-established transfection protocol can be used in future experiments to knock-down these genes and to study the long-term effects of their depletion on the processing of 5mC. Likewise, this strategy may also be used to interfere with the expression of other genes previously implicated in DNA demethylation, such as NEIL DNA glycosylases[302] and GADD45A[303,304] to increase our understanding about the complete DNA methylation turnover process (5mC → 5C).

7 Summary

Epigenetic processes, such as DNA (de)methylation and chromatin remodeling are fundamental for hematopoietic cell differentiation and lineage specification. Hence, their dysregulation is associated with many hematological malignancies.

In this thesis, the analyses were focused on DNA demethylation processes during monocyte (MO) to immature dendritic cell (iDC) differentiation in the presence of the cytokines IL4 and GM-CSF. Due to the absence of cell proliferation in this system, we were able to investigate TET-mediated active demethylation events occurring in a replication-independent context.

In hematopoietic cells, comprising MO and MO-derived cells, the TET2 enzyme is the main hydroxylase catalysing the initial oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) and likely the remaining oxidative steps as well. In this thesis, the TET2 gene was transiently knocked-down in MO using a well-established siRNA-mediated approach and its functional consequences during iDC differentiation were assessed.

Overall, TET2 depletion was associated with very mild effects on iDC differentiation, transcriptional programs and chromatin remodeling processes. Nevertheless, the data was reproducible among donors and corroborated the importance of TET2-mediated active DNA demethylation processes during MO differentiation.

Based on preliminary genome-wide methylation analyses in MO and iDCs, 7610 iDC-specific differentially methylated regions (DMRs) were identified. These regions were characterized by a progressive active DNA demethylation (5hmC enrichment) and an increase in chromatin accessibility. Since we were mostly interested on differentiation-associated processes, the analyses were centered on transcription factors (TFs) enriched across DMRs and strongly induced in iDCs, such as IRF4 and EGR2. Consequently, both TFs were individually and transiently knocked-down in MO and their effects on iDC differentiation and associated epigenetic processes were determined.

IRF4 depletion prevented the differentiation of iDC, and instead cells acquired a macrophage-like morphology confirming previous mouse work[244]. Interestingly, we found that IRF4 might be involved in open chromatin at its binding sites regardless of changes in DNA demethylation, as reflected by the reduced chromatin accessibility and the minor effects on DNA methylation upon its depletion.

Remarkably, this work is the first to describe an essential role of EGR2 in IL4/GM-CSF-mediated MO differentiation. Indeed, data showed a strong impact of EGR2 depletion on cell morphology and viability as well as profound changes in iDC transcriptional programs.

Besides its importance on MO biology, EGR2 is implicated in active DNA demethylation and chromatin remodeling processes as indicated by iDC DMRs that remained methylated and failed to gain accessibility upon EGR2-silencing.

Notably, the presence of a methylation footprint in the EGR2 motif associated with its ability to bind methylated DNA, suggest that the epigenetic pioneer EGR2 can target *de novo* demethylation processes at its transient and stable binding sites. In fact, co-immunoprecipitation data showed an interaction between EGR2 and TET2, emphasizing the EGR2 role in targeting the DNA demethylation machinery.

In summary, this thesis provides new insights into dynamic epigenetic changes through IL4/GM-CSF-mediated MO differentiation. Accordingly, we identified key regulators on MO biology and differentiation-associated epigenetic processes, specifically IRF4 and EGR2. Importantly, EGR2 was found to recruit TET2 to its binding sites even before chromatin opening or TF binding detection, suggesting that at certain iDC DMRs active DNA demethylation and chromatin accessibility changes are uncoupled. In addition, this work demonstrated that DNA methylation-spikes (identified in EGR2 as well as other consensus motifs of TFs) are cell type-specific and protected from demethylation by bound key epigenetic pioneer factors.

8 Zusammenfassung

Epigenetische Prozesse wie DNA (De)Methylierung und „Chromatin-Remodeling“ sind für die Differenzierung hämatopoetischer Zellen und für ihre spezifische Zellabstammung grundlegend. Daher ist deren Dysregulation mit vielen hämatologischen Malignitäten verbunden.

Diese Doktorarbeit beschreibt in wesentlichen Zügen die DNA-Demethylierungsprozesse während der IL4 / GM-CSF-gesteuerten Differenzierung von Monozyten (MO) zu inaktiven dendritischen Zellen (iDCs). Da die Zellen in diesem System nicht proliferieren, konnten wir die enzymatischen TET-vermittelten aktiven Demethylierungsereignisse untersuchen, die in einem Replikations-unabhängigen Kontext auftreten. In hämatopoetischen Zellen, welche unter anderem die MO- und MO-abstammenden Zellen umfassen, ist das TET2-Enzym hauptsächlich die Hydroxylase, welche die anfängliche Oxidation von 5 mC zu 5 hmC und wahrscheinlich auch die verbleibenden Oxidationsschritte (5fC und 5caC) katalysiert. In dieser Doktorarbeit wurde das TET2-Gen in MO vorübergehend mittels siRNA unterdrückt und dessen funktionelle Konsequenz während der iDC-Differenzierung untersucht.

Insgesamt hatte die TET2-Depletion sehr geringe Auswirkungen auf die iDC-Differenzierung, das Transkriptionsprogramm und die Chromatin-Remodeling-Prozesse. Dennoch waren die Daten unter den Spendern reproduzierbar und bestätigten die Bedeutung von TET2-vermittelten aktiven DNA-Demethylierungsprozessen während der MO-Differenzierung.

Basierend auf vorläufigen genomweiten Methylierungsanalysen in MO und iDCs wurden 7610 iDC-spezifische differentiell methylierte Regionen (DMRs) identifiziert. Diese Regionen waren durch eine fortschreitende aktive Demethylierung der DNA (5hmC-Anreicherung) und eine erhöhte Zugänglichkeit von Chromatin (Euchromatin) charakterisiert.

Da wir hauptsächlich an differenzierungsassoziierten Prozessen interessiert waren, wurden diejenigen Transkriptionsfaktoren (TFs), welche in den iDC-DMRs angereichert sind, analysiert. Zusätzlich zeigten die TFs IRF4 und EGR2 eine starke Induktion in iDC. Folglich wurden beide TFs in MO einzeln und vorübergehend unterdrückt und dessen Auswirkungen auf die iDC-Differenzierung und die damit verbundenen epigenetischen Prozesse bestimmt.

Die IRF4-Depletion verhinderte den iDC-Differenzierungsprozess und die Zellen erwarben stattdessen eine makrophagenähnliche Morphologie, was auch in früheren Arbeiten in Mausmodellen gezeigt werden konnte. Interessanterweise wurde beobachtet, dass IRF4, unabhängig von Änderungen in der DNA-Demethylierung, an seinen Bindungsstellen an der Regulierung von offenem Chromatin beteiligt sein könnte. Dies zeigte sich zusätzlich in der verringerten Zugänglichkeit des Chromatins und den geringen Auswirkungen auf die DNA-Methylierung nach IRF4 Gen-Knock-down.

Bemerkenswerterweise ist diese Arbeit die erste, die eine wesentliche Rolle von EGR2 bei der IL4 / GM-CSF-gesteuerten MO-Differenzierung beschreibt. In der Tat zeigten die Daten einen starken

Einfluss der EGR2-Depletion auf die Zellmorphologie und -lebensfähigkeit sowie signifikante Veränderungen in den iDC-Transkriptionsprogrammen.

Neben seiner Bedeutung für die MO-Biologie ist EGR2 an aktiven DNA-Demethylierungs- und Chromatin-Remodellierungsprozessen beteiligt, was durch eine Reihe von charakteristischen iDCs-DMRs angezeigt wurde. Diese blieben methyliert und erlangten auch nach EGR2-Stummschaltung keine Zugänglichkeit.

Insbesondere das Vorhandensein einer spezifischen DNA-Methylierung im EGR2-Motiv (im Einklang mit seiner Fähigkeit, methylierte DNA zu binden) schlagen vor, dass die EGR2 zu einem „epigenetischen Pionier“ am Beginn von *De-novo*-Demethylierungsprozessen an seinen transienten und stabilen Bindungsstellen.

Tatsächlich zeigten die CoIP-Daten eine Interaktion zwischen EGR2 und TET2, was die Rolle von EGR2 als Teil der DNA-Demethylierungsmaschinerie unterstreicht.

Zusammenfassend liefert diese Arbeit neue Einblicke in dynamische DNA-Methylierungsänderungen während der IL4 / GM-CSF-gesteuerten MO-Differenzierung. Dementsprechend wurden Hauptregulatoren für die MO-Biologie und differenzierungsassoziierte epigenetische Prozesse identifiziert, insbesondere IRF4 und EGR2. Es wurde festgestellt, dass EGR2 TET2 bereits vor dem Öffnen des Chromatins oder dem Nachweis der TF-Bindung an seine Bindungsstellen rekrutiert, was darauf hindeutet, dass bei bestimmten iDC-DMRs Änderungen der aktiven DNA-Demethylierung und der Zugänglichkeit des Chromatins entkoppelt sind. Insgesamt zeigt diese Arbeit, dass DNA-Methylierungsanreicherungen, die möglicherweise den Methylierungsfußabdruck an EGR2-Bindungsstellen widerspiegeln, zelltypspezifisch sind und durch gebundene epigenetische Schlüsselpionierfaktoren vor Demethylierung geschützt werden können.

9 References

1. Handy DE, Castro R, Loscalzo J. Epigenetic modifications: basic mechanisms and role in cardiovascular disease. *Circulation*. 2011;123(19):2145-2156.
2. Sadakierska-Chudy A, Kostrzewa RM, Filip M. A comprehensive view of the epigenetic landscape part I: DNA methylation, passive and active DNA demethylation pathways and histone variants. *Neurotoxicity research*. 2015;27(1):84-97.
3. Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics*. 2019. https://doi.org/10.1038/s41576-018-0089-8.
4. Tsompana M, Buck MJ. Chromatin accessibility: a window into the genome. *Epigenetics & chromatin*. 2014;7(1):33.
5. Lawrence M, Daujat S, Schneider R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends in genetics : TIG*. 2016;32(1):42-56.
6. Benetatos L, Vartholomatos G. Enhancer DNA methylation in acute myeloid leukemia and myelodysplastic syndromes. *Cellular and molecular life sciences : CMLS*. 2018;75(11):1999-2009.
7. Božić T, Frobel J, Raic A, et al. Variants of DNMT3A cause transcript-specific DNA methylation patterns and affect hematopoiesis. *Life science alliance*. 2018;1(6):e201800153.
8. Bröske A-M, Vockentanz L, Kharazi S, et al. DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nature Genetics*. 2009;41:1207 EP -. https://doi.org/10.1038/ng.463.
9. Gu J, Stevens M, Xing X, et al. Mapping of Variable DNA Methylation Across Multiple Cell Types Defines a Dynamic Regulatory Landscape of the Human Genome. *G3 (Bethesda, Md.)*. 2016;6(4):973-986.
10. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*. 2013;502(7472):472-479.
11. Lessard S, Beaudoin M, Benkirane K, Lettre G. Comparison of DNA methylation profiles in human fetal and adult red blood cell progenitors. *Genome medicine*. 2015;7(1):1.
12. Agirre X, Castellano G, Pascual M, et al. Whole-epigenome analysis in multiple myeloma reveals DNA hypermethylation of B cell-specific enhancers. *Genome research*. 2015;25(4):478-487.
13. Álvarez-Errico D, Vento-Tormo R, Sieweke M, Ballestar E. Epigenetic control of myeloid cell differentiation, identity and function. *Nature reviews. Immunology*. 2015;15(1):7-17.
14. Bestor TH, Bourc'his D. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harbor symposia on quantitative biology*. 2004;69:381-387.
15. Bird A. DNA methylation patterns and epigenetic memory. *Genes & development*. 2002;16(1):6-21.
16. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics*. 2003;33 Suppl:245-254.
17. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nature reviews. Genetics*. 2017;18(9):517-534.
18. Huang D, Cui L, Ahmed S, et al. An overview of epigenetic agents and natural nutrition products targeting DNA methyltransferase, histone deacetylases and microRNAs. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association*. 2019;123:574-594.
19. Lamadema N, Burr S, Brewer AC. Dynamic regulation of epigenetic demethylation by oxygen availability and cellular redox. *Free radical biology & medicine*. 2019;131:282-298.
20. Portela A, Esteller M. Epigenetic modifications and human disease. *Nature biotechnology*. 2010;28(10):1057-1068.
21. Joseph DB, Strand DW, Vezina CM. DNA methylation in development and disease: an overview for prostate researchers. *American Journal of Clinical and Experimental Urology*. 2018;6(6):197-218.
22. Hon GC, Rajagopal N, Shen Y, et al. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nature Genetics*. 2013;45:1198 EP -. https://doi.org/10.1038/ng.2746.
23. Ziller MJ, Gu H, Müller F, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature*. 2013;500:477 EP -. https://doi.org/10.1038/nature12433.
24. Ravichandran M, Jurkowska RZ, Jurkowski TP. Target specificity of mammalian DNA methylation and demethylation machinery. *Organic & biomolecular chemistry*. 2018;16(9):1419-1435.
25. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science (New York, N.Y.)*. 2009;324(5929):929-930.
26. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (New York, N.Y.)*. 2009;324(5929):930-935.
27. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nature reviews. Molecular cell biology*. 2013;14(6):341-356.
28. Saitou M, Kagiwada S, Kurimoto K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development (Cambridge, England)*. 2012;139(1):15-31.

29. Guo JU, Su Y, Zhong C, Ming G-l, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*. 2011;145(3):423-434.
30. Li X, Wei W, Ratnu VS, Bredy TW. On the potential role of active DNA demethylation in establishing epigenetic states associated with neural plasticity and memory. *Neurobiology of learning and memory*. 2013;105:125-132.
31. Klug M, Heinz S, Gebhard C, et al. Active DNA demethylation in human postmitotic cells correlates with activating histone modifications, but not transcription levels. *Genome biology*. 2010;11(6):R63.
32. Klug M, Schmidhofer S, Gebhard C, Andreesen R, Rehli M. 5-Hydroxymethylcytosine is an essential intermediate of active DNA demethylation processes in primary human monocytes. *Genome biology*. 2013;14(5):R46.
33. Rasmussen KD, Helin K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes & development*. 2016;30(7):733-750.
34. Bochtler M, Kolano A, Xu G-L. DNA demethylation pathways: Additional players and regulators. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 2017;39(1):1-13.
35. Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia*. 2003;17:637 EP -. https://doi.org/10.1038/sj.leu.2402834.
36. He Y-F, Li B-Z, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science (New York, N.Y.)*. 2011;333(6047):1303-1307.
37. Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science (New York, N.Y.)*. 2011;333(6047):1300-1303.
38. Liu MY, Torabifard H, Crawford DJ, et al. Mutations along a TET2 active site scaffold stall oxidation at 5-hydroxymethylcytosine. *Nature chemical biology*. 2017;13(2):181-187.
39. Tamanaha E, Guan S, Marks K, Saleh L. Distributive Processing by the Iron(II)/α-Ketoglutarate-Dependent Catalytic Domains of the TET Enzymes Is Consistent with Epigenetic Roles for Oxidized 5-Methylcytosine Bases. *Journal of the American Chemical Society*. 2016;138(30):9345-9348.
40. Bachman M, Uribe-Lewis S, Yang X, Williams M, Murrell A, Balasubramanian S. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nature Chemistry*. 2014;6:1049 EP -. https://doi.org/10.1038/nchem.2064.
41. Bachman M, Uribe-Lewis S, Yang X, et al. 5-Formylcytosine can be a stable DNA modification in mammals. *Nature chemical biology*. 2015;11:555 EP -. https://doi.org/10.1038/nchembio.1848.
42. Raiber E-A, Murat P, Chirgadze DY, Beraldi D, Luisi BF, Balasubramanian S. 5-Formylcytosine alters the structure of the DNA double helix. *Nature structural & molecular biology*. 2015;22(1):44-49.
43. Pfaffeneder T, Spada F, Wagner M, et al. Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nature chemical biology*. 2014;10(7):574-581.
44. Nakajima H, Kunimoto H. TET2 as an epigenetic master regulator for normal and malignant hematopoiesis. *Cancer science*. 2014;105(9):1093-1099.
45. Chen Q, Chen Y, Bian C, Fujiki R, Yu X. TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature*. 2013;493(7433):561-564.
46. Kreppel LK, Blomberg MA, Hart GW. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *The Journal of biological chemistry*. 1997;272(14):9308-9315.
47. Vosseller K, Sakabe K, Wells L, Hart GW. Diverse regulation of protein function by O-GlcNAc: a nuclear and cytoplasmic carbohydrate post-translational modification. *Current Opinion in Chemical Biology*. 2002;6(6):851-857.
48. Deplus R, Delatte B, Schwinn MK, et al. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *The EMBO journal*. 2013;32(5):645-655.
49. Lu X, Zhao BS, He C. TET family proteins: oxidation activity, interacting molecules, and functions in diseases. *Chemical reviews*. 2015;115(6):2225-2239.
50. Losman J-A, Kaelin WG. What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. *Genes & development*. 2013;27(8):836-852.
51. Blaschke K, Ebata KT, Karimi MM, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature*. 2013;500(7461):222-226.
52. Gustafson CB, Yang C, Dickson KM, et al. Epigenetic reprogramming of melanoma cells by vitamin C treatment. *Clinical epigenetics*. 2015;7:51.
53. Mingay M, Chaturvedi A, Bilenky M, et al. Vitamin C-induced epigenomic remodelling in IDH1 mutant acute myeloid leukaemia. *Leukemia*. 2018;32(1):11-20.
54. Minor EA, Court BL, Young JI, Wang G. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *The Journal of biological chemistry*. 2013;288(19):13669-13674.
55. Sasidharan Nair V, Song MH, Oh KI. Vitamin C Facilitates Demethylation of the Foxp3 Enhancer in a Tet-Dependent Manner. *Journal of immunology (Baltimore, Md. : 1950)*. 2016;196(5):2119-2131.

56. Yin R, Mao S-Q, Zhao B, et al. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *Journal of the American Chemical Society*. 2013;135(28):10396-10403.
57. Garcia-Gomez A, Li T, Kerick M, et al. TET2- and TDG-mediated changes are required for the acquisition of distinct histone modifications in divergent terminal differentiation of myeloid cells. *Nucleic acids research*. 2017;45(17):10002-10017.
58. Delhommeau F, Dupont S, Della Valle V, et al. Mutation in TET2 in myeloid cancers. *The New England journal of medicine*. 2009;360(22):2289-2301.
59. Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *The New England journal of medicine*. 2010;363(25):2424-2433.
60. Tefferi A, Lim K-H, Abdel-Wahab O, et al. Detection of mutant TET2 in myeloid malignancies other than myeloproliferative neoplasms: CMML, MDS, MDS/MPN and AML. *Leukemia*. 2009;23(7):1343-1345.
61. Medeiros BC, Fathi AT, DiNardo CD, Pollyea DA, Chan SM, Swords R. Isocitrate dehydrogenase mutations in myeloid malignancies. *Leukemia*. 2017;31(2):272-281.
62. Dang L, White DW, Gross S, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature*. 2010;465(7300):966.
63. Pollard PJ, Brière JJ, Alam NA, et al. Accumulation of Krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations. *Human molecular genetics*. 2005;14(15):2231-2239.
64. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer cell*. 2010;18(6):553-567.
65. Xu W, Yang H, Liu Y, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer cell*. 2011;19(1):17-30.
66. An J, Rao A, Ko M. TET family dioxygenases and DNA demethylation in stem cells and cancers. *Experimental & molecular medicine*. 2017;49(4):e323.
67. Laukka T, Mariani CJ, Ihantola T, et al. Fumarate and Succinate Regulate Expression of Hypoxia-inducible Genes via TET Enzymes. *The Journal of biological chemistry*. 2016;291(8):4256-4265.
68. Xiao M, Yang H, Xu W, et al. Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes & development*. 2012;26(12):1326-1338.
69. Hashimoto H, Liu Y, Upadhyay AK, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic acids research*. 2012;40(11):4841-4849.
70. Valinluck V, Tsai H-H, Rogstad DK, Burdzy A, Bird A, Sowers LC. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic acids research*. 2004;32(14):4100-4108.
71. Shearstone JR, Pop R, Bock C, Boyle P, Meissner A, Socolovsky M. Global DNA demethylation during mouse erythropoiesis in vivo. *Science (New York, N.Y.)*. 2011;334(6057):799-802.
72. Frauer C, Hoffmann T, Bultmann S, et al. Recognition of 5-Hydroxymethylcytosine by the Uhrf1 SRA Domain. *PLoS ONE*. 2011;6(6).
73. Spruijt CG, Gnerlich F, Smits AH, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell*. 2013;152(5):1146-1159.
74. Cortellino S, Xu J, Sannai M, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*. 2011;146(1):67-79.
75. Hegde ML, Banerjee S, Hegde PM, et al. Enhancement of NEIL1 protein-initiated oxidized DNA base excision repair by heterogeneous nuclear ribonucleoprotein U (hnRNP-U) via direct interaction. *The Journal of biological chemistry*. 2012;287(41):34202-34211.
76. Olinski R, Starczak M, Gackowski D. Enigmatic 5-hydroxymethyluracil: Oxidatively modified base, epigenetic mark or both? *Mutation research. Reviews in mutation research*. 2016;767:59-66.
77. Scourzic L, Mouly E, Bernard OA. TET proteins and the control of cytosine demethylation in cancer. *Genome medicine*. 2015;7(1):9.
78. Lio C-WJ, Yue X, Lopez-Moyado IF, Tahiliani M, Aravind L, Rao A. TET methylcytosine oxidases: new insights from a decade of research. *Journal of biosciences*. 2020;45.
79. Han J-A, An J, Ko M. Functions of TET Proteins in Hematopoietic Transformation. *Molecules and cells*. 2015;38(11):925-935.
80. Cortázar D, Kunz C, Selfridge J, et al. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature*. 2011;470(7334):419-423.
81. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *The Journal of biological chemistry*. 2011;286(41):35334-35338.
82. Schomacher L, Han D, Musheev MU, et al. Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation. *Nature structural & molecular biology*. 2016;23(2):116-124.

83. Fromme JC, Verdine GL. Base Excision Repair. In: Parry DAD, Squire JM, eds. *Fibrous proteins: coiled-coils, collagen and elastomers*. Vol. 69. Amsterdam: Elsevier; 2005:1-41. *Advances in Protein Chemistry*; vol. 70.
84. Lambert SA, Jolma A, Campitelli LF, et al. The Human Transcription Factors. *Cell*. 2018;172(4):650-665.
85. Schütte J, Wang H, Antoniou S, et al. An experimentally validated network of nine haematopoietic transcription factors reveals mechanisms of cell state stability. *eLife*. 2016;5:e11469.
86. Boeva V. Analysis of Genomic Sequence Motifs for Deciphering Transcription Factor Binding and Transcriptional Regulation in Eukaryotic Cells. *Frontiers in genetics*. 2016;7:24.
87. Wang J, Zhuang J, Iyer S, et al. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome research*. 2012;22(9):1798-1812.
88. Burda P, Laslo P, Stopka T. The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis. *Leukemia*. 2010;24(7):1249-1257.
89. György I, Boller S, Nechanitzky R, et al. Transcription factor Ebf1 regulates differentiation stage-specific signaling, proliferation, and survival of B cells. *Genes & development*. 2012;26(7):668-682.
90. Di Tullio A, Vu Manh TP, Schubert A, Castellano G, Månsson R, Graf T. CCAAT/enhancer binding protein alpha (C/EBP(alpha))-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(41):17016-17021.
91. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature*. 2007;448(7151):313-317.
92. Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell*. 2015;161(3):555-568.
93. Mayran A, Sochodolsky K, Khetchoumian K, et al. Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening. *Nature Communications*. 2019;10(1):3807.
94. Héberlé É, Bardet AF. Sensitivity of transcription factors to DNA methylation. *Essays in biochemistry*. 2019;63(6):727-741.
95. Ambrosi C, Manzo M, Baubec T. Dynamics and Context-Dependent Roles of DNA Methylation. *Journal of molecular biology*. 2017;429(10):1459-1475.
96. Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schübeler D. Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature*. 2015;528(7583):575-579.
97. La Rica L de, Rodríguez-Ubreva J, García M, et al. PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome biology*. 2013;14(9):R99.
98. Sardina JL, Collombet S, Tian TV, et al. Transcription Factors Drive Tet2-Mediated Enhancer Demethylation to Reprogram Cell Fate. *Cell stem cell*. 2018;23(5):727-741.e9.
99. Sun Z, Xu X, He J, et al. EGR1 recruits TET1 to shape the brain methylome during development and upon neuronal activity. *Nature Communications*. 2019;10(1):3892.
100. Stadler MB, Murr R, Burger L, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature*. 2011;480(7378):490-495.
101. Yin Y, Morgunova E, Jolma A, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science (New York, N.Y.)*. 2017;356(6337).
102. Xu Y, Xu C, Kato A, et al. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. *Cell*. 2012;151(6):1200-1213.
103. Rampal R, Alkalin A, Madzo J, et al. DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. *Cell reports*. 2014;9(5):1841-1855.
104. Wang Y, Xiao M, Chen X, et al. WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. *Molecular cell*. 2015;57(4):662-673.
105. Guilhamon P, Eskandarpour M, Halai D, et al. Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2. *Nature Communications*. 2013;4:2166.
106. Suzuki T, Shimizu Y, Furuhata E, et al. RUNX1 regulates site specificity of DNA demethylation by recruitment of DNA demethylation machineries in hematopoietic cells. *Blood advances*. 2017;1(20):1699-1711.
107. Chen L-L, Lin H-P, Zhou W-J, et al. SNIP1 Recruits TET2 to Regulate c-MYC Target Genes and Cellular DNA Damage Response. *Cell reports*. 2018;25(6):1485-1500.e4.
108. Costa Y, Ding J, Theunissen TW, et al. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*. 2013;495(7441):370-374.
109. Okashita N, Kumaki Y, Ebi K, et al. PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development (Cambridge, England)*. 2014;141(2):269-280.
110. Neri F, Incarnato D, Krepelova A, et al. Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome biology*. 2013;14(8):R91.

111. Zeng Y, Yao B, Shin J, et al. Lin28A Binds Active Promoters and Recruits Tet1 to Regulate Gene Expression. *Molecular cell*. 2016;61(1):153-160.
112. Guilliams M, Ginhoux F, Jakubzick C, et al. Dendritic cells, monocytes and macrophages: a unified nomenclature based on ontogeny. *Nature reviews. Immunology*. 2014;14(8):571-578.
113. Roussel M, Ferrell PB, Greenplate AR, et al. Mass cytometry deep phenotyping of human mononuclear phagocytes and myeloid-derived suppressor cells from human blood and bone marrow. *Journal of leukocyte biology*. 2017;102(2):437-447.
114. Gordon S, Plüddemann A. The Mononuclear Phagocytic System. Generation of Diversity. *Frontiers in immunology*. 2019;10:1893.
115. Vegh P, Fletcher J, Dixon D, Haniffa M. Mononuclear Phagocyte System. In: Ltd JW&S, ed. *Encyclopedia of life sciences*. London, New York, Vols. 21-32, Chichester, West Sussex, U.K.: Nature Pub. Group; Wiley; 2002-2010:1-8.
116. Pahari S, Kaur G, Negi S, et al. Reinforcing the Functionality of Mononuclear Phagocyte System to Control Tuberculosis. *Frontiers in immunology*. 2018;9:193.
117. Germic N, Frangez Z, Yousefi S, Simon H-U. Regulation of the innate immune system by autophagy: monocytes, macrophages, dendritic cells and antigen presentation. *Cell Death & Differentiation*. 2019. https://doi.org/10.1038/s41418-019-0297-6.
118. Sokol CL, Luster AD. The chemokine system in innate immunity. *Cold Spring Harbor perspectives in biology*. 2015;7(5).
119. Valledor AF, Comalada M, Santamaría-Babi LF, Lloberas J, Celada A. Macrophage proinflammatory activation and deactivation: a question of balance. *Advances in immunology*. 2010;108:1-20.
120. Engblom C, Pfirschke C, Pittet MJ. The role of myeloid cells in cancer therapies. *Nature reviews. Cancer*. 2016;16(7):447-462.
121. Lavin Y, Mortha A, Rahman A, Merad M. Regulation of macrophage development and function in peripheral tissues. *Nature reviews. Immunology*. 2015;15(12):731-744.
122. Merad M, Sathe P, Helft J, Miller J, Mortha A. The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annual review of immunology*. 2013;31:563-604.
123. Hume DA. The mononuclear phagocyte system. *Current opinion in immunology*. 2006;18(1):49-53.
124. Lavoie PM, Levy O. Mononuclear Phagocyte System. In: Polin RA, Abman SH, Rowitch DH, Benitz WE, Fox WW, eds. *Fetal and neonatal physiology*. Fifth edition. Philadelphia, PA: Elsevier; 2016:1208-1216.e3.
125. Sallusto F, Lanzavecchia A. Efficient presentation of soluble antigen by cultured human dendritic cells is maintained by granulocyte/macrophage colony-stimulating factor plus interleukin 4 and downregulated by tumor necrosis factor alpha. *The Journal of Experimental Medicine*. 1994;179(4):1109-1118.
126. Smedt T de, Pajak B, Muraille E, et al. Regulation of dendritic cell numbers and maturation by lipopolysaccharide in vivo. *The Journal of Experimental Medicine*. 1996;184(4):1413-1424.
127. Collin M, Bigley V. Human dendritic cell subsets: an update. *Immunology*. 2018;154(1):3-20.
128. Richter L, Landsverk OJB, Atlasy N, et al. Transcriptional profiling reveals monocyte-related macrophages phenotypically resembling DC in human intestine. *Mucosal immunology*. 2018;11(5):1512-1523.
129. Coillard A, Segura E. In vivo Differentiation of Human Monocytes. *Frontiers in immunology*. 2019;10:1907.
130. Guttman-Yassky E, Lowes MA, Fuentes-Duculan J, et al. Major differences in inflammatory dendritic cells and their products distinguish atopic dermatitis from psoriasis. *The Journal of allergy and clinical immunology*. 2007;119(5):1210-1217.
131. Laoui D, Keirsse J, Morias Y, et al. The tumour microenvironment harbours ontogenically distinct dendritic cell populations with opposing effects on tumour immunity. *Nature Communications*. 2016;7:13720.
132. Michea P, Noël F, Zakine E, et al. Adjustment of dendritic cells to the breast-cancer microenvironment is subset specific. *Nature Immunology*. 2018;19(8):885-897.
133. Andreesen R, Picht J, Löhr GW. Primary cultures of human blood-born macrophages grown on hydrophobic teflon membranes. *Journal of immunological methods*. 1983;56(3):295-304.
134. Ohradanova-Repic A, Machacek C, Fischer MB, Stockinger H. Differentiation of human monocytes and derived subsets of macrophages and dendritic cells by the HLDA10 monoclonal antibody panel. *Clinical & translational immunology*. 2016;5(1):e55.
135. Bajpai G, Schneider C, Wong N, et al. The human heart contains distinct macrophage subsets with divergent origins and functions. *Nature medicine*. 2018;24(8):1234-1245.
136. Eguíluz-Gracia I, Schultz HHL, Sikkeland LIB, et al. Long-term persistence of human donor alveolar macrophages in lung transplant recipients. *Thorax*. 2016;71(11):1006-1011.
137. MacParland SA, Liu JC, Ma X-Z, et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nature Communications*. 2018;9(1):4383.
138. Goudot C, Coillard A, Villani A-C, et al. Aryl Hydrocarbon Receptor Controls Monocyte Differentiation into Dendritic Cells versus Macrophages. *Immunity*. 2017;47(3):582-596.e6.

139. Seta N, Kuwana M. Human circulating monocytes as multipotential progenitors. *The Keio journal of medicine*. 2007;56(2):41-47.
140. Olingy CE, Dinh HQ, Hedrick CC. Monocyte heterogeneity and functions in cancer. *Journal of leukocyte biology*. 2019.
141. Hettinger J, Richards DM, Hansson J, et al. Origin of monocytes and macrophages in a committed progenitor. *Nature Immunology*. 2013;14:821 EP -. https://doi.org/10.1038/ni.2638.
142. Martin AP, Rankin S, Pitchford S, Charo IF, Furtado GC, Lira SA. Increased expression of CCL2 in insulin-producing cells of transgenic mice promotes mobilization of myeloid cells from the bone marrow, marked insulitis, and diabetes. *Diabetes*. 2008;57(11):3025-3033.
143. Qian B-Z, Li J, Zhang H, et al. CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature*. 2011;475(7355):222-225.
144. Swirski FK, Nahrendorf M, Etzrodt M, et al. Identification of splenic reservoir monocytes and their deployment to inflammatory sites. *Science (New York, N.Y.)*. 2009;325(5940):612-616.
145. Tacke F, Alvarez D, Kaplan TJ, et al. Monocyte subsets differentially employ CCR2, CCR5, and CX3CR1 to accumulate within atherosclerotic plaques. *The Journal of clinical investigation*. 2007;117(1):185-194.
146. Randolph GJ, Inaba K, Robbiani DF, Steinman RM, Muller WA. Differentiation of phagocytic monocytes into lymph node dendritic cells in vivo. *Immunity*. 1999;11(6):753-761.
147. Auffray C, Fogg D, Garfa M, et al. Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. *Science (New York, N.Y.)*. 2007;317(5838):666-670.
148. Carlin LM, Stamatiades EG, Auffray C, et al. Nr4a1-dependent Ly6C(low) monocytes monitor endothelial cells and orchestrate their disposal. *Cell*. 2013;153(2):362-375.
149. Cros J, Cagnard N, Woollard K, et al. Human CD14dim monocytes patrol and sense nucleic acids and viruses via TLR7 and TLR8 receptors. *Immunity*. 2010;33(3):375-386.
150. Hanna RN, Cekic C, Sag D, et al. Patrolling monocytes control tumor metastasis to the lung. *Science (New York, N.Y.)*. 2015;350(6263):985-990.
151. Nagamura-Inoue T, Tamura T, Ozato K. Transcription factors that regulate growth and differentiation of myeloid cells. *International reviews of immunology*. 2001;20(1):83-105.
152. Zhu YP, Thomas GD, Hedrick CC. 2014 Jeffrey M. Hoeg Award Lecture: Transcriptional Control of Monocyte Development. *Arteriosclerosis, thrombosis, and vascular biology*. 2016;36(9):1722-1733.
153. Murray PJ. Immune regulation by monocytes. *Seminars in immunology*. 2018;35:12-18.
154. Sichien D, Scott CL, Martens L, et al. IRF8 Transcription Factor Controls Survival and Function of Terminally Differentiated Conventional and Plasmacytoid Dendritic Cells, Respectively. *Immunity*. 2016;45(3):626-640.
155. Ziegler-Heitbrock L, Ancuta P, Crowe S, et al. Nomenclature of monocytes and dendritic cells in blood. *Blood*. 2010;116(16):e74-80.
156. Thomas GD, Hamers AAJ, Nakao C, et al. Human Blood Monocyte Subsets: A New Gating Strategy Defined Using Cell Surface Markers Identified by Mass Cytometry. *Arteriosclerosis, thrombosis, and vascular biology*. 2017;37(8):1548-1558.
157. Schmidl C, Renner K, Peter K, et al. Transcription and enhancer profiling in human monocyte subsets. *Blood*. 2014;123(17):e90-9.
158. Chimen M, Yates CM, McGettrick HM, et al. Monocyte Subsets Coregulate Inflammatory Responses by Integrated Signaling through TNF and IL-6 at the Endothelial Cell Interface. *Journal of immunology (Baltimore, Md. : 1950)*. 2017;198(7):2834-2843.
159. Sampath P, Moideen K, Ranganathan UD, Bethunaickan R. Monocyte Subsets: Phenotypes and Function in Tuberculosis Infection. *Frontiers in immunology*. 2018;9:1726.
160. Wong KL, Tai JJ-Y, Wong W-C, et al. Gene expression profiling reveals the defining features of the classical, intermediate, and nonclassical human monocyte subsets. *Blood*. 2011;118(5):e16-31.
161. Steinman RM, Cohn ZA. IDENTIFICATION OF A NOVEL CELL TYPE IN PERIPHERAL LYMPHOID ORGANS OF MICE : I. MORPHOLOGY, QUANTITATION, TISSUE DISTRIBUTION. *The Journal of Experimental Medicine*. 1973;137(5):1142-1162.
162. Al-Ashmawy GMZ. Dendritic Cell Subsets, Maturation and Function. In: Chapoval SP, ed. *Dendritic Cells*: InTech; 2018.
163. Lin Q, Chauvistré H, Costa IG, et al. Epigenetic program and transcription factor circuitry of dendritic cell development. *Nucleic acids research*. 2015;43(20):9680-9693.
164. Schinnerling K, García-González P, Aguillón JC. Gene Expression Profiling of Human Monocyte-derived Dendritic Cells - Searching for Molecular Regulators of Tolerogenicity. *Frontiers in immunology*. 2015;6:528.
165. Jackson SH, Yu C-R, Mahdi RM, Ebong S, Egwuagu CE. Dendritic cell maturation requires STAT1 and is under feedback regulation by suppressors of cytokine signaling. *Journal of immunology (Baltimore, Md. : 1950)*. 2004;172(4):2307-2315.
166. Qian C, Cao X. Dendritic cells in the regulation of immunity and inflammation. *Seminars in immunology*. 2018;35:3-11.
167. Tian Y, Meng L, Zhang Y. Epigenetic Regulation of Dendritic Cell Development and Function. *Cancer journal (Sudbury, Mass.)*. 2017;23(5):302-307.

168. Barrat FJ, Su L. A pathogenic role of plasmacytoid dendritic cells in autoimmunity and chronic viral infection. *The Journal of Experimental Medicine*. 2019;216(9):1974-1985.
169. Poulin LF, Lasseaux C, Chamaillard M. Understanding the Cellular Origin of the Mononuclear Phagocyte System Sheds Light on the Myeloid Postulate of Immune Paralysis in Sepsis. *Frontiers in immunology*. 2018;9:823.
170. Swiecki M, Colonna M. The multifaceted biology of plasmacytoid dendritic cells. *Nature reviews. Immunology*. 2015;15(8):471-485.
171. Jongbloed SL, Kassianos AJ, McDonald KJ, et al. Human CD141+ (BDCA-3)+ dendritic cells (DCs) represent a unique myeloid DC subset that cross-presents necrotic cell antigens. *The Journal of Experimental Medicine*. 2010;207(6):1247-1260.
172. Murphy TL, Grajales-Reyes GE, Wu X, et al. Transcriptional Control of Dendritic Cell Development. *Annual review of immunology*. 2016;34:93-119.
173. Tussiwand R, Everts B, Grajales-Reyes GE, et al. Klf4 expression in conventional dendritic cells is required for T helper 2 cell responses. *Immunity*. 2015;42(5):916-928.
174. Reis ES, Barbuto JAM, Köhl J, Isaac L. Impaired dendritic cell differentiation and maturation in the absence of C3. *Molecular immunology*. 2008;45(7):1952-1962.
175. van Tendeloo VF, Ponsaerts P, Lardon F, et al. Highly efficient gene delivery by mRNA electroporation in human hematopoietic cells: superiority to lipofection and passive pulsing of mRNA and to electroporation of plasmid cDNA for tumor antigen loading of dendritic cells. *Blood*. 2001;98(1):49-56.
176. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*. 2009;6(5):343-345.
177. Ehrich M, Nelson MR, Stanssens P, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102(44):15785-15790.
178. Roux KJ, Kim DI, Burke B. BioID: a screen for protein-protein interactions. *Current protocols in protein science*. 2013;74:Unit 19.23.
179. Yaciuk P. Co-immunoprecipitation of protein complexes. *Methods in molecular medicine*. 2007;131:103-111.
180. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature methods*. 2012;9(4):357-359.
181. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell*. 2010;38(4):576-589.
182. E.P. Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57 EP -. https://doi.org/10.1038/nature11247.
183. Derrien T, Estellé J, Marco Sola S, et al. Fast computation and applications of genome mappability. *PLoS ONE*. 2012;7(1):e30377.
184. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current protocols in molecular biology*. 2015;109:21.29.1-21.29.9.
185. Corces MR, Trevino AE, Hamilton EG, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nature methods*. 2017;14(10):959-962.
186. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*. 2010;26(1):139-140.
187. Hansen KD, Irizarry RA, Wu Z. Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics (Oxford, England)*. 2012;13(2):204-216.
188. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*. 2013;29(1):15-21.
189. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications*. 2019;10(1):1523.
190. Suzuki S, Honma K, Matsuyama T, et al. Critical roles of interferon regulatory factor 4 in CD11bhighCD8alpha- dendritic cell development. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101(24):8981-8986.
191. Tailor P, Tamura T, Ozato K. IRF family proteins and type I interferon induction in dendritic cells. *Cell research*. 2006;16(2):134-140.
192. Tamura T, Tailor P, Yamaoka K, et al. IFN regulatory factor-4 and -8 govern dendritic cell subset development and their functional diversity. *Journal of immunology (Baltimore, Md. : 1950)*. 2005;174(5):2573-2581.
193. Srinivasan R, Mager GM, Ward RM, Mayer J, Svaren J. NAB2 represses transcription by interacting with the CHD4 subunit of the nucleosome remodeling and deacetylase (NuRD) complex. *The Journal of biological chemistry*. 2006;281(22):15129-15137.
194. Chiba S. Dysregulation of TET2 in hematologic malignancies. *International journal of hematology*. 2017;105(1):17-22.
195. Hu D, Shilatifard A. Epigenetics of hematopoiesis and hematological malignancies. *Genes & development*. 2016;30(18):2021-2041.

196. Ko M, Huang Y, Jankowska AM, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*. 2010;468(7325):839-843.
197. Rasmussen KD, Berest I, Keβler S, et al. TET2 binding to enhancers facilitates transcription factor recruitment in hematopoietic cells. *Genome research*. 2019;29(4):564-575.
198. Solary E, Bernard OA, Tefferi A, Fuks F, Vainchenker W. The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. *Leukemia*. 2014;28(3):485-496.
199. Mullighan CG. TET2 mutations in myelodysplasia and myeloid malignancies. *Nature Genetics*. 2009;41(7):766-767.
200. Pronier E, Delhommeau F. Role of TET2 mutations in myeloproliferative neoplasms. *Current hematologic malignancy reports*. 2012;7(1):57-64.
201. Duployez N, Goursaud L, Fenwarth L, et al. Familial myeloid malignancies with germline TET2 mutation. *Leukemia*. 2019.
202. Dominguez PM, Ghamlouch H, Rosikiewicz W, et al. TET2 Deficiency Causes Germinal Center Hyperplasia, Impairs Plasma Cell Differentiation, and Promotes B-cell Lymphomagenesis. *Cancer discovery*. 2018;8(12):1632-1653.
203. Kubuki Y, Yamaji T, Hidaka T, et al. TET2 mutation in diffuse large B-cell lymphoma. *Journal of clinical and experimental hematopathology : JCEH*. 2017;56(3):145-149.
204. Hiasa M, Abe M, Nakano A, et al. GM-CSF and IL-4 induce dendritic cell differentiation and disrupt osteoclastogenesis through M-CSF receptor shedding by up-regulation of TNF-alpha converting enzyme (TACE). *Blood*. 2009;114(20):4517-4526.
205. Sander J, Schmidt SV, Cirovic B, et al. Cellular Differentiation of Human Monocytes Is Regulated by Time-Dependent Interleukin-4 Signaling and the Transcriptional Regulator NCOR2. *Immunity*. 2017;47(6):1051-1066.e12.
206. Vento-Tormo R, Company C, Rodríguez-Ubreva J, et al. IL-4 orchestrates STAT6-mediated DNA demethylation leading to dendritic cell differentiation. *Genome biology*. 2016;17:4.
207. Pacis A, Tailleux L, Morin AM, et al. Bacterial infection remodels the DNA methylation landscape of human dendritic cells. *Genome research*. 2015;25(12):1801-1811.
208. Zhang X, Ulm A, Somineni HK, et al. DNA methylation dynamics during ex vivo differentiation and maturation of human dendritic cells. *Epigenetics & chromatin*. 2014;7:21.
209. Vento-Tormo R, Álvarez-Errico D, Rodríguez-Ubreva J, Ballestar E. Gains of DNA methylation in myeloid terminal differentiation are dispensable for gene silencing but influence the differentiated phenotype. *The FEBS journal*. 2015;282(9):1815-1825.
210. Tagoh H, Melnik S, Lefevre P, Chong S, Riggs AD, Bonifer C. Dynamic reorganization of chromatin structure and selective DNA demethylation prior to stable enhancer complex formation during differentiation of primary hematopoietic cells in vitro. *Blood*. 2004;103(8):2950-2955.
211. Alver BH, Kim KH, Lu P, et al. The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nature Communications*. 2017;8:14648.
212. Romero OA, Sanchez-Cespedes M. The SWI/SNF genetic blockade: effects in cell differentiation, cancer and developmental diseases. *Oncogene*. 2014;33(21):2681-2689.
213. Witzel M, Petersheim D, Fan Y, et al. Chromatin-remodeling factor SMARCD2 regulates transcriptional networks controlling differentiation of neutrophil granulocytes. *Nature Genetics*. 2017;49(5):742-752.
214. Sepulveda H, Villagra A, Montecino M. Tet-Mediated DNA Demethylation Is Required for SWI/SNF-Dependent Chromatin Remodeling and Histone-Modifying Activities That Trigger Expression of the Sp7 Osteoblast Master Gene during Mesenchymal Lineage Commitment. *Molecular and Cellular Biology*. 2017;37(20).
215. Hu L, Lu J, Cheng J, et al. Structural insight into substrate preference for TET-mediated oxidation. *Nature*. 2015;527(7576):118-122.
216. Xiong J, Zhang Z, Chen J, et al. Cooperative Action between SALL4A and TET Proteins in Stepwise Oxidation of 5-Methylcytosine. *Molecular cell*. 2016;64(5):913-925.
217. Wu H, Wu X, Shen L, Zhang Y. Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. *Nature biotechnology*. 2014;32(12):1231-1240.
218. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010;466(7310):1129-1133.
219. Shen L, Inoue A, He J, Liu Y, Lu F, Zhang Y. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell stem cell*. 2014;15(4):459-471.
220. Bagci H, Fisher AG. DNA demethylation in pluripotency and reprogramming: the role of tet proteins and cell division. *Cell stem cell*. 2013;13(3):265-269.
221. Secardin L, Limia CEG, Di Stefano A, et al. TET2 haploinsufficiency alters reprogramming into induced pluripotent stem cells. *Stem cell research*. 2020;44:101755.
222. Pronier E, Almire C, Mokrani H, et al. Inhibition of TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine disturbs erythroid and granulomonocytic differentiation of human hematopoietic progenitors. *Blood*. 2011;118(9):2551-2555.

223. Koh KP, Yabuuchi A, Rao S, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell stem cell*. 2011;8(2):200-213.
224. Cimmino L, Dolgalev I, Wang Y, et al. Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression. *Cell*. 2017;170(6):1079-1095.e20.
225. Montagner S, Leoni C, Emming S, et al. TET2 Regulates Mast Cell Differentiation and Proliferation through Catalytic and Non-catalytic Activities. *Cell reports*. 2016;15(7):1566-1579.
226. Calvanese V, Fernández AF, Urdinguio RG, et al. A promoter DNA demethylation landscape of human hematopoietic differentiation. *Nucleic acids research*. 2012;40(1):116-131.
227. Pham T-H, Benner C, Lichtinger M, et al. Dynamic epigenetic enhancer signatures reveal key transcription factors associated with monocytic differentiation states. *Blood*. 2012;119(24):e161-71.
228. Ito T, Nishiyama C, Nakano N, et al. Roles of PU.1 in monocyte- and mast cell-specific gene regulation: PU.1 transactivates CIITA pIV in cooperation with IFN-gamma. *International immunology*. 2009;21(7):803-816.
229. Friedman AD. Transcriptional control of granulocyte and monocyte development. *Oncogene*. 2007;26(47):6816-6828.
230. Miah MA, Byeon SE, Ahmed MS, Yoon C-H, Ha S-J, Bae Y-S. Egr2 induced during DC development acts as an intrinsic negative regulator of DC immunogenicity. *European journal of immunology*. 2013;43(9):2484-2496.
231. Agnarelli A, Chevassut T, Mancini EJ. IRF4 in multiple myeloma-Biology, disease and therapeutic target. *Leukemia research*. 2018;72:52-58.
232. Remesh SG, Santosh V, Escalante CR. Structural Studies of IRF4 Reveal a Flexible Autoinhibitory Region and a Compact Linker Domain. *The Journal of biological chemistry*. 2015;290(46):27779-27790.
233. Savitsky D, Tamura T, Yanai H, Taniguchi T. Regulation of immunity and oncogenesis by the IRF transcription factor family. *Cancer immunology, immunotherapy : CII*. 2010;59(4):489-510.
234. Tamura T, Yanai H, Savitsky D, Taniguchi T. The IRF family transcription factors in immunity and oncogenesis. *Annual review of immunology*. 2008;26:535-584.
235. Nam S, Lim J-S. Essential role of interferon regulatory factor 4 (IRF4) in immune cell development. *Archives of pharmacal research*. 2016;39(11):1548-1555.
236. Pongubala JM, Nagulapalli S, Klemsz MJ, McKercher SR, Maki RA, Atchison ML. PU.1 recruits a second nuclear factor to a site important for immunoglobulin kappa 3' enhancer activity. *Molecular and Cellular Biology*. 1992;12(1):368-378.
237. Nam S, Kang K, Cha JS, et al. Interferon regulatory factor 4 (IRF4) controls myeloid-derived suppressor cell (MDSC) differentiation and function. *Journal of leukocyte biology*. 2016;100(6):1273-1284.
238. Yamamoto M, Kato T, Hotta C, et al. Shared and distinct functions of the transcription factors IRF4 and IRF8 in myeloid cell development. *PLoS ONE*. 2011;6(10):e25812.
239. Heo MH, Park HY, Ko YH, Kim WS, Kim SJ. IRF4/MUM1 expression is associated with poor survival outcomes in patients with peripheral T-cell lymphoma. *Journal of Cancer*. 2017;8(6):1018-1024.
240. Wong RWJ, Tan TK, Amanda S, et al. Feed-forward regulatory loop driven by IRF4 and NF-κB in adult T-cell leukemia/lymphoma. *Blood*. 2020;135(12):934-947.
241. El Chartouni C, Schwarzfischer L, Rehli M. Interleukin-4 induced interferon regulatory factor (Irf) 4 participates in the regulation of alternative macrophage priming. *Immunobiology*. 2010;215(9-10):821-825.
242. Gupta S, Jiang M, Anthony A, Pernis AB. Lineage-specific modulation of interleukin 4 signaling by interferon regulatory factor 4. *The Journal of Experimental Medicine*. 1999;190(12):1837-1848.
243. Honma K, Udono H, Kohno T, et al. Interferon regulatory factor 4 negatively regulates the production of proinflammatory cytokines by macrophages in response to LPS. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102(44):16001-16006.
244. Briseno CG, Haldar M, Kretzer NM, et al. Distinct Transcriptional Programs Control Cross-Priming in Classical and Monocyte-Derived Dendritic Cells. *Cell reports*. 2016;15(11):2462-2474.
245. Vincent MS, Leslie DS, Gumperz JE, Xiong X, Grant EP, Brenner MB. CD1-dependent dendritic cell instruction. *Nature Immunology*. 2002;3(12):1163-1168.
246. Lim TS, Goh JKH, Mortellaro A, Lim CT, Hämmerling GJ, Ricciardi-Castagnoli P. CD80 and CD86 differentially regulate mechanical interactions of T-cells with antigen-presenting dendritic cells and B-cells. *PLoS ONE*. 2012;7(9):e45185.
247. Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of human genetics*. 2009;54(1):15-39.
248. Schefold JC, Porz L, Uebe B, et al. Diminished HLA-DR expression on monocyte and dendritic cell subsets indicating impairment of cellular immunity in pre-term neonates: a prospective observational analysis. *Journal of perinatal medicine*. 2015;43(5):609-618.
249. Kushchayeva Y, Mishchuk D, Ugarova T. The Role of Beta 2 Integrins in Macrophage Migration During Resolution of Inflammation. *Blood*. 2009;114(22):3600.
250. Tagashira A, Nishi K, Matsumoto S, Sugahara T. Anti-inflammatory effect of lysozyme from hen egg white on mouse peritoneal macrophages. *Cytotechnology*. 2018;70(3):929-938.

251. Underhill DM, Shimada T, Park BG, et al. Macrophage phagocytosis and lysozyme-based digestion of peptidoglycan are required for IL-1β production in response to S. aureus (134.38). *The Journal of Immunology*. 2009;182(1 Supplement):134.38.
252. Antonczyk A, Krist B, Sajek M, et al. Direct Inhibition of IRF-Dependent Transcriptional Regulatory Mechanisms Associated With Disease. *Frontiers in immunology*. 2019;10:1176.
253. Escalante CR, Shen L, Escalante MC, et al. Crystallization and characterization of PU.1/IRF-4/DNA ternary complex. *Journal of Structural Biology*. 2002;139(1):55-59.
254. Lloberas J, Soler C, Celada A. The key role of PU.1/SPI-1 in B cells, myeloid cells and macrophages. *Immunology Today*. 1999;20(4):184-189.
255. Yashiro T, Yamaguchi M, Watanuki Y, Kasakura K, Nishiyama C. The Transcription Factors PU.1 and IRF4 Determine Dendritic Cell-Specific Expression of RALDH2. *Journal of immunology (Baltimore, Md. : 1950)*. 2018;201(12):3677-3682.
256. Pang SHM, Minnich M, Gangatirkar P, et al. PU.1 cooperates with IRF4 and IRF8 to suppress pre-B-cell leukemia. *Leukemia*. 2016;30(6):1375-1387.
257. Li P, Leonard WJ. Chromatin Accessibility and Interactions in the Transcriptional Regulation of T Cells. *Frontiers in immunology*. 2018;9:2738.
258. Gabet Y, Baniwal SK, Leclerc N, et al. Krox20/EGR2 deficiency accelerates cell growth and differentiation in the monocytic lineage and decreases bone mass. *Blood*. 2010;116(19):3964-3971.
259. Taefehshokr S, Key YA, Khakpour M, Dadebighlu P, Oveisi A. Early growth response 2 and Egr3 are unique regulators in immune system. *Central-European journal of immunology*. 2017;42(2):205-209.
260. Chavrier P, Zerial M, Lemaire P, Almendral J, Bravo R, Charnay P. A gene encoding a protein with zinc fingers is activated during G0/G1 transition in cultured cells. *The EMBO journal*. 1988;7(1):29-35.
261. Swirnoff AH, Milbrandt J. DNA-binding specificity of NGFI-A and related zinc finger transcription factors. *Molecular and Cellular Biology*. 1995;15(4):2275-2287.
262. Luciano RL, Wilson AC. HCF-1 functions as a coactivator for the zinc finger protein Krox20. *The Journal of biological chemistry*. 2003;278(51):51116-51124.
263. Russo MW, Sevetson BR, Milbrandt J. Identification of NAB1, a repressor of NGFI-A- and Krox20-mediated transcription. *Proceedings of the National Academy of Sciences of the United States of America*. 1995;92(15):6873-6877.
264. Svaren J, Sevetson BR, Apel ED, Zimonjic DB, Popescu NC, Milbrandt J. NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by proliferative and differentiative stimuli. *Molecular and Cellular Biology*. 1996;16(7):3545-3553.
265. Leblanc SE, Srinivasan R, Ferri C, et al. Regulation of cholesterol/lipid biosynthetic genes by Egr2/Krox20 during peripheral nerve myelination. *Journal of neurochemistry*. 2005;93(3):737-748.
266. Poirier R, Cheval H, Mailhes C, et al. Distinct functions of egr gene family members in cognitive processes. *Frontiers in neuroscience*. 2008;2(1):47-55.
267. Crist SA, Elzey BD, Ahmann MT, Ratliff TL. Early growth response-1 (EGR-1) and nuclear factor of activated T cells (NFAT) cooperate to mediate CD40L expression in megakaryocytes and platelets. *The Journal of biological chemistry*. 2013;288(47):33985-33996.
268. Wieland GD, Nehmann N, Müller D, et al. Early growth response proteins EGR-4 and EGR-3 interact with immune inflammatory mediators NF-kappaB p50 and p65. *Journal of cell science*. 2005;118(Pt 14):3203-3212.
269. Gururajan M, Simmons A, Dasu T, et al. Early growth response genes regulate B cell development, proliferation, and immune response. *Journal of immunology (Baltimore, Md. : 1950)*. 2008;181(7):4590-4602.
270. Li S, Miao T, Sebastian M, et al. The transcription factors Egr2 and Egr3 are essential for the control of inflammation and antigen-induced proliferation of B and T cells. *Immunity*. 2012;37(4):685-696.
271. Miao T, Symonds ALJ, Singh R, et al. Egr2 and 3 control adaptive immune responses by temporally uncoupling expansion from T cell differentiation. *The Journal of Experimental Medicine*. 2017;214(6):1787-1808.
272. Laslo P, Spooner CJ, Warmflash A, et al. Multilineage Transcriptional Priming and Determination of Alternate Hematopoietic Cell Fates. *Cell*. 2006;126(4):755-766.
273. Barbieri E, Trizzino M, Welsh SA, et al. Targeted Enhancer Activation by a Subunit of the Integrator Complex. *Molecular cell*. 2018;71(1):103-116.e7.
274. Du N, Kwon H, Li P, et al. EGR2 is critical for peripheral naïve T-cell differentiation and the T-cell response to influenza. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(46):16484-16489.
275. Zheng Y, Zha Y, Driessens G, Locke F, Gajewski TF. Transcriptional regulator early growth response gene 2 (Egr2) is required for T cell anergy in vitro and in vivo. *The Journal of Experimental Medicine*. 2012;209(12):2157-2163.
276. Li S, Symonds ALJ, Zhu B, et al. Early growth response gene-2 (Egr-2) regulates the development of B and T cells. *PLoS ONE*. 2011;6(4):e18498.
277. Li K, Wu Y, Li Y, et al. *Landscape and dynamics of the transcriptional regulatory network during natural killer cell differentiation*; 2019; 395.
278. Comrie WA, Li S, Boyle S, Burkhardt JK. The dendritic cell cytoskeleton promotes T cell adhesion and activation by constraining ICAM-1 mobility. *The Journal of cell biology*. 2015;208(4):457-473.

279. Malinova D, Fritzsche M, Nowosad CR, et al. WASp-dependent actin cytoskeleton stability at the dendritic cell immunological synapse is required for extensive, functional T cell contacts. *Journal of leukocyte biology*. 2016;99(5):699-710.
280. Al-Alwan MM, Rowden G, Lee TD, West KA. The dendritic cell cytoskeleton is critical for the formation of the immunological synapse. *Journal of immunology (Baltimore, Md. : 1950)*. 2001;166(3):1452-1456.
281. Baranov MV, Revelo NH, Dingjan I, et al. SWAP70 Organizes the Actin Cytoskeleton and Is Essential for Phagocytosis. *Cell reports*. 2016;17(6):1518-1531.
282. Frittoli E, Matteoli G, Palamidessi A, et al. The signaling adaptor Eps8 is an essential actin capping protein for dendritic cell migration. *Immunity*. 2011;35(3):388-399.
283. Vargas P, Maiuri P, Bretou M, et al. Innate control of actin nucleation determines two distinct migration behaviours in dendritic cells. *Nature cell biology*. 2016;18(1):43-53.
284. Saita D, Ferrarese R, Foglieni C, et al. Adaptive immunity against gut microbiota enhances apoE-mediated immune regulation and reduces atherosclerosis and western-diet-related inflammation. *Scientific reports*. 2016;6:29353.
285. Hartman ZC, Kiang A, Everett RS, et al. Adenovirus infection triggers a rapid, MyD88-regulated transcriptome response critical to acute-phase and adaptive immune responses in vivo. *Journal of virology*. 2007;81(4):1796-1812.
286. Averill MM, Barnhart S, Becker L, et al. S100A9 differentially modifies phenotypic states of neutrophils, macrophages, and dendritic cells: implications for atherosclerosis and adipose tissue inflammation. *Circulation*. 2011;123(11):1216-1226.
287. Mayran A, Drouin J. Pioneer transcription factors shape the epigenetic landscape. *The Journal of biological chemistry*. 2018;293(36):13795-13804.
288. Li H, Ta N, Long C, et al. The spatial binding model of the pioneer factor Oct4 with its target genes during cell reprogramming. *Computational and structural biotechnology journal*. 2019;17:1226-1233.
289. Mann IK, Chatterjee R, Zhao J, et al. CG methylated microarrays identify a novel methylated sequence bound by the CEBPB|ATF4 heterodimer that is active in vivo. *Genome research*. 2013;23(6):988-997.
290. Liu Y, Olanrewaju YO, Zheng Y, et al. Structural basis for Klf4 recognition of methylated DNA. *Nucleic acids research*. 2014;42(8):4859-4867.
291. Radtke F, Hug M, Georgiev O, Matsuo K, Schaffner W. Differential sensitivity of zinc finger transcription factors MTF-1, Sp1 and Krox-20 to CpG methylation of their binding sites. *Biological chemistry Hoppe-Seyler*. 1996;377(1):47-56.
292. Mager GM, Ward RM, Srinivasan R, Jang S-W, Wrabetz L, Svaren J. Active gene repression by the Egr2.NAB complex during peripheral nerve myelination. *The Journal of biological chemistry*. 2008;283(26):18187-18197.
293. Srinivasan R, Jang S-W, Ward RM, Sachdev S, Ezashi T, Svaren J. Differential regulation of NAB corepressor genes in Schwann cells. *BMC molecular biology*. 2007;8:117.
294. Le N, Nagarajan R, Wang JYT, et al. Nab proteins are essential for peripheral nervous system myelination. *Nature neuroscience*. 2005;8(7):932-940.
295. Pradhan SK, Su T, Yen L, et al. EP400 Deposits H3.3 into Promoters and Enhancers during Gene Activation. *Molecular cell*. 2016;61(1):27-38.
296. Guerrero-Martínez JA, Reyes JC. High expression of SMARCA4 or SMARCA2 is frequently associated with an opposite prognosis in cancer. *Scientific reports*. 2018;8(1):2043.
297. Li XS, Trojer P, Matsumura T, Treisman JE, Tanese N. Mammalian SWI/SNF--a subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B. *Molecular and Cellular Biology*. 2010;30(7):1673-1688.
298. Bosselut R. Pleiotropic Functions of H3K27Me3 Demethylases in Immune Cell Differentiation. *Trends in immunology*. 2016;37(2):102-113.
299. Bedke N, Swindle EJ, Molnar C, et al. A method for the generation of large numbers of dendritic cells from CD34+ hematopoietic stem cells from cord blood. *Journal of immunological methods*. 2020;477:112703.
300. Shen L, Wu H, Diep D, et al. Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell*. 2013;153(3):692-706.
301. Weber AR, Krawczyk C, Robertson AB, et al. Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nature Communications*. 2016;7:10806.
302. Slyvka A, Mierzejewska K, Bochtler M. Nei-like 1 (NEIL1) excises 5-carboxylcytosine directly and stimulates TDG-mediated 5-formyl and 5-carboxylcytosine excision. *Scientific reports*. 2017;7(1):9001.
303. Li Z, Gu T-P, Weber AR, et al. Gadd45a promotes DNA demethylation through TDG. *Nucleic acids research*. 2015;43(8):3986-3997.
304. Niehrs C, Schäfer A. Active DNA demethylation by Gadd45 and DNA repair. *Trends in cell biology*. 2012;22(4):220-227.
305. Ballester B, Medina-Rivera A, Schmidt D, et al. Multi-species, multi-transcription factor binding highlights conserved control of tissue-specific biological pathways. *eLife*. 2014;3:e02626.
306. Corces MR, Buenrostro JD, Wu B, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nature Genetics*. 2016;48(10):1193-1203.

10 Appendix

The following tables list all published high-throughput-sequencing data used in this thesis.

Table 10.1 Published ChIP-sequencing data

Cell type	Sample	IP	Accession number	Total reads[1]	FRIP (%)[2]	Peaks[3]	Reference
MO	freshly isolated	CEBPB ChIP	SRR333633	5489459	4.45	13275	[227]
MAC	7d culture in 2% AB serum	CEBPB ChIP	SRR333649, SRR333650, SRR333651, SRR333652	14793649	14.01	67734	[227]
MO/MAC	7d IL4/GM-CSF	IgG control	SRR333634, SRR333635, SRR333636	13754210	–	–	[227]
Liver		CEBPA ChIP	ERR235748, ERR235729, ERR235723, ERR235766	37071639	12.76	72762	[305]
Liver		Input	ERR235788, ERR235759	30055419	–	–	[305]

[1]Unique reads after mapping to human reference genome GRCh38

[2]Fraction of reads in peaks (FRIP), determined by running HOMER's findPeaks program in ''factor'' mode using default parameters and the matching background (input)

[3]Number of peaks (determined by HOMER's findPeaks program in ''factor'' mode using default parameters and the matching background (input))

Table 10.2 Published ATAC-sequencing data

Cell type	Accession number	Total read pairs	FRIP (%)	Peaks	Reference
MO	SRR2920543	63854968	33.86	63647	[306]
MO	SRR2920487	21181211	21.25	24324	[306]
MO	SRR2920475	9713700	29.68	36278	[306]
iDC	SRR1725732	204024037	41.00	285569	[207]

Table 10.3 Published WGBS-sequencing data

Cell type	Accession number	Data	CpGs	Reference
iDC	GSM1565940	methylation levels for each covered CpG (hg19) converted to bedGraph	24617968	[207]
iDC	GSM1565942	methylation levels for each covered CpG (hg19) converted to bedGraph	23551701	[207]
iDC	GSM1565944	methylation levels for each covered CpG (hg19) converted to bedGraph	24445760	[207]
iDC	GSM1565946	methylation levels for each covered CpG (hg19) converted to bedGraph	24479420	[207]
iDC	GSM1565948	methylation levels for each covered CpG (hg19) converted to bedGraph	24427703	[207]
iDC	GSM1565950	methylation levels for each covered CpG (hg19) converted to bedGraph	24560919	[207]
CD14+CD16– MO from venous blood	EGAX00001086967	methylation levels for each covered CpG (hg19), bigWig converted to bedGraph	22667869	http://dcc.blueprint-epigenome.eu/
CD14+CD16– MO from venous blood	EGAX00001086968	methylation levels for each covered CpG (hg19), bigWig converted to bedGraph	23768442	http://dcc.blueprint-epigenome.eu/
CD14+CD16– MO from venous blood	EGAX00001086970	methylation levels for each covered CpG (hg19), bigWig converted to bedGraph	23652501	http://dcc.blueprint-epigenome.eu/
CD14+CD16– MO from venous blood	EGAX00001097774	methylation levels for each covered CpG (hg19), bigWig converted to bedGraph	24145754	http://dcc.blueprint-epigenome.eu/
CD38- B cells	ERS214672	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	24601631	https://epigenomesportal.ca
CD38- B cells	ERS214675	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	24932472	https://epigenomesportal.ca
CD38- B cells	ERS222266	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	23878548	https://epigenomesportal.ca
CD38- B cells	ERS523625	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	23633379	https://epigenomesportal.ca
CD8+ T cells	ERS214674	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	25029123	https://epigenomesportal.ca
CD8+ T cells	ERS222241	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	24488031	https://epigenomesportal.ca
CD8+ T cells	ERS317233	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	25482590	https://epigenomesportal.ca
CD8+ T cells	ERS433791	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	22165393	https://epigenomesportal.ca
NK cells	ERS222243	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	24729720	https://epigenomesportal.ca
NK cells	ERS317235	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	25327131	https://epigenomesportal.ca
NK cells	ERS763560	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	26961816	https://epigenomesportal.ca
Neutrophils	ERS208313	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	26244136	https://epigenomesportal.ca
Neutrophils	ERS227748	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	24960155	https://epigenomesportal.ca
Neutrophils	ERS227749	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	26025313	https://epigenomesportal.ca
Neutrophils	ERS661057	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	25195761	https://epigenomesportal.ca
Neutrophils	ERS661059	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	23925309	https://epigenomesportal.ca
Neutrophils	ERS661060	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	24906321	https://epigenomesportal.ca
Hepatocytes	IHECRE00001879.2	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	21708615	https://epigenomesportal.ca
Hepatocytes	IHECRE00001876.2	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	22059276	https://epigenomesportal.ca
Hepatocytes	IHECRE00001878.2	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	21382628	https://epigenomesportal.ca
Hepatocytes	IHECRE00001877.2	methylation levels for each covered CpG (hg38), bigWig converted to bedGraph	22177806	https://epigenomesportal.ca

11 Publications

Mendes K, Schmidhofer S, Minderjahn J, Glatz D, Kiesewetter C, Raithel J, Wimmer J, Gebhard C, Rehli M (2020). The epigenetic pioneer EGR2 initiates DNA demethylation in differentiating monocytes at both stable and transient binding sites. *Submitted*

Minderjahn J, Schmidt A, Fuchs A, Schill R, Raithel J, Babina M, Schmidl C, Gebhard C, Schmidhofer S, **Mendes K**, Ratermann A, Glatz D, Nützel M, Edinger M, Hoffmann P, Spang R, Längst G, Imhof A, Rehli M (2020). Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat Commun* 11, 402.

Srivastava JK, Awatade NT, Bhat HR, Kmit A, **Mendes K**, Ramos M, Amaral MD, Singh UP (2015). Pharmacological evaluation of hybrid thiazolidin-4-one-1,3,5-triazines for NF-κB, biofilm and CFTR activity. *RSC Advances* 5, 88710-88718

Awatade NT, Uliyakina I, Farinha CM, Clarke LA, **Mendes K**, Solé A, Pastor J, Ramos M, Amaral MD (2015). Measurements of Functional Responses in Human Primary Lung Cells as a Basis for Personalized Therapy for Cystic Fibrosis. *EBioMedicine* 2, 147–153

Acknowledgment

I would like to thank all people that in one way or another contributed to the work presented in this PhD thesis.

First and foremost, I would like to express my sincere gratitude to my supervisor, **Prof. Dr. Michael Rehli**, who kindly accepted me into his research group and guided me throughout this exciting project. His outstanding knowledge, support, encouragement and availability for data discussion, were fundamental to keep me motivated and complete my thesis.

To my mentors, **Prof. Dr. Gernot Längst** and **Prof. Dr. Sven Heinz**, I would like to take them for their valuable input and data discussion during my progress reports.

I would like to thank **Julia W.** for her help and guidance at the initial steps of this thesis and to **Sandra S.** for the generation of all the hydroxymethylation data presented in this thesis.

To **Claudia K.**, **Julia M.** and **Johanna**, I would like to thank them for their support in the experimental procedures and all the troubleshooting tips. Also thank you for all the great moments.

I would like to thank **Claudia G**. for her availability to read this thesis and to **Jan** for his help in translating the summary.

I also wish to thank **AG Rehli** (former and current members) as well as **AG Kreutz** and **AG Thomas** for accompanying me in the everyday lab work and for helping me when I needed. Also thank you for the parties and time we spent together.

A special thanks to **Carina** and **Sakhila**, for always listening to me and encouraging me to move forward. Also many thanks to **Sara** and **Tobi** for their friendship and help in translating the summary. Obrigada **Rui** e **Carina** pelas vossas sugestões e tempo despendido para ler esta tese. A vós e restantes membros do “nosso grupo de portugueses”, obrigada por todos os convívios e boa disposição ao longo deste percurso.

Um agradecimento especial a ti, **Flávio**. Obrigada por te teres mudado para perto de mim apesar de todos os desafios. Por todo o teu apoio, paciência e ajuda. Por tornares as coisas mais fáceis e me fazeres sentir em casa.

Por último, um muito obrigada aos **meus pais** e aos **meus irmãos**. Pelo vosso apoio, positivismo e por sempre acreditarem em mim. Por serem o meu refúgio, onde me sinto tão segura. Aos **meus pais**, a quem dedico esta tese, obrigada por todas as oportunidades e amor incondicional.