Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the

Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M. Sc. Laura Wiehle

born in Waren (Müritz), Germany

Oral examination: 5th October 2017

TET-dependent DNA methylation patterns in mammalian development and disease

Referees:

Prof. Dr. Frank Lyko

Prof. Dr. Jan Lohmann

Abstract

TET enzymes are relatively novel players in the epigenetic regulation of mammalian DNA methylation. They participate in DNA demethylation, but their precise roles in different developmental and disease scenarios are not fully understood. The aim of this work was to investigate the biological roles of TET enzymes in lineage-committed normal and cancer cells. To this end, murine primary cells with genetic deletion of TET enzymes and human cancer cells with recurrent mutations in the cofactor providing isocitrate dehydrogenases (IDH), provoking competitive inhibition of TET enzymes, were analyzed.

By characterizing mouse embryonic fibroblasts adipogenic differentiation defects, inefficient activation of genes relevant to adipogenesis and widespread gene deregulation upon *TET1/2*-deficiency were discovered. Examination of the genome-wide DNA methylation landscape demonstrated the hypermethylation of DNA methylation canyons as a main characteristic of the *TET1/2*-deficient methylome. Canyons were associated with developmentally important genes and canyon collapse due to hypermethylation coincided with developmental gene deregulation, defective induction of adipogenic markers and the hypermethylation of their promoters. Together, these findings uncovered a novel epigenetic regulatory role in the maintenance of DNA methylation canyons for TET1 and TET2 that is essential for epigenetic programming during differentiation.

In the second part of this thesis, published array-based DNA methylation profiles of a large acute myeloid leukemia (AML) patient cohort were used to examine mutant *IDH*-(*mIDH*) and TET-dependent DNA methylation changes. This confirmed the known association between *mIDH* and genome-wide hypermethylation. However, similar global methylation changes were not present in *TET2* mutant patients and *mIDH* carrying patients lacked specific canyon hypermethylation. Intriguingly, neither overexpression of *mIDH*, nor treatment of a leukemia cell line with D-2-hydroxyglutarate, which is a putative TET inhibitor produced by mIDH, recapitulated the *mIDH*-associated hypermethylation. Instead, comparison with hematopoietic reference methylomes revealed high similarity between *mIDH*-associated and myeloid progenitor methylation profiles, suggesting the involvement of differentiation state rather than TET inhibition in the hypermethylation phenotype. These findings implicate a previously unnoted factor in the epigenomic changes of AML cells with *mIDH*, which may be critical to understand and therapeutically target mIDH-dependent pathogenesis.

Zusammenfassung

TET Enzyme sind relativ neue Akteure in der epigenetischen Regulation von DNA-Methylierung in Säugern. Sie wirken an DNA-Demethylierung mit, jedoch sind ihre konkreten Rollen in verschiedenen Differenzierungs- und Krankheitsszenarien nicht vollständig aufgeklärt. In der vorliegenden Arbeit sollten die biologischen Funktionen von TET-Enzymen in determinierten normalen sowie Krebszellen erforscht werden. Dazu wurden primäre Mauszellen mit genetischer *TET*-Deletion und humane Krebszellen mit wiederkehrenden Mutationen in den Cofaktor-bereitstellenden Isocitrat-Dehydrogenasen (*IDH*), welche zu einer kompetitiven Enzymhemmung der TET Proteine führen, untersucht.

Mittels Charakterisierung von embryonalen Mausfibroblasten wurden Defekte in der adipogenen Differenzierung, der Aktivierung von Adipogenese-relevanten Genen und der Genexpression bei *TET1/2*-Defizienz entdeckt. Die Untersuchung der genomweiten DNA-Methylierung identifizierte die Hypermethylierung von sogenannten DNA-Methylierungs-*Canyons* als wesentliches Merkmal des *TET1/2*-defizienten Methyloms. *Canyons* waren mit entwicklungsbiologisch relevanten Genen assoziiert und ihr Zusammenbruch durch Hypermethylierung war begleitet von Regulationsdefekten in Entwicklungsgenen, fehlerhafter Induktion von adipogenen Markern und Promoter-Hypermethylierung. Diese Ergebnisse schreiben TET1 und TET2 eine neuartige regulatorische Rolle in der Erhaltung von *Canyons* zu, die essentiell für die epigenetische Programmierung während der Differenzierung ist.

Im zweiten Teil dieser Arbeit wurden publizierte Array-basierte DNA Methylierungsprofile einer großen akuten myeloischen Leukämie (AML) Patienten-Kohorte verwendet, um Methylierungsveränderungen durch mutante *IDH* (*mIDH*) und *TET* Enzyme zu analysieren. Dies bestätigte die bekannte Assoziation zwischen *mIDH* und genomweiter DNA Hypermethylierung. Allerdings waren ähnliche Veränderungen nicht in *TET*2 mutanten Patienten vorhanden und Patienten mit *mIDH* wiesen keine spezifische *Canyon* Hypermethylierung auf. Interessanterweise wurde die *mIDH*-assoziierte Hypermethylierung auch weder durch die Überexpression von *mIDH*, noch die Behandlung von Leukämiezellen mit D-2-Hydroxyglutarat, welches ein durch mIDH produzierter, mutmaßlicher TET Inhibitor ist, exakt nachgebildet. Stattdessen offenbarte der Vergleich mit hämatopoetischen Referenzmethylomen eine hohe Ähnlichkeit zwischen *mIDH*-assoziierten und myeloischen Progenitorzell-Methylierungsprofilen, was auf eine Beteiligung des Differenzierungsgrades anstelle der TET Inhibition an der Hypermethylierung hinweist. Diese Ergebnisse involvieren einen bisher unbeachteten Faktor in die epigenomischen Veränderungen von AML-Zellen mit *mIDH*, welcher entscheidend für das weitere Verständnis und die gezielte Therapie von mIDH-abhängiger Pathogenese sein könnte.

Contents

List of Abbreviations

5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AML	acute myeloid leukemia
bp	base pairs
CGI	CpG island
ChIP(-seq)	chromatin immunoprecipitation (sequencing)
CIMP	CpG island hypermethylator phenotype
CLP	common lymphoid progenitor
CMML	chronic myelomonocytic leukemia
CMP	common myeloid progenitor
D-2-HG	D-2-hydroxyglutarate
(d)dNTP	(di)deoxynucleotide triphosphate
DKO	double knockout
g	gravitational acceleration constant
GMP	granulocyte macrophage progenitor
HGDH	hydroxyglutarate dehydrogenase
HSC	hematopoietic stem cell
kb	kilo bases
KD	knockdown
KO	knockout
L-2-HG	L-2-hydroxyglutarate
LSK	lineage$^-$ SCA1$^+$ KIT$^+$ cells (includes HSCs and MPPs)
MDS	myelodysplastic syndrome
MEF	mouse embryonic fibroblasts
MEP	megakaryocyte erythrocyte progenitor
(m)ESC	(mouse) embryonic stem cell
(m)IDH	(mutant) isocitrate dehydrogenase
min	minutes
MPP	multipotent progenitor
OD	optical density
PMD	partially methylated domain
qRT-PCR	quantitative real time PCR
R&D	research and development

RNA-seq	RNA sequencing
rpm	rotations per minute
sec	seconds
TCGA	The Cancer Genome Atlas
TET	Ten-eleven translocation enzyme
TSS	transcriptional start site
WGBS	whole genome bisulfite sequencing
WT	wildtype

All gene names in introduction and discussion are italic and in capitals independent of species, while protein names are in regular font. Spelling in the results section is according to the species-specific conventions.

List of Figures

List of Tables

1 Introduction

The genomic sequence of a living creature contains the entire information needed to build the organism. But how can the static genome translate into a variety of specialized cell types and functions as observed in complex beings? In 1942 Waddington described that “between genotype and phenotype […] there lies a whole complex of developmental processes” and termed this phenomenon “epigenetics” (Waddington, 1942). Nowadays, epigenetic research investigates (mitotically and/or meiotically heritable) molecular changes, modulating gene function and eventually the phenotype, independently of alterations in the DNA sequence (Russo, Martienssen and Riggs, 1996). Epigenetic processes establish stable gene expression patterns defining cellular identity, but at the same time are reversible and dynamic, allowing cellular plasticity and differentiation. Several components orchestrate the complex epigenetic interplay, among them covalent modifications of nucleic acids and histones, chromatin remodelers and non-coding RNAs.

1.1 The mammalian DNA methylation machinery

Presently, known covalent modifications of DNA bases include N^6-methyladenine (6mA), N^7-methylguanine, C^5-methylcytosine (5mC) and its oxidized derivatives 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5mC), 5-carboxylcytosine (5caC) as well as 5-hydroxymethyluracil (Achwal et al., 1983; Breiling and Lyko, 2015).

DNA methylation at the carbon-5 position of cytosine was first described in 1948 (Hotchkiss, 1948) and due to intensive studies is now the best characterized epigenetic modification. It is highly conserved and present in all three domains of life. In mammals, it is the most common modification termed the fifth base of the genome and has crucial roles in development, differentiation and disease (Li and Zhang, 2014; Smith and Meissner, 2013). The family of DNA methyltransferases (DNMTs) is responsible for the establishment and maintenance of this mark.

1.1.1 DNA methyltransferases

DNMTs catalyze the transfer of a methyl group from its donor S-adenosylmethionine (SAM) to position five of the cytosine carbon ring (Figure 1.3). While DNMT3A and B are *de novo* methyltransferases responsible for initial establishment of DNA methylation, DNMT1 is the maintenance transferase ensuring faithful propagation of the methylation mark during cell division (Figure 1.1A). Since replication is semi-conservative in mammals, palindromic hemimethylated CpG residues can be recognized by DNMT1 and copied to the daughter

DNA strand, resulting in symmetrically methylated CpGs (Figure 1.1C). This process has an error rate of ~1% per division and thus can ensure the gross maintenance of this mark (Reik, 2007). *DNMT1* is ubiquitously expressed in proliferating cells and most abundant during S phase (Kishikawa et al., 2003). Its targeting to DNA replication forks relies on the physical interaction with PCNA and UHRF1, which is capable of binding hemimethylated CpGs (Bostick et al., 2007; Chuang et al., 1997; Smith and Meissner, 2013). Homozygous deletion of *DNMT1* in mESCs leads to a drastic reduction of global DNA methylation, disruption of imprinted gene regulation, complete X-inactivation and derepression of retrotransposons (Goll and Bestor, 2005).

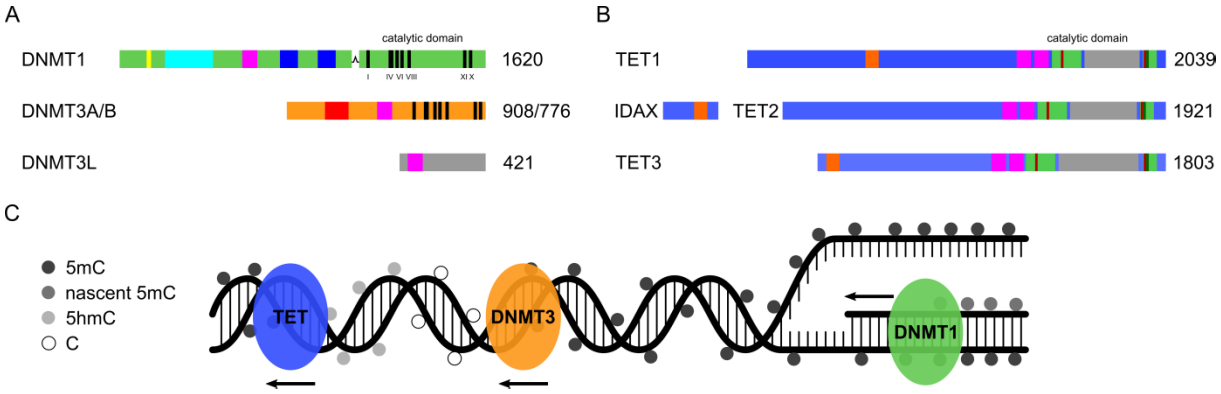

Figure 1.1: Mammalian DNA (de)methylation machinery.
(A) Domain structure of DNA methyltransferases (adapted from Goll and Bestor, 2005). **(B)** Domain structure of the TET family (adapted from Rasmussen and Helin, 2016). Nuclear localization signal (yellow), replication foci (cyan), Cys-rich- (pink), BAH- (blue), PWWP- (red), CXXC- (orange) and DSBH domain (green), methyltransferase motifs (black), low complexity insert (gray), Fe(II)- (dark red) and D-2-HG-interacting motifs (dark green). **(C)** Schematic of DNMT and TET activity on DNA.

De novo methylation activity is recruited to repeat sequences by an unknown mechanism and unmethylated H3K4 harboring regions, such as inactive promoters and methylated sequences, by DNMT3L, which occurs in a complex with DNMT3 enzymes in germ cells (Edwards et al., 2017; Ooi et al., 2007). The disruption of the PIWI-interacting RNA pathway in male germ cells results in abrogation of *de novo* methylation, arguing for an involvement of piRNAs in DNA methylation, but the exact connection is unclear (Aravin and Bourc'his, 2008). Other reports suggest that DNMT3A and B are targeted to active genes by recognition of H3K36me3 through their PWWP domains (Baubec et al., 2015; Rondelet et al., 2016). *DNMT3A/B*-deficient mESCs failed to methylate introduced retrovirus sequences and lost methylation from various repeat elements and genes (Chen et al., 2003; Okano et al., 1999). Knockout of *DNMT3A/B* in human ESCs revealed redundant targeting of DNMT3A or B to intergenic regions, CpG island shores, promoters with intermediate or low CpG density

and introns, while selective targeting happens at CpG islands (Liao et al., 2015). Moreover, DNMT3A seems to play an essential role in the methylation of imprints (Kaneda et al., 2004), whereas DNMT3B is crucial for methylating centromeric minor satellite repeats. *DNMT3A/B* deficient mouse embryos display globally reduced methylation (Okano et al., 1999). These studies indicate that DNMT3A and B have overlapping as well as distinct functions during different developmental stages.

1.1.2 The mammalian DNA methylation landscape

DNA methylation predominantly occurs symmetrically in the sequence context of 5'-CpG-3' (Ramsahoye et al., 2000) and 60 - 80% of these dinucleotides in adult mammalian genomes are usually methylated resulting in a total of ~1% methylated cytosines among all bases (Ehrlich et al., 1982). CpG residues are less frequent than statistically expected in mammalian genomes, presumably because 5mC is prone to spontaneous deamination to thymine, which is not recognized as an erroneous base due to its natural occurrence in DNA. In contrast, unmethylated cytosines are deaminated to uracil and hence would be recognized and repaired by uracil-DNA glycosylase (Coulondre et al., 1978).

This phenomenon most likely resulted in the emergence of CpG islands (CGIs) – clustered regions of high CpG density – which are mostly devoid of methylation at any time point during development in many tissues and thus have been protected from C to T transition. Mammalian methylomes are therefore characterized by a bimodal distribution with the majority of CpGs being highly methylated and the remaining proportion largely unmethylated (Meissner et al., 2008). CGIs comprise regions of ~1000 bp average length with an elevated CG content (>50%) and reduced CpG depletion (observed/expected CpG ratio >60%; Deaton and Bird, 2011). About 72% of annotated promoters are associated with CGIs (Saxonov et al., 2006), suggesting functional relevance of these features. The related genes are often ubiquitously expressed and include nearly all housekeeping genes, many developmental regulators, but also tissue-specific genes with limited expression (Larsen et al., 1992).

Refined analysis of CGIs and their methylation variation in cancer has led to the discovery of less CpG dense shores and shelves surrounding CGIs, which are interspersed into the open sea of single isolated CpGs (Figure 1.2A). These epigenomic elements are highly conserved between species and most tissue-specific and cancer-associated variation appears to occur in CGI shores (Doi et al., 2009; Irizarry et al., 2009).

DNA methylation canyons (Figure 1.2B) have been recently identified as another poorly methylated feature of mammalian genomes (Jeong et al., 2013). In contrast to CGIs, these elements are considerably larger (>3.5 kb; some extending over 25 kb) and 90% of

them harbor at least one CGI. Identification requires at least five CpGs/kb and an average methylation of ≤10%. This identified 1,104 canyons in purified mouse hematopoietic stem cells (HSC). Canyon-associated genes are of particular developmental relevance as they are enriched for homeobox genes, transcriptional regulators and genes involved in embryonic morphogenesis. Furthermore, canyons are characterized by inter-species and cell-type conservation, depletion for transposable elements, 5hmC presence at their borders and dependency on DNMT3A (Jeong et al., 2013).

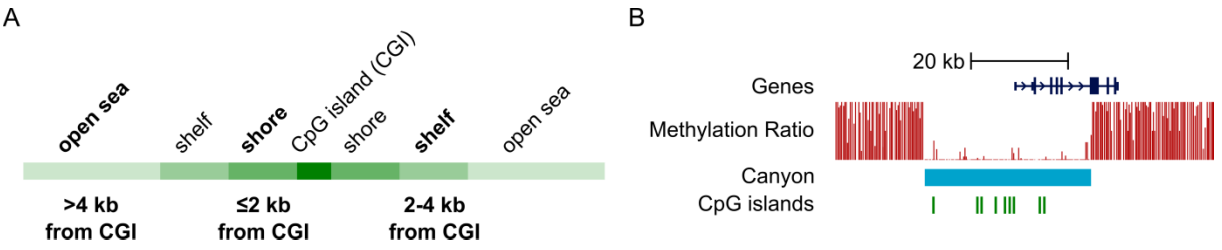

Figure 1.2: Schematic presentation of key features of the mammalian methylome.
(A) Conceptual scheme showing the relationship between CpG island, shore, shelf and open sea. Green color intensity indicates CpG density. **(B)** Schematic example of a DNA methylation canyon at the mouse *PAX6* locus (adapted from Jeong et al., 2013).

Notably, a range of other hypomethylated elements has been identified to date. Some of them are probably identical to canyons such as DNA methylation valleys (Xie et al., 2013), but others are rather small and coincide with DNAse-hypersensitive and transcription factor binding sites such as low-methylated regions (LMRs; Stadler et al., 2011). Partially methylated domains (PMDs) are large regions with an average methylation level of <70% (Lister et al., 2009) and have been reported to coincide with nuclear lamina-associated domains (Berman et al., 2011). It was suggested that cancer cells or tissue cultures, i.e., systems that proliferate quickly, might progressively lose methylation from PMDs due to inefficient maintenance methylation in gene-poor regions (Gaidatzis et al., 2014; Hon et al., 2012; Lister et al., 2011).

1.1.3 Function of DNA methylation

DNA methylation plays a major role in transcription and chromatin structure regulation, X-chromosome inactivation, silencing of repetitive DNA and transposable elements, and genomic imprinting (Bird, 2002). Since numerous studies showed the inhibition of expression upon methylation (Vardimon et al., 1982), it has been widely accepted that promoter DNA methylation is not compatible with gene expression. However, on a genome-wide level this correlation is not always true (e.g. Eckhardt et al., 2006), emphasizing that the mechanistic relationship between methylation and expression has to be further elucidated.

Methylation of CGI promoters during development has been linked to permanent silencing and can be employed to stably repress germline, stem cell and specific genes of other lineages not required anymore in differentiating cells. Consistently, differentiation relies on the presence of DNMTs and genetically deficient ESCs fail to differentiate and regulate gene expression of germ layer und pluripotency markers accordingly (Jackson et al., 2004; Lei et al., 1996; Li et al., 1992; Okano et al., 1999). Typical pluripotency genes whose promoters are silenced by DNA methylation are *OCT4* and *NANOG.* However, repression is a coordinated sequence of silencing events that also employ histone modification and nucleosome remodeling. In most cases, sequence-specific repressors initiate silencing and only later DNA methylation follows to inevitably lock in the repressive chromatin state. Nevertheless, an instructive role for DNA methylation in initiation of gene silencing is still conceivable (Smith and Meissner, 2013).

Mechanistically, DNA methylation can lead to silencing by inhibiting the recruitment of basal transcription machinery and methylation-sensitive transcription factors, as has been described for SP1, CREB, C-MYC or NRF1 (Clark et al., 1997; Domcke et al., 2015; Mancini et al., 1999; Prendergast et al., 1991). However, there are also a number of factors that are attracted by DNA methylation in their consensus sequence (Yin et al., 2017), potentially explaining positive correlations between methylation and transcription. The insulator protein CTCF can block enhancer promoter interactions, change chromatin architecture and attenuate POLII, which then includes exons during splicing (Ong and Corces, 2014) and thus translates its methylation sensitivity into transcriptional regulation (Bell and Felsenfeld, 2000; Narendra et al., 2015; Shukla et al., 2011). The general applicability of this principle is under debate, especially since CTCF also seems to create local regions of low methylation (Domcke et al., 2015; Feldmann et al., 2013; Stadler et al., 2011).

Furthermore, methylation could affect nucleosome positioning, eventually re-directing the transcriptional machinery and can induce the formation of heterochromatin by recruiting methyl-CpG-binding domain (MBD) proteins such as MeCP2. In turn, histone deacetylases or repressive complexes, such as the Sin3A, Mi2/NuRD or MeCP1 complex, are recruited, which condense the chromatin and facilitate the formation of a repressive chromatin state (Bird and Wolffe, 1999).

In contrast to promoter methylation, gene body methylation in mammals has been correlated with increased transcription of the corresponding genes (Baubec et al., 2015; Hellman and Chess, 2007; Lister et al., 2009). It has been suggested that it is essential to block transcription of active transposable elements while maintaining the elongation of the proper gene (Jones, 2012; Yoder et al., 1997), reduce transcriptional noise (Bird, 1995), regulate cell type specific alternative promoters (Maunakea et al., 2010) or prevent erroneous initiation of cryptic transcription (Neri et al., 2017).

While mammalian genomes are usually highly methylated, two striking waves of global DNA methylation reprogramming are observed during early development. The first takes place during germline development and the second in the early zygote. Primordial germ cell (PGC) demethylation is thought to enable the establishment of gender specific methylation patterns at imprinted genes, reset the cells to pluripotency and reduce mutational burden due to 5mC deamination (Reik et al., 2001). Zygotic demethylation is probably required to erase acquired epigenetic changes during gametogenesis due to environmental or individual genetic events and facilitate the expression of pluripotency-associated genes (Reik, 2007; Reik et al., 2001). The involvement of active and passive demethylation processes in these events will be introduced in chapter 1.2.3.1.

1.1.4 DNA methylation changes in cancer

Epigenetic dysregulation is an emerging hallmark of cancer and comprises alterations in DNA and histone modifications, histone modifiers and readers, chromatin remodelers, and microRNAs (Baylin and Jones, 2016). Together with genetic defects, it participates in tumor promotion and initiation and therefore genetic changes of epigenetic modifiers are now becoming increasingly investigated. The earliest studies on cancer epigenetics noted that 5mC content in different cancers was globally reduced (Feinberg and Vogelstein, 1983a; Gama-Sosa et al., 1983). Later on, it was realized that cancer epigenomes redistribute their genomic methylation resulting in global hypomethylation and focused CGI hypermethylation. Deregulation of oncogenes, tumor suppressor, stem cell and differentiation genes, chromatin architecture, transposable elements, satellite DNA and genomic imprinting are the rational consequences, which could be observed in cancer to date (Dawson, 2017; Feinberg and Tycko, 2004).

Global hypomethylation can affect normally methylated CGIs and lead to derepression of associated genes (Feinberg and Tycko, 2004). These can be potential oncogenes, such as *RAS* or *CCND2* (Feinberg and Vogelstein, 1983b; Oshimo et al., 2003), cancer/testis antigens such as *MAGE-1* (De Smet et al., 1996) or even viral oncogenes among them *HPV16* (Badal et al., 2003; de Capoa et al.). Another prominent example of promoter hypomethylation correlating with increased expression is the *MDR1* gene, which is associated with increased drug resistance in AML (Nakayama et al., 1998). Hypomethylation can induce genomic instability, another hallmark of cancer (Hanahan and Weinberg, 2011). Mechanistically, affected satellites tend to break and recombine, resulting in chromosomal fusions of pericentromeric regions (Qu et al., 1999). Extensive transposition due to derepression of transposable elements might additionally facilitate DNA breaks.

In contrast, hypermethylation of CGIs has been linked to the silencing of tumor suppressor genes. Classical examples that acquire promoter hypermethylation and decline in expression in cancers are *RB, CDKN2A, VHL* and *E-Cadherin* (Graff et al., 1995; Herman et al., 1994; Merlo et al., 1995; Sakai et al., 1991). This evolved into the concept of epigenetic caretakers – genes that usually prevent unlimited self-renewal and are induced upon differentiation but cannot be activated anymore in cancer cells and thus enable clonal expansion (Baylin and Jones, 2016; Feinberg and Tycko, 2004). Similarly, it has been observed that poised genes, which are bivalently marked (H3K4me3 and H3K27me3) in embryonic stem cells, have a ~12-fold increased propensity to undergo promoter hypermethylation in cancer (Dawson, 2017; Widschwendter et al., 2007). Moreover, silencing of genes involved in DNA mismatch repair, for example *MLH1*, has been correlated with the presence of mismatch repair defective cancers (Kane et al., 1997) and might facilitate mutator phenotypes and microsatellite instability.

The hypermethylation of a distinctive set of CGIs in a subset of colorectal cancers was identified in 1999 and termed the CpG island hypermethylator phenotype (CIMP; Toyota et al., 1999). It was correlated with mismatch repair deficiency and *BRAF* mutations in colon cancer (Weisenberger et al., 2006) and many CIMP loci are targets of embryonic Polycomb group proteins (Widschwendter et al., 2007). However, the concept and underlying mechanisms of this coordinated acquisition of CGI hypermethylation are still controversial (Yamashita et al., 2003). Nevertheless, it has been extended to other cancer entities meanwhile, among them glioma and AML (Kelly et al., 2017; Noushmehr et al., 2010).

Additionally, disturbed methylation in cancer hampers CTCF binding, which appears to result in defective insulator function and chromatin architecture (Flavahan et al., 2015).

These observations implicate DNA methylation as a contributing, if not causative factor in the deregulation of gene expression and genomic integrity in cancer. It is thus not surprising that the epigenetic modifiers *DNMT3A* and *TET2* are among the most frequently mutated genes in clonally expanded stem cell populations of the hematopoietic system (Dawson, 2017). Even so, it is still an unresolved issue whether a driving role in pathogenesis is attributable to aberrant DNA methylation or it is merely a reflection of cellular history and differentiation state.

1.2 The mammalian DNA demethylation machinery

The unique maintenance activity of DNMT1 implies not only stable inheritance of 5mC, but also its theoretically indefinite propagation. Demethylation would thus have to rely on the tight regulation of DNMT1 or its targeting factors and dilution of 5mC during cell divisions, removal of methylated cytosines or deaminated products by repair mechanisms or direct enzymatic reversal. The discovery of 5hmC in mammalian DNA opened up a new avenue for passive

and active DNA demethylation since hydroxymethylated CpGs are poorly recognized by DNMT1 (Hashimoto et al., 2012).

1.2.1 DNA hydroxymethylation

DNA hydroxymethylation has first been described in bacteriophage nucleic acids (Wyatt and Cohen, 1952) and in the 1970s in vertebrate DNA (Penn et al., 1972). In 2009 it was demonstrated that 5hmC is present in mouse brain (Kriaucionis and Heintz, 2009) and ESCs and that the ten-eleven translocation 1 (TET1) enzyme is responsible for the conversion of 5mC to 5hmC (Tahiliani et al., 2009). It was also speculated that this conversion might represent part of an active DNA demethylation pathway (Tahiliani et al., 2009).

The 5hmC mark is highly conserved among animals ranging from mammals to chicken and sea urchin and is also present in plants, albeit levels are quite variable between these species (Terragni et al., 2012). Overall levels of 5hmC are variable between tissues and cell types. In most mouse and human tissues, including mESCs and several human cell lines, the total 5hmC percentage in the genome ranges from roughly 0.1 to 0.3. The highest levels were observed in neuronal tissues, i.e., human and mouse adult brains, with 0.48% and 0.9%, respectively, resulting in 1 – 2% 5hmC among all cytosines (Terragni et al., 2012). It was claimed that in Purkinje cells, 5hmC reached almost 40% of 5mC levels (Kriaucionis and Heintz, 2009), suggesting that post-mitotic cells accumulate particularly high 5hmC levels. The content of 5hmC does not correlate with 5mC levels or TET expression and declines in cultured cells (Nestor et al., 2012). It was further noted that 5hmC levels were decreased in different cancer tissues and cell lines (Globisch et al., 2010; Kriaucionis and Heintz, 2009; Terragni et al., 2012).

Since its discovery, 5hmC has been implicated in DNA demethylation and epigenetic signaling. The evidence for active DNA demethylation originates from *in vitro* and cell culture studies describing TET-mediated 5mC oxidation followed by excision of 5fC and 5caC employing thymine DNA glycosylase (TDG) and base excision repair (He et al., 2011; Maiti and Drohat, 2011). Consistently, 5caC and 5fC accumulate to detectable levels in TDG-deficient cells (He et al., 2011; Shen et al., 2013). Evidence for active DNA demethylation *in vivo* comes from reports showing the accumulation of oxidized mC forms, the activation of base excision repair and enrichment of single-strand breaks during zygotic and germ cell DNA methylation reprogramming (Hajkova et al., 2010; Ladstätter and Tachibana-Konwalski, 2016; Wossidlo et al., 2010). Passive demethylation by dilution has been supported by strongly reduced DNMT1 activity on hemimodified 5hmC, 5fC and 5caC compared to 5mC (Hashimoto et al., 2012) and has for example been observed in zygotic methylation erasure (Inoue and Zhang, 2011).

While in DNA demethylation 5hmC would be a short-lived intermediate, epigenetic signaling would require reasonable turn over times of 5hmC. Analysis of steady states of 5hmC, 5fC and 5caC suggested that 5hmC is slowly established on the replicated strand and stably maintained over several days in cultured cells and *in vivo* (Bachman et al., 2014). Global 5fC and 5caC content was not detectable in mouse olfactory bulb DNA using HPLC-MS (Globisch et al., 2010), indicating that the levels of conversion products of 5hmC are negligible (0.3 to 1.4% of 5hmC) and thus are either very short-lived or cannot be responsible for large-scale DNA demethylation. The kinetics of TET processivity also argue against a systematic, iterative and thus demethylating oxidation process (Tamanaha et al., 2016). Base resolution analysis of 5hmC, 5fC and 5caC using TDG-deficient mESCs suggested that distinct regions of the genome carry different oxidized 5mC forms and that TET processivity correlates with local chromatin accessibility, suggesting targeted conversion at functionally relevant regions (Shen et al., 2013; Wu and Zhang, 2017; Wu et al., 2014). Consistent with a signaling role of oxidized 5mC forms, specifically interacting reader proteins, such as MeCP2, MBD3, Uhrf2, CTCF, SALL4A and RNA polymerase II elongation complex, which are involved in transcriptional regulation, have been discovered (Iurlaro et al., 2013; Marina et al., 2016; Mellén et al., 2012; Spruijt et al., 2013; Wang et al., 2015a; Xiong et al., 2016; Yildirim et al., 2011). It was also suggested that 5hmC could function in marking DNA damage sites (Kafer et al., 2016).

To elucidate the genomic location of 5hmC conversion and demethylation, numerous studies mapped the 5hmC distribution by various methods. Whole genome bisulfite sequencing, which is the gold standard for base resolution 5mC mapping, has been used to monitor regions of demethylation in TET-deficient models – assuming that 5mC accumulates in regions where it would usually be removed by TETs. However, it can neither distinguish 5hmC from 5mC nor 5fC and 5caC from unmodified C (Booth et al., 2012; He et al., 2011; Yu et al., 2012). Advanced technologies have thus been developed to directly map 5hmC, such as 5hmC-DNA-IP, oxidative or TET-assisted bisulfite sequencing.

Firstly, 5hmC is widespread in the mammalian genome occurring at inter- and intragenic, DNAse hypersensitive regions, TSS, CTCF and p300 sites. In ESCs, 5hmC deposition seems to correlate negatively with CpG density, so that promoters with low or intermediate CpG density are enriched, while unmethylated, H3K4me3-marked CGI promoters of highly expressed genes are depleted of 5hmC (Booth et al., 2012; Pastor et al., 2011; Wu et al., 2011a; Xu et al., 2011b; Yu et al., 2012). Elevated levels of 5hmC were also found at bivalent promoters of poised genes, enhancers and gene bodies (Hon et al., 2014; Stroud et al., 2011; Sun et al., 2015; Williams et al., 2011). Immediate transcription factor binding sites are depleted of 5hmC, but flanking areas are enriched (Yu et al., 2012). TET1 and TET2 activity have been assigned to TSS and gene bodies, respectively (Huang et al.,

2014). Together, these correlations suggest relevance of 5mC oxidation at various genomic functional elements either by demethylation or conferring specific epigenetic signals.

1.2.2 Ten-eleven translocation enzymes

The mammalian TET enzyme family comprises three members – TET1, TET2 and TET3 – all of which can successively convert 5mC to 5hmC, 5fC and 5caC (Ito et al., 2010, 2011; Tahiliani et al., 2009). TET proteins require α-ketoglutarate, oxygen and iron (Fe(II)) as cofactors to exert their dioxygenase activity (Figure 1.3). According to biochemical studies, TET enzymes display the highest affinity for mCpG compared to other neighboring bases (Hu et al., 2013) and 5hmC consistently occurs almost exclusively in a CpG context in mammalian genomes (Yu et al., 2012). Moreover, TET proteins prefer binding to 5mC over 5hmC and 5fC (Hu et al., 2015; Ito et al., 2011) and display similar conversion rates independently of different modification combinations or hemimodification of a CpG dyad (Wu and Zhang, 2017).

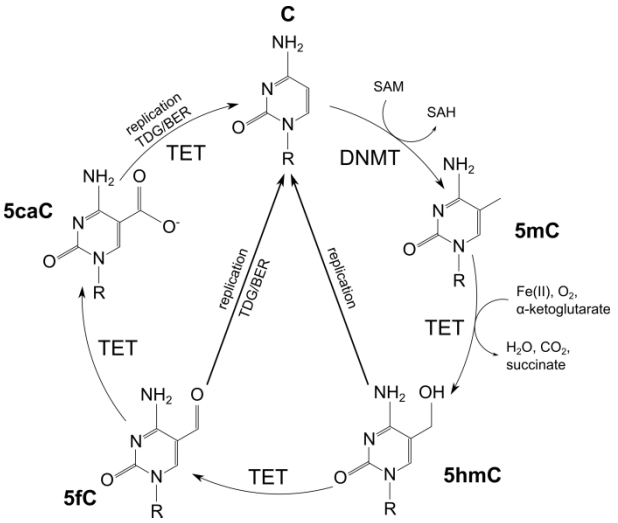

Figure 1.3: Mechanisms of DNA methylation and active and passive demethylation.
DNMTs establish 5mC from unmodified cytosines (C). All successive oxidation steps can be mediated by TET1, TET2 and TET3 and require the indicated cofactors (adapted from Wu and Zhang, 2017).

The *TET1* gene was initially identified in a translocation between chromosomes 10 and 11 resulting in a fusion with the *MLL* gene in acute myeloid leukemia (AML; Ono et al., 2002). All three TET enzymes have a C-terminal catalytic domain consisting of a double-stranded β-helix (DSBH) domain, a cysteine-rich domain and an Fe(II)-interacting HxD motif. The N-terminal CXXC domain, which includes two zinc finger motifs, is present only in TET1 and TET3 and allows DNA binding (Figure 1.1B). The C-terminus of *TET2* most likely was separated from its putative CXXC domain by a chromosomal inversion during evolution yielding the gene *IDAX*. However, IDAX binds to unmethylated DNA and is able to recruit

TET2 (An et al., 2017; Ko et al., 2013). Both *TET1* and *TET3* also express truncated isoforms at certain developmental stages lacking the CXXC domain (Zhang et al., 2016a) and these forms have differential demethylation activity (Wu and Zhang, 2017), suggesting that their distinct expression during developmental stages might serve to adapt overall oxidative conversion activity.

TET enzymes are subject to many regulatory processes including substrate and cofactor availability, transcriptional, posttranscriptional and -translational regulation, cellular and genomic localization. TET activity is linked to cellular metabolism via α-ketoglutarate, which is provided by isocitrate dehydrogenase (IDH) enzymes. As evidenced by two studies, 5hmC levels directly seem to correlate with IDH expression, suggesting that α-ketoglutarate is a limiting factor in 5hmC generation (Lian et al., 2012; Xu et al., 2011a). Consistent with this, enhanced glucose metabolism in mice was found to elevate α-ketoglutarate in conjunction with 5hmC (Yang et al., 2014). In contrast, accumulation of fumarate and succinate due to cancer-associated mutations in fumarase and succinate dehydrogenase negatively impact on TET-dependent oxidation due to competition with α-ketoglutarate (Xiao et al., 2012). TET activity furthermore depends on oxygen, as hypoxia was shown to induce DNA hypermethylation by reducing 5mC oxidation (Thienpont et al., 2016). Conversely, vitamin C stimulates TET catalytic activity (Blaschke et al., 2013).

Another layer of regulation is provided by miRNA-dependent TET transcript degradation, various posttranslational modifications affecting DNA binding or enzymatic activity as well as caspase, calpain and proteasome-mediated TET protein degradation (Wu and Zhang, 2017). The *TET1* gene is furthermore subject to regulation by a network of pluripotency factors and rapidly downregulated upon ESC differentiation (Ficz et al., 2011; Ito et al., 2010; Koh et al., 2011; Neri et al., 2015).

Lastly, ChIP-seq studies have determined the genomic localization of TET1, identifying CGIs, active and bivalent promoters as preferentially bound (Williams et al., 2011; Wu et al., 2011b; Xu et al., 2011b). However, the assumption that the DNA binding domain of TET1 dictates its localization was disproven, since full-length and CXXC-deficient TET1 displayed very similar binding patterns (Zhang et al., 2016a). Instead, certain chromatin states or other proteins could recruit TET proteins by physical interaction, as was shown for NANOG (Costa et al., 2013), PRDM14 (Okashita et al., 2014), PRC2 (Neri et al., 2013), WT1 (Rampal et al., 2014; Wang et al., 2015b), REST (Perera et al., 2015), PPARγ (Fujiki et al., 2013), CTCF (Dubois-Chevalier et al., 2014), TDG (Weber et al., 2016; Zhang et al., 2017) or the SIN3A complex (Williams et al., 2011).

1.2.3 Biological roles of TET enzymes

1.2.3.1 Developmental role

TET enzymes are involved in many developmental processes and stages. *TET1* and *TET2* are highly expressed in mouse ESCs, but rapidly decline while *TET3* is upregulated upon *in vitro* differentiation (Ito et al., 2010; Koh et al., 2011). High levels of TET1 and TET2 are also found in the inner cell mass of blastocysts and developing PGCs. In mouse oocytes and zygotes of the one cell stage *TET3* is exclusively and highly expressed, but is downregulated with progressing cell cleavage (Iqbal et al., 2011). TET2 and TET3 are the major adult forms with wide-spread expression in many tissues (Wu and Zhang, 2014).

Phenotypes of TET-deficient mice

Phenotypic characterization of *TET*-deficient mice suggested a role for TET enzymes in differentiation and imprinting. *TET1* KO mice appear viable, healthy and fertile, but display reduced body size, weight and litter (Dawlaty et al., 2011). However, when *TET1*-deficient males were mated with WT females in a different study, paternal imprinting erasure failed and various imprinting-related phenotypic defects were observed (Yamaguchi et al., 2013). These data suggest a requirement for TET1 in imprinting regulation, whereas it is not essential for pluripotency maintenance or development. Similarly, *TET2*-deficient mice developed overtly normal, but had an augmented stem cell pool in the bone marrow accompanied by increased proliferation and impaired differentiation leading to the frequent emergence of a range of myeloid malignancies (Li et al., 2011). *TET1/2*-deficient mice displayed stochastic developmental plasticity with frequent perinatal lethality due to malformations or growth retardation, but some individuals developed normally. These exhibited hypermethylation in various tissues, disturbed methylation at several imprinted loci and reduced fertility in females. As *TET3* deletion in mice is neonatally lethal, a conditional germline-specific *TET3* KO mouse model was established (Gu et al., 2011). This revealed maternal contribution of TET3 to the zygote, which is then recruited to the paternal pronucleus. Female mice with germline *TET3*-deficiency were less fertile and descending embryos showed a pronounced increase of developmental failure. Recently however, *TET3* haploinsufficiency rather than TET3-dependent 5mC conversion was determined to be causative for the observed phenotypes (Inoue et al., 2015). *TET* triple-deficient (TKO) ESCs are completely devoid of 5hmC, display gastrulation defects and cannot contribute to embryonic development (Dai et al., 2016; Dawlaty et al., 2014). Together, these phenotypes suggest that TET enzymes have redundant, but also specific, context-dependent functions. Whether these solely depend on their dioxygenase activity or on potential non-catalytic functions remains to be studied.

Involvement in pluripotency and differentiation

The embryonic stem cell state, self-renewal and pluripotency seem to be unaffected by *TET*-deficiency. However, *TET1* KO or KD in mESCs skews their differentiation (Dawlaty et al., 2011; Ficz et al., 2011; Koh et al., 2011) and *TET2*-deficiency delays transcriptional adaptation during differentiation by enhancer suppression (Hon et al., 2014). DKO for *TET1* and *TET2* in mESCs likewise retained pluripotency and three germ layer competence. However, differentiation was skewed similar to *TET1*-deficient cells (Dawlaty et al., 2013). TKO mESCs are characterized by normal morphology and pluripotency gene expression, however they cannot differentiate properly (Dawlaty et al., 2014). Furthermore, *TDG*-deficient mice are embryonic lethal, but their ESCs self-renew normally, suggesting that demethylation is not essential for ESC maintenance but differentiation (Cortázar et al., 2011; Cortellino et al., 2011; Shen et al., 2013).

DNA methylation seems to be an obstacle in somatic cell reprogramming. Indeed, it was found that during Yamanaka factor-mediated reprogramming, TET enzymes are essential to induce NANOG and ESRRB expression (Doege et al., 2012) and to activate some key miRNAs required for mesenchymal-to-epithelial transition (Hu et al., 2014). Interestingly, TET1 was also able to substitute OCT4 as one of the four reprogramming factors and additionally augment reprogramming efficiency (Costa et al., 2013; Gao et al., 2013).

In the adult organism, constitutive differentiation processes take place for instance in the hematopoietic compartment. TET enzymes were found to be essential for normal myeloid and B-cell differentiation (Li et al., 2011; Orlanski et al., 2016) as well as proper NK cell differentiation (Tsagaratou et al., 2017). Furthermore, TET-dependent demethylation occurs in a variety of hematopoietic cell types and differentiation scenarios (Kallin et al., 2012; Klug et al., 2013; de la Rica et al., 2013; Yue et al., 2016).

In conclusion, TET enzymes seem to be dispensable for maintaining the pluripotent state, but crucial for proper differentiation due to their activities on genomic regulatory elements.

Involvement in methylation reprogramming

The two prominent DNA methylation reprogramming events in primordial germ cells and the zygote represent interesting settings to study demethylation and thus have been revised since the discovery of TET activity. While the maternal genome is successively demethylated during cell division suggestive of a dilution-dependent mechanism, the paternal genome displays immediate 5mC loss after fertilization but before replication (Mayer et al., 2000), indicating active removal. At this time TET3 specifically localizes to the paternal pronucleus

and large-scale generation of all oxidized mC forms takes place (Gu et al., 2011; Inoue and Zhang, 2011; Inoue et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011). However, the driving force of demethylation nevertheless was suggested to be replication-dependent dilution (Inoue and Zhang, 2011; Inoue et al., 2011), as inhibition of DNA synthesis resulted in a greater reduction of DNA demethylation than TET3 abrogation (Guo et al., 2014; Shen et al., 2014a). Interestingly, a recent study provided evidence that the paternal genome might undergo methylation loss even before 5mC oxidation, which argues for a completely TET-independent demethylation mechanism (Amouroux et al., 2016).

Once zygotic genomes are demethylated, they completely remethylate and only then PGCs emerge, which demethylate their genomes again in two subsequent waves. While the first major wave employs mainly dilution of 5mC by downregulating UHRF1, DNMT3A and B (Kagiwada et al., 2013; Kurimoto et al., 2008), the second one makes use of TET1/2-mediated 5hmC production followed by replication-dependent dilution. Consistently, *TET1* and *TET2* are upregulated during this phase in PGCs (Hackett et al., 2013). Most global demethylation in PGCs is however TET-independent as *TET1/2*-deficiency has only locus-specific effects on methylation. For instance, TET1 is important to demethylate specific imprinted genes (Yamaguchi et al., 2012, 2013).

Involvement in DNA damage repair

Recent studies indicate that TET activity might be required for DNA damage repair. Firstly, transcriptional regulation of DNA repair genes seems to rely on TET enzymes and thus *TET* deficiency results in impaired telomeres, increased chromosomal abnormalities and the accumulation of DNA breaks and γH2AX (An et al., 2015; Cimmino et al., 2015; Kafer et al., 2016; Lu et al., 2014; Yamaguchi et al., 2012; Yang et al., 2016a). Secondly, a direct role for 5hmC has been suggested as it was enriched at DNA damage sites and mitotic chromosome segregation defects were observed in replication-stressed TKO mESCs (Kafer et al., 2016).

1.2.3.2 Role in disease

Somatic mutations of *TET* enzymes have been detected in a variety of cancers, particularly in hematological malignancies. While *TET1* and *TET3* mutations are rare, *TET2* mutations have been frequently identified in a spectrum of hematological diseases with both myeloid and lymphoid origin (Abdel-Wahab et al., 2009; Delhommeau et al., 2009; Gaidzik et al., 2012; Nibourel et al., 2010; Quivoron et al., 2011; Scopim-Ribeiro et al., 2015; Tefferi et al., 2009). CMML (20-58%), AML (2-32%), MDS (6-26%) and T-cell lymphoma (20-83%) are the diseases with the highest observed mutation frequency in *TET2* (Scourzic et al., 2015). As the phenotype of *TET2*-deficient mice was dominated by the development of myeloid malignancies after long latency, a tumor suppressor role was attributed to TET2 (Ko et al.,

2011; Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011). However, it was hypothesized that other oncogenic hits are required for malignant transformation. Consistently, *TET2* mutations were also observed in healthy elderly individuals with clonal hematopoiesis (Busque et al., 2012; Genovese et al., 2014; Jaiswal et al., 2014).

TET1 seems to have context-dependent functions in leukemogenesis. While its deficiency alone or in combination with *TET2* has been associated with the development of B-cell malignancies (Cimmino et al., 2015; Zhao et al., 2015), another study reported that it has a critical oncogenic role in MLL-rearranged leukemia (Huang et al., 2013a).

Since TET2 and TET3 are the major expressed forms in the hematopoietic system, a mouse model with inducible DKO in hematopoietic precursor cells was established recently, resulting in the rapid emergence of aggressive myeloid leukemia. This was accompanied by skewed differentiation towards the myeloid lineage and accumulation of DNA damage (An et al., 2015), suggesting TET requirement for lineage priming and genomic integrity.

In solid cancers, all three TET enzymes have been found mutated, but with low frequencies. The functional importance of these mutations remains to be clarified, especially since these tumors accumulate many somatic mutations. However, reduced TET enzyme and global 5hmC levels are a general hallmark of a wide range of cancers (Rasmussen and Helin, 2016; Shen et al., 2014b), but it is currently unclear whether this is due to increased proliferation rates that have been shown to negatively correlate with 5hmC amounts (Bachman et al., 2014).

1.3 The interaction between IDH und TET enzymes

1.3.1 Molecular functions of IDH and mutated IDH (mIDH)

Isocitrate dehydrogenases (IDH) are metabolic enzymes that decarboxylate isocitrate to α-ketoglutarate. There are three members comprising IDH1, IDH2 and IDH3. While IDH1 is located in the cytosol, IDH2 and IDH3 localize to the mitochondria, where they participate in the tricarboxylic acid cycle. IDH1 and IDH2 form homodimers and use $NADP^+$ as a cofactor, whereas IDH3 requires NAD^+ (Figure 1.4). In addition to crucial metabolites for cellular anabolism, IDH synthesize NADH for respiration and NADPH for fatty acid metabolism and redox homeostasis (Dang et al., 2016). Over 80 dioxygenases rely on α-ketoglutarate to exert their catalytic functions, including TET enzymes, histone demethylases and a range of protein hydroxylases (Inoue et al., 2016a).

Recurrent cancer-associated mutations occur exclusively in *IDH1* and *IDH2* and invariably affect key arginines of the enzyme´s active site – R132 in IDH1 and R140 or R172 in IDH2. These mutations occur early in clonal expansion, are heterozygous and mostly mutually exclusive, suggesting an oncogenic gain of function (Dang et al., 2016; Medeiros et

al., 2017). Indeed, it was shown that the observed mutations confer neomorphic activity, allowing the conversion of α-ketoglutarate into D-2-hydroxyglutarate (D-2-HG) using NADPH (Figure 1.4; Dang et al., 2009; Gross et al., 2010; Ward et al., 2010). Since all of the reported mutations converge to the accumulation of D-2-HG to millimolar concentrations in cancer cells, it has been recognized as an oncometabolite (Dang et al., 2016).

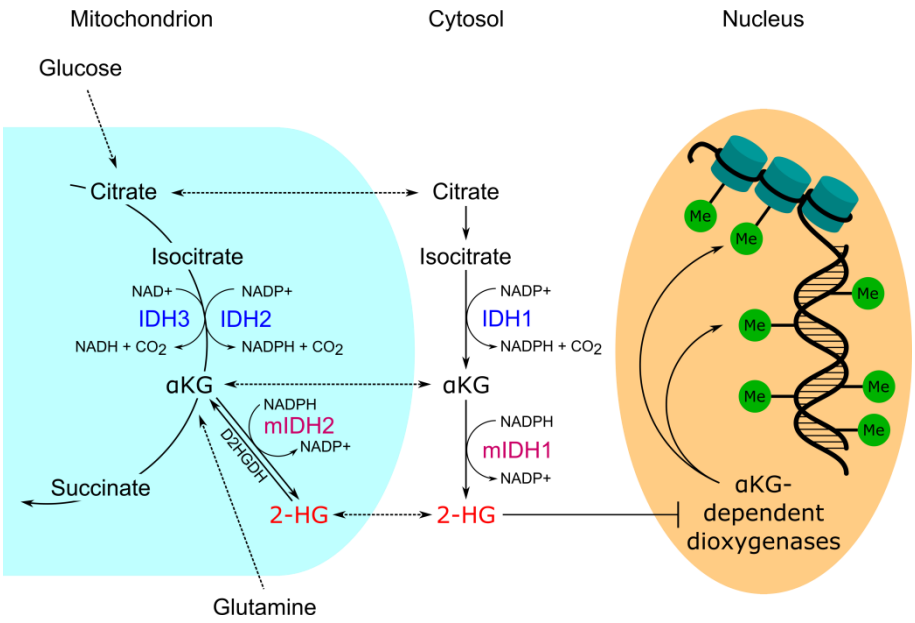

Figure 1.4: Actions of WT IDH enzymes in homeostasis and mutant IDH1/2 in disease.
αKG – α-ketoglutarate, 2-HG – D-2-hydroxyglutarate (adapted from Dang et al., 2016).

D-2-HG was subsequently shown to impair a variety of cellular processes. Firstly, it was reported to competitively inhibit histone demethylases (Chowdhury et al., 2011) and TET enzymes (Koivunen et al., 2012; Xu et al., 2011a), thereby affecting epigenetic regulation (Figure 1.4). Secondly, it seems to affect HIF prolyl-4-hydroxylases, which regulate the stability of hypoxia-inducible factor (HIF) and thereby might aberrantly influence hypoxia signaling (Koivunen et al., 2012; Xu et al., 2011a; Zhao et al., 2009). In addition, D-2-HG inhibited certain prolyl-hydroxylases involved in proper basement membrane function in the brain (Sasaki et al., 2012a), cytochrome c oxidase IV of the respiratory chain in AML cells (Chan et al., 2015) and fumarase as well as succinate dehydrogenase, resulting in hypersuccinylation of lysine residues, mitochondrial depolarization, impaired respiration and increased apoptosis resistance (Li et al., 2015). In another study D-2-HG was associated with paracrine stimulation of NFkB-signaling in bone marrow stromal cells, which supported AML cell proliferation and chemoresistance (Chen et al., 2016). Finally, *mIDH* produces $NADP^+$ instead of NADPH, which potentially hampers the removal of reactive oxygen species and confounds the tricarboxylic acid cycle by reducing α-ketoglutarate levels. Eventually, this may

compromise the normal synthesis of several biomolecules (Inoue et al., 2016a). Together, these effects result in a pronounced disturbance of the cellular homeostasis.

1.3.2 *IDH* mutations in cancer

Neomorphic *IDH* mutations have been observed in glioma, AML, cholangiocarcinoma, chondrosarcoma and premalignant disorders like MDS and myeloproliferative neoplasms (Inoue et al., 2016a). In glioma, CIMP was associated with *mIDH* (Noushmehr et al., 2010) and it was claimed that *mIDH1* is sufficient to establish this hypermethylation pattern (Turcan et al., 2012). Similar reports showed *mIDH*-associated genomic hypermethylation in AML (Akalin et al., 2012) that supposedly was caused by inhibition of the demethylating TET activity (Figueroa et al., 2010; Turcan et al., 2012). A number of studies since then have confirmed the association of *mIDH* with hypermethylation in AML and glioma patients, mouse and *in vitro* models (Chaturvedi et al., 2016; Duncan et al., 2012; Flavahan et al., 2015; Kernytsky et al., 2014; Sasaki et al., 2012b; Sturm et al., 2012; Voigt and Reinberg, 2013).

Additionally, *mIDH* has been connected to histone hypermethylation and a block in cellular differentiation using different *in vitro* and *in vivo* models (Figueroa et al., 2010; Kernytsky et al., 2014; Losman et al., 2013; Lu et al., 2012; Sasaki et al., 2012b). For example, a mouse model of conditional myeloid-specific *mIDH1* knock-in showed an increased hematopoietic progenitor pool, splenomegaly and anemia resembling human MDS with concurrent histone and DNA hypermethylation (Sasaki et al., 2012b). Recently, long-term HSCs were shown to downregulate DNA damage repair and self-renewal using the same model. Interestingly, this was independent of TET2 (Inoue et al., 2016b). Other transgenic mice expressing *mIDH2* developed AML or resembling malignancies upon cooperation with oncogenic *FLT3/NRAS* mutations or overexpression of *HOXA9/MEIS1A*. This suggests a proto-oncogenic role for *mIDH* that requires additional driving mutations or gene expression changes (Chen et al., 2013a; Kats et al., 2014). Additionally, D-2-HG itself was shown to promote proliferation, cytokine independence and differentiation inhibition in cultured cells (Losman et al., 2013) and this was supported by a recent study using a patient-derived AML xenotransplantation mouse model, in which administration of D-2-HG accelerated the onset of leukemia (Chaturvedi et al., 2016). Together, these data established the hypothesis that epigenetic deregulation is the driving factor in cancers with *mIDH*.

The *mIDH*-associated differentiation block can be resolved by suppressing D-2-HG production using small molecule inhibitors against *mIDH* (Losman et al., 2013; Rohle et al., 2013; Wang et al., 2013a). In two recent studies, the primary effect of mIDH inhibition was differentiation of AML cells *ex vivo*, in xenograft mouse models and in *in vitro* models (Shih et al., 2017; Yen et al., 2017). However, it is currently unclear how this is mediated: Several

studies claim that histone and DNA methylation patterns normalize under small molecule treatment (Kernytsky et al., 2014; Shih et al., 2017), whereas others find no reversal of DNA hypermethylation although efficiently depleting D-2-HG (Chaturvedi et al., 2017; Pusch et al., 2017; Rohle et al., 2013).

1.3.3 The hematopoietic system and acute myeloid leukemia (AML)

Hematopoiesis provides the organism with all blood and immune cells and occurs throughout lifetime. Adult hematopoietic stem cells mainly reside in the bone marrow and constantly give rise to self-renewing and differentiating daughter cells through asymmetric cell divisions. In a stepwise differentiation hierarchy, cells successively restrict their lineage potential and commit to multipotent progenitors (MPPs) from which precursors of the myeloid (common myeloid progenitors - CMPs) and lymphoid (common lymphoid progenitors - CLPs) lineages arise. Through multiple intermediate states including granulocyte macrophage progenitors (GMPs) and promyelocytes (PMNs), the specialized cells of the myeloid compartment, such as macrophages or granulocytes (PMC – polymorphonuclear cell), are generated. CLPs in turn give rise to natural killer cells, B-, and T-lymphocytes (Figure 1.5).

The different cell fates are accompanied by dynamic changes in the methylation landscape and it has been proposed that DNA methylation plays an important role in hematopoietic lineage choice (Bröske et al., 2009). Indeed, the methylomes of hematopoietic cells are distinct (Bock et al., 2012; Farlik et al., 2016). For example, HSCs change their methylation levels gradually during differentiation with a general loss of methylation upon myelopoiesis (Figure 1.5). In contrast, lymphoid cells become hypermethylated compared to HSCs and myeloid cells (Accomando et al., 2014; Bocker et al., 2011; Hodges et al., 2011; Ji et al., 2010; Rönnerblad et al., 2014). Using conditional *DNMT1* hypomorphic or KO mice, it was shown that DNMT1-mediated DNA methylation is required to maintain the HSC pool and silence premature expression of myeloid-specific genes. Impairment of maintenance methylation consequently prevented the generation of lymphoid progeny (Bröske et al., 2009; Trowbridge et al., 2009). In contrast, DNMT3A and DNMT3B are required to methylate and silence HSC specific genes like *RUNX1* or *GATA3*. Conditional ablation of DNMT3A alone or in combination with DNMT3B conferred enhanced self-renewal and repopulating capacity to HSCs and impaired their differentiation (Challen et al., 2014, 2011). Therefore, initial increases in methylation upon early progenitor commitment might be related to the silencing of stem cell genes and restriction of pluripotency (Álvarez-Errico et al., 2014).

AML is a malignant disease of the myeloid compartment that is characterized by the abnormal expansion of immature myeloid progenitor cells due to an early differentiation block and increased proliferative capacity. These cells mostly originate from the highly proliferative

progenitor pool (CD34+/CD38+) or even the HSC pool (CD34+/CD38-). As a result, various tissues are infiltrated with aberrantly or poorly differentiated hematopoietic cells, while normal hematopoiesis is prevented. To diagnose AML the percentage of immature blasts in bone marrow or blood has to exceed ~25% (Passegué et al., 2003).

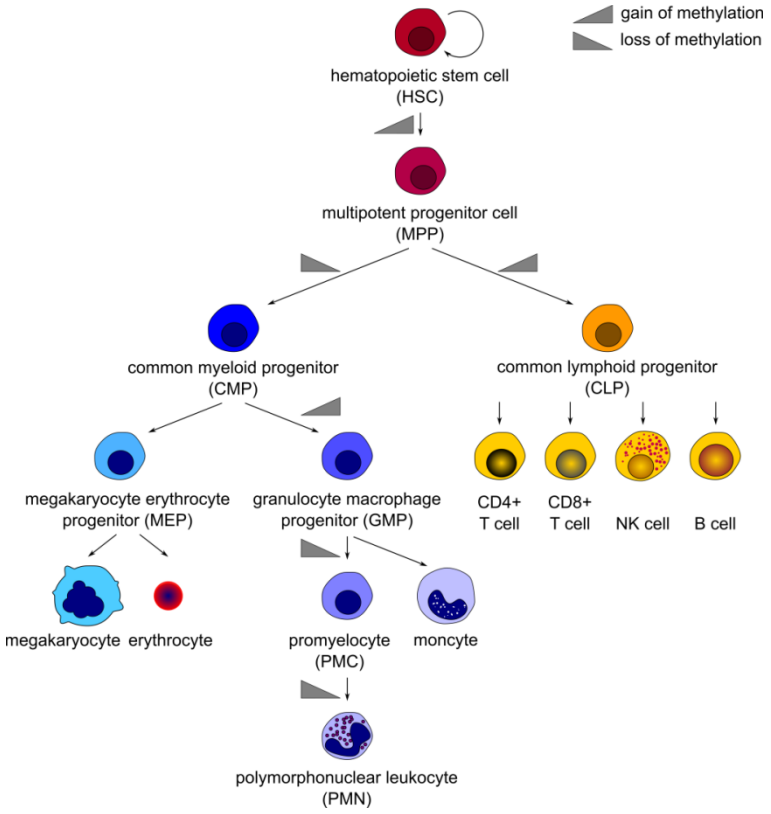

Figure 1.5: Schematic of human hematopoiesis and accompanying DNA methylation changes. Major cell fates during adult hematopoiesis are shown (modified from Farlik et al., 2016). Alterations in global DNA methylation at certain transitions were adapted from Álvarez-Errico et al., 2014.

AML is marked by inter-individual, cytogenetic, molecular, morphological and clonal heterogeneity (Döhner et al., 2015; Heuser et al., 2011). Frequently recurring mutations in AML affect *FLT3, NPM1, N- or KRAS, RUNX1, TP53, WT1, PTPN11, KIT, CEBPA, TET2, DNMT3A* and *IDH1* or *IDH2* (Voigt and Reinberg, 2013). Functionally, these genes relate to different categories among them tumor suppressor genes, receptor tyrosine kinase signaling proteins, transcriptional regulators and epigenetic modifiers (Döhner et al., 2015). The comparison of mutational status in different AML subtypes and healthy adult HSCs suggested that many of the mutations in epigenetic regulators including *IDH1, IDH2, TET2* and *DNMT3A* are found in preleukemic cells and early stages of AML. Consequently, it was hypothesized that these mutations and epigenetic deregulation might be initiating events in leukemogenesis (Corces-Zimmerman et al., 2014; Shlush et al., 2014; Welch et al., 2012). *DNMT3A* mutations occur in roughly a quarter of newly diagnosed AML patients and of these more than half

display a mutation in arginine 882, most commonly R882H (Ley et al., 2010; Voigt and Reinberg, 2013). These heterozygous mutations hamper the homotetramerization of DNMT3A, inhibit remaining WT activity in a dominant-negative manner and lead to focal methylation loss (Holz-Schietinger et al., 2012; Russler-Germain et al., 2014). Recent studies have indicated that *TET2* and *DNMT3A* mutations cooperate to induce leukemia (Scourzic et al., 2016; Zhang et al., 2016b), suggesting that antagonizing molecular functions might exacerbate epigenetic deregulation and accelerate the onset of transformation. In contrast, *TET2* and *IDH* mutations have been described as mutually exclusive in AML (Figueroa et al., 2010) and it was thus hypothesized that both contribute by similar mechanisms – aberrant DNA hypermethylation – to leukemogenesis. There is currently no consensus on the prognosis of *IDH* mutated AML patients as results are diverging across studies and seem to vary with the specific amino acid substitution and enzyme. However, a recent study did not find differences in overall survival between AML patients with *mIDH* or *IDH* WT status (Dang et al., 2016; DiNardo et al., 2015).

1.4 Aims of the thesis

TET enzymes provide a novel mechanism of epigenetic regulation antagonizing DNA methylation in mammals. While the majority of studies focused on early developmental stages to examine the functions of TET enzymes, their biological roles in differentiated cells are not well described. This is also valid in cancer cells, where their impairment through mutations in cofactor providing metabolic enzymes remains controversial. The aims of this thesis were to characterize the effects of TET-deficiency and inhibition in a selected developmental stage and relevant disease scenario, respectively, with special emphasis on DNA methylation changes as a result of defective DNA demethylation.

1. The majority of *TET1/2* deficient mice die perinatally due to developmental failures, suggesting a biological role of TET enzymes also in later developmental stages. The aim of this project was to characterize the phenotype and methylome of DKO mouse embryonic fibroblasts as a paradigm of a differentiated primary cell type with a disrupted DNA demethylation pathway.

2. Neomorphic *IDH* mutations have been associated with DNA hypermethylation in cancer. The goal of this work was to examine IDH-dependent DNA methylation changes in AML and improve the understanding of their origin with special regard to TET impairment.

2 Results

2.1 TET function in mouse development

DNA methylation patterns are dynamically changing during mammalian development and are essential to regulate pluripotency and lineage-specific genes (Smith and Meissner, 2013). In *Tet*-deficient mice developmental defects mainly occur during embryogenesis (Dawlaty et al., 2011, 2013, 2014). While *Tet* triple KO ESCs do not contribute to embryonic development (Dawlaty et al., 2014), *Tet1/2* double deficient ESCs remain pluripotent, but embryos exhibit developmental plasticity with various defects and frequent perinatal lethality (Dawlaty et al., 2013). This is accompanied by global DNA hypermethylation in all analyzed tissues. These phenotypes suggest a role for TET enzymes in early development. In order to analyze the role of TET enzymes in somatic cells, primary mouse embryonic fibroblasts (MEFs) representing a differentiated embryonic cell type, were isolated from *Tet1/2* KO mice and characterized with regard to phenotype, gene expression and DNA methylation. The following data have partially been published in Wiehle et al., 2016.

2.1.1 Phenotypic characterization of *Tet1/2*-deficient MEFs

Tet1/2-deficient (DKO) mice were created previously by intercrossing *Tet1*- with *Tet2*-deficient mice (Dawlaty et al., 2013). These mice had been generated before by gene targeting in mESCs. In the case of *Tet1,* exon 4 was deleted resulting in an out of frame fusion of exons 3 and 5, which truncated the TET1 protein and led to the loss of its catalytic domain (Dawlaty et al., 2011). The *Tet2* gene was disrupted by insertion of a GFP-cassette into exon 3, which destroyed the endogenous start codon and resulted in the expression of a truncated, non-functional protein (Li et al., 2011). Primary MEFs were isolated from DKO animals at embryonic day 13.5 according to standard procedures by M. Dawlaty and transferred to Germany for further analysis.

To confirm the disruption of *Tet1* and *Tet2* genes in the primary MEF model, qRT-PCR analysis of *Tet* transcripts was conducted using primers annealing to the corrupted regions. DKO MEFs were devoid of detectable *Tet1* and *Tet2* mRNA, whereas *Tet3* transcript amounts were not significantly changed compared to WT MEFs (Figure 2.1A). This was confirmed by *Tet3* transcript levels detected in RNA-seq analysis (Figure 2.1A, right panel). These data confirm *Tet1/2* transcript disruption and the absence of compensatory *Tet3* upregulation in response to genetic *Tet1*- and *Tet2*-deficiency.

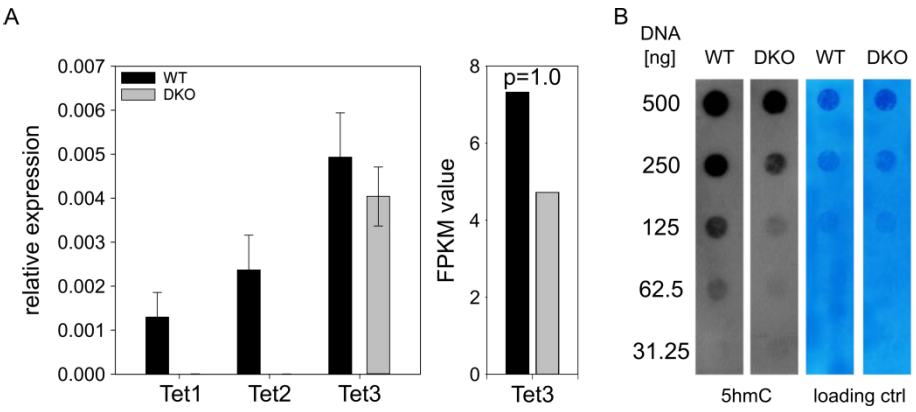

Figure 2.1: Validation of the *Tet1/2*-deficient MEF model.

(A) Expression of the three *Tet* enzymes in WT and DKO MEFs was measured by qRT-PCR. The right panel shows FPKM values for *Tet3* extracted from RNA-seq data. **(B)** Immunostaining of hydroxymethylation in genomic DNA of WT and DKO MEFs. Methylene blue staining was used as loading control.

TET enzymes catalyze the oxidative conversion of 5mC to 5hmC. To confirm functional impairment of TET enzymes in the DKO model, global 5hmC levels were examined. To this end genomic DNA of WT and DKO MEFs was immobilized by dotblotting and stained with an antibody specific to 5hmC. This showed a moderate overall reduction of the 5hmC mark in the DNA of *Tet1/2*-deficient MEFs (Figure 2.1B). DNA hydroxymethylation was not completely abrogated though, indicating remaining TET3 activity. Together, these data demonstrate disruption of the *Tet1* and *Tet2* genes in DKO MEFs, which is accompanied by reduced DNA hydroxylating activity. Thus these cells are a valid model to investigate further *Tet1/2*-dependent changes.

Phenotypic characterization of DKO MEFs revealed a flattened morphology with widespread protrusions (Figure 2.2A) and a reduced growth rate (Figure 2.2B), indicating that DKO MEFs have reduced proliferation capacity.

Previous studies have reported that mESCs display skewed or defective differentiation upon *Tet* -deficiency (Dawlaty et al., 2013, 2014; Hon et al., 2014; Ito et al., 2010; Koh et al., 2011; Li et al., 2014). MEFs isolated between embryonic day 12 and 14 are mesenchymal cells that generate normal tissue fibroblasts and fibrocytes. However, they are still multipotent and retain the capacity to differentiate into adipocytes, chondrocytes and osteoblasts (Rastegar et al., 2010; Rosen and MacDougald, 2006).

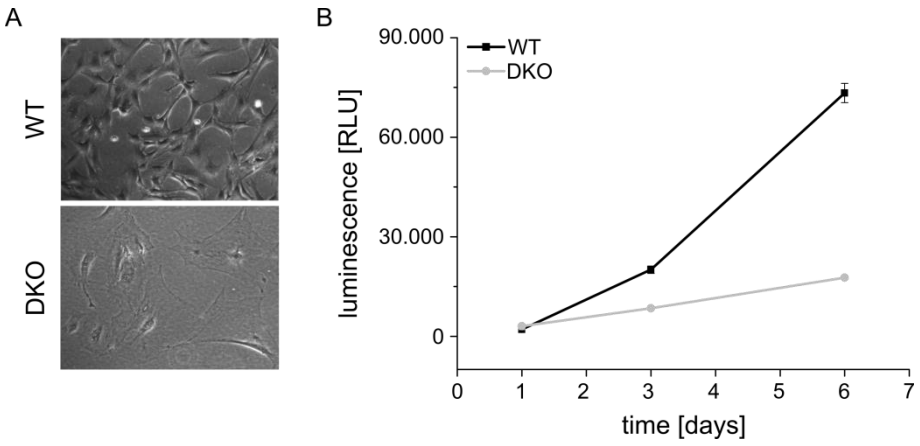

Figure 2.2: Phenotypic changes in *Tet1/2*-deficient MEFs.

(A) Representative bright field microscopic image of WT and DKO MEFs. **(B)** Growth curve of WT and DKO MEFs based on metabolic activity.

In order to test the differentiation potential of *Tet1/2*$^{-/-}$ MEFs, *in vitro* adipogenesis was performed by incubating the cells in adipogenic differentiation medium (ADM). While WT MEFs started to form microscopically visible lipid droplets around day 7 of treatment, the amount of droplets generated in DKO MEFs was significantly decreased (Figure 2.3), indicating inefficient adipogenesis in *Tet1/2*$^{-/-}$ MEFs.

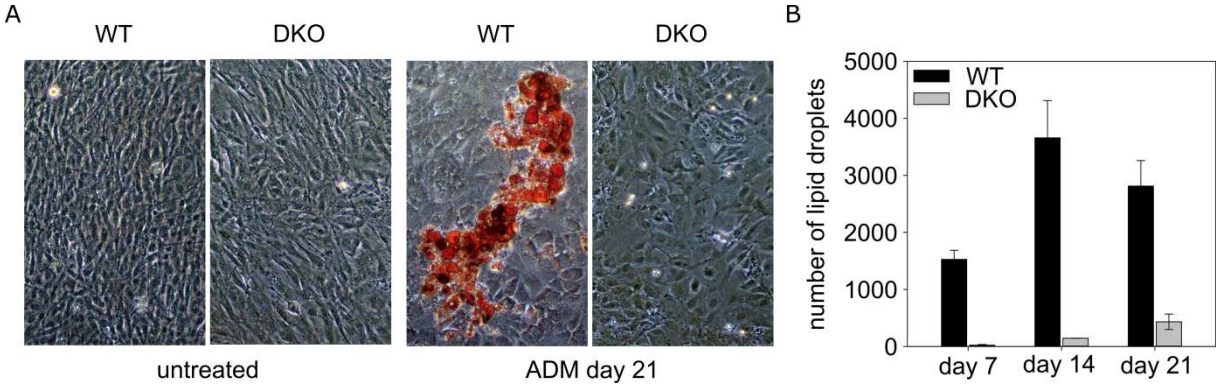

Figure 2.3: Incomplete adipogenesis in *Tet1/2*-deficient MEFs.

(A) Bright field microscopic images of Oil Red O stained MEFs that were incubated in standard culture medium (untreated) or ADM for 21 days. **(B)** Quantification of lipid droplets at the indicated time points. Shown are the means of the numbers of lipid droplets counted in five view fields ± standard deviations.

To test whether this was due to a failure in transcriptional adaptation, essential adipogenic transcription factors and effector genes were analyzed by qRT-PCR. Adipogenesis was accompanied by the ADM-dependent transcriptional induction of the established adipogenic markers genes *Pparγ*, *C/Ebpα* and *Igf1* in WT MEFs (Rosen and MacDougald, 2006). However, induction of all three genes was strongly reduced in DKO

MEFs (Figure 2.4), suggesting that the cells fail to efficiently activate adipogenic differentiation genes.

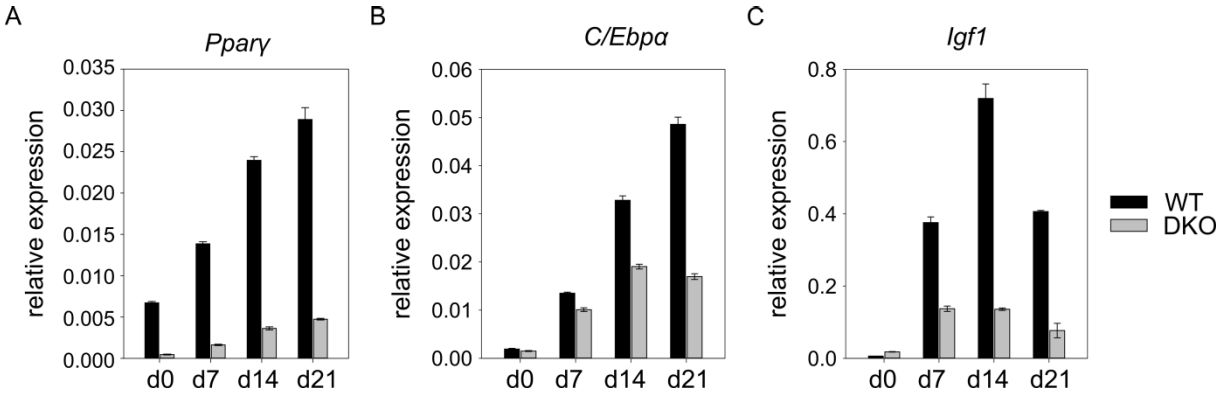

Figure 2.4: Inefficient transcription of adipogenic marker genes in *Tet1/2*-deficient MEFs.
Expression of adipogenic markers *Pparγ* **(A),** *C/Ebpα* **(B)** and *Igf1* **(C)** was analyzed by qRT-PCR in WT and DKO MEFs before (d0) and at day 7, 14 and 21 of ADM treatment. *Actb* and *Gapdh* were used as reference genes. Bars represent means and error bars standard deviations of three replicates.

As TET1 and TET2 seemed to be required for adipogenesis their expression dynamics were analyzed during this process. All three *Tet* transcripts were upregulated during adipogenesis in WT MEFs (Figure 2.5). While *Tet1* and *Tet3* were only transiently induced (Figure 2.5A, C), *Tet2* expression was continuously increasing over time (Figure 2.5B). Overall transcript amounts of *Tet2* and *Tet3* were higher than *Tet1*, consistent with predominant *Tet3* (and depending on the tissue *Tet2*) expression upon differentiation (Koh et al., 2011; Wu and Zhang, 2014). Interestingly, DKO MEFs expressed increased amounts of *Tet3* upon adipogenesis, suggesting that in a setting where dynamic adaptation of TET levels is required, the cell might try to compensate for the loss of *Tet1* and *Tet2* by upregulating *Tet3*.

Taken together, these data indicate an altered phenotype with a pronounced differentiation and gene expression defect in *Tet1/2*-deficient MEFs. Moreover, they point towards a functional requirement of TET enzymes in adipogenesis.

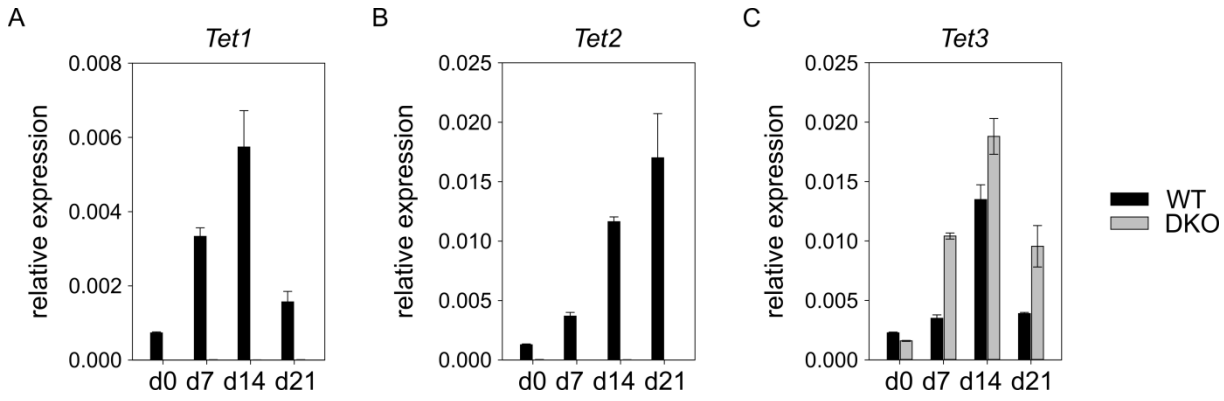

Figure 2.5: *Tet* expression is induced during adipogenesis.

Tet1 **(A)**, *Tet2* **(B)** and *Tet3* **(C)** expression dynamics during adipogenic differentiation were analyzed by qRT-PCR in WT and DKO MEFs. *Actb* and *Gapdh* were used as reference genes. Bars represent means and error bars standard deviations of three replicates.

2.1.2 DNA methylation changes in *Tet1/2*-deficient MEFs

The observed differentiation and gene regulation defects in DKO MEFs together with the reported demethylating activity of TET enzymes (He et al., 2011; Inoue and Zhang, 2011) suggested defects in epigenetic programming upon *Tet1/2* KO. To investigate potential alterations due to reduced DNA demethylation in the methylomes of WT and *Tet1/2*-deficient MEFs, whole genome bisulfite sequencing was performed. This technique yields a genome-wide base resolution map of methylated cytosines and, depending on the coverage, high confidence methylation ratios.

Analysis of two independent biological replicates per genotype yielding a combined average coverage of 20x for each WT and DKO methylome showed that average methylation levels were comparable between replicates, but moderately increased in DKO MEFs (Figure 2.6A). To conduct an unbiased analysis of the distribution of hypermethylation within the DKO genome, it was segmented into different subgenomic features based on the presence of certain histone modifications, PolII and CTCF. To this end, previously published genome-wide maps of these marks in WT MEFs were used (Mikkelsen et al., 2007; Shen et al., 2012). The combined patterns of modified histones, PolII and CTCF allowed assigning 15 functional classes and the average methylation level of each class was calculated. This revealed widespread hypermethylation in DKO MEFs, which affected almost all analyzed features (Figure 2.6B). Enhancers and promoters showed the most pronounced hypermethylation. These results demonstrate hypermethylation of functional regulatory elements upon *Tet1/2*-deficiency, suggesting a demethylating activity for TET1 and TET2 at these regions.

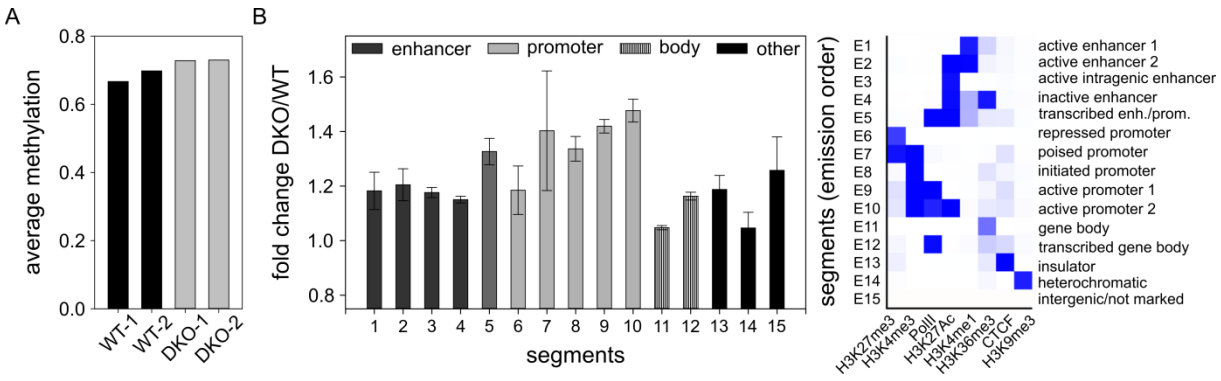

Figure 2.6: *Tet1/2*-deficient MEFs show widespread DNA hypermethylation.
(A) Average methylation ratios determined by whole genome bisulfite sequencing of two WT and DKO replicates. **(B)** Fold change of average methylation between DKO and WT MEFs of 15 subgenomic features identified by analysis of chromatin states. Error bars represent standard deviations of the two biological replicates. The right panel shows the 15 identified segments. The intensity of blue color reflects the appearance of histone modifications, PolII and CTCF in each segment. Functional elements were identified by the combined histone modification and protein patterns and are indicated on the right. These analyses were performed by Günter Raddatz.

As promoters were the most prominently hypermethylated elements in DKO MEFs, these features were analyzed in more detail. Previously described DNA methylation canyons are associated with the genes of developmentally important transcriptional regulators, mostly harbor at least one CGI and their borders are enriched for 5hmC (Jeong et al., 2013). As such they represented interesting promoter-associated features to be analyzed in a *Tet*-deficient context.

Mapping of DNA methylation canyons in WT MEFs showed that 95% of canyons overlapped with a transcriptional start site (Figure 2.7A). The methylome of WT MEFs exhibited on average 1330 canyons. In DKO MEFs fewer canyons were detected, possibly due to the failure to meet the minimal size of 3.5 kb (Figure 2.7B). Next, the average methylation levels of WT and DKO canyons were analyzed. As expected, methylation ratios of canyons in general were very low. However, DKO MEFs showed a significantly increased median canyon methylation compared to WT MEFs (Figure 2.7C). Determination of canyon sizes in WT and DKO MEFs revealed that the majority of canyons in DKO cells were smaller than in WT MEFs (30%) or completely disappeared (38%) from the analysis as they did not meet the minimal size criterion (Figure 2.7D). A quarter of the canyons detected in WT MEFs were expanded in DKO MEFs and 7% did not display significant changes. These findings show that canyon hypermethylation leading to canyon shrinkage or collapse is a prominent feature of the methylome of *Tet1/2*-deficient MEFs.

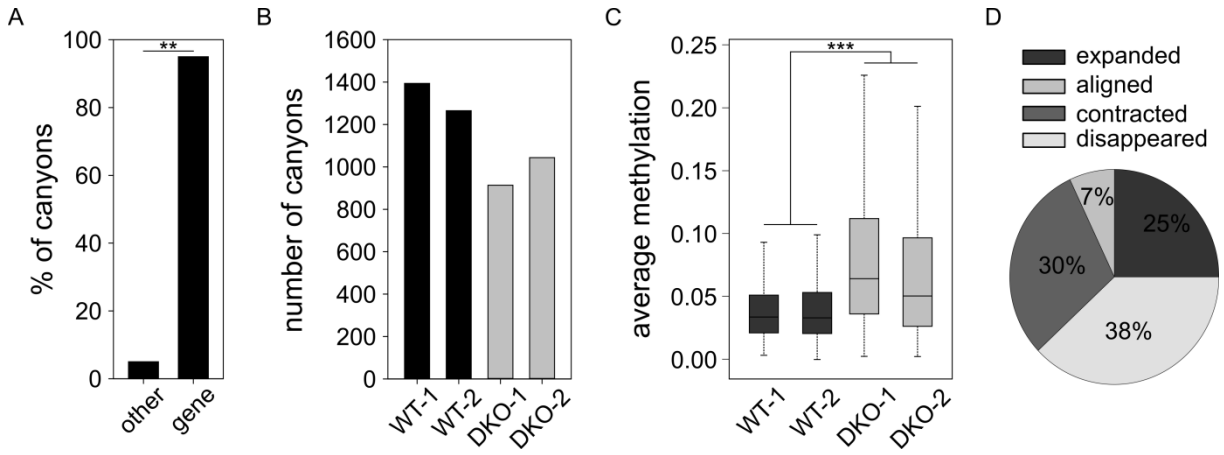

Figure 2.7: *Tet1/2*-deficiency in MEFs results in canyon hypermethylation.
(A) Association of canyons with transcriptional start sites (gene). Significance level was calculated using the chi-squared test (**: P<0.01). **(B)** Numbers of canyons in the two WT and DKO replicates. **(C)** Average methylation of canyons identified in the two WT and DKO replicates. The two-sided paired *t*-test was used to calculate the significance level of the difference between the two groups (***: P=2.2 x 10^{-16}). **(D)** Piechart showing size changes of canyons in DKO MEFs compared to the corresponding canyons in WT MEFs. These analyses were performed by Günter Raddatz.

The loss and size reductions of canyons in DKO cells suggested that they are collapsing due to invading hypermethylation at canyon borders. To confirm this, methylation profiles were visually examined at canyons. A prominent example of a large canyon with border hypermethylation was the one at the anterior *Hoxa* gene cluster (Figure 2.8A). This canyon did not only show increased methylation at the margins but also in the center. This is in accordance with *Tet2* hydroxylating and potentially removing 5mC in the *HOXA* cluster during differentiation of NT2 cells (Bocker et al., 2012) and suggests that Tet enzymes have conserved functions at the *Hoxa* cluster in humans and mice.

To examine canyon borders on a global scale, all canyons were size-normalized and the average methylation ratio at each relative position was displayed. This showed a general increase of average methylation in both DKO replicates, which was enhanced at the margins of the canyons (Figure 2.8B), indicating that hypermethylation from adjacent methylated regions invades the canyon due to the lack of TET1/2 activity.

To validate canyon border hypermethylation, amplicon bisulfite sequencing was performed targeting selected canyon margins. Individual amplicons were covered by at least 100 reads providing high confidence information on the methylation status of the examined CpGs. All selected loci displayed hypermethylation at all the interrogated CpG residues in DKO MEFs (Figure 2.9), confirming canyon border hypermethylation independently of whole genome bisulfite sequencing.

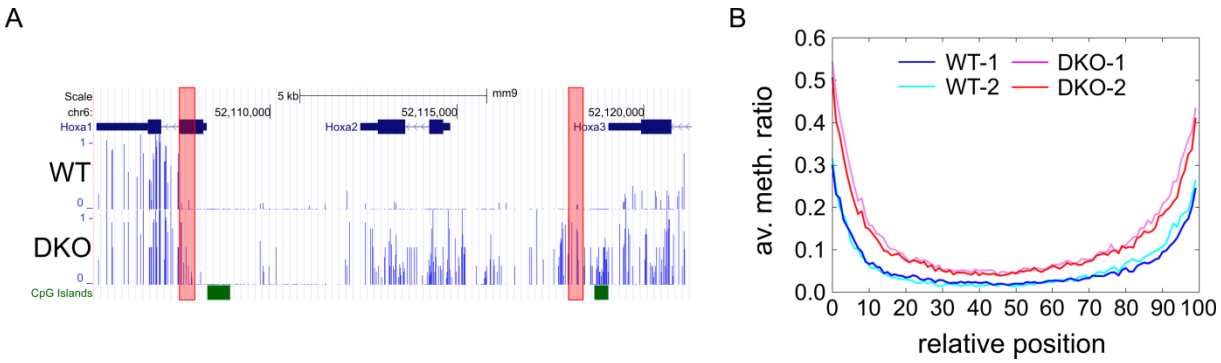

Figure 2.8: Canyon borders are maintained by TET1/2.

(A) Example of a hypermethylated canyon at the *Hoxa* gene cluster. Methylation ratios of individual CpGs in WT-1 (top) and DKO-1 (bottom) MEFs are indicated by vertical blue lines. Red squares show the positions of the first two amplicons analyzed in Figure 2.9. **(B)** Superposition of the methylation profiles of all size-normalized canyons in WT and DKO MEFs. This analysis was performed by Günter Raddatz.

Taken together, these results show that TET1 and TET2 prevent genomic hypermethylation of enhancers, promoters and specifically DNA methylation canyons. In particular, their demethylating activity is critical to protect canyon borders from aberrantly invading methylation.

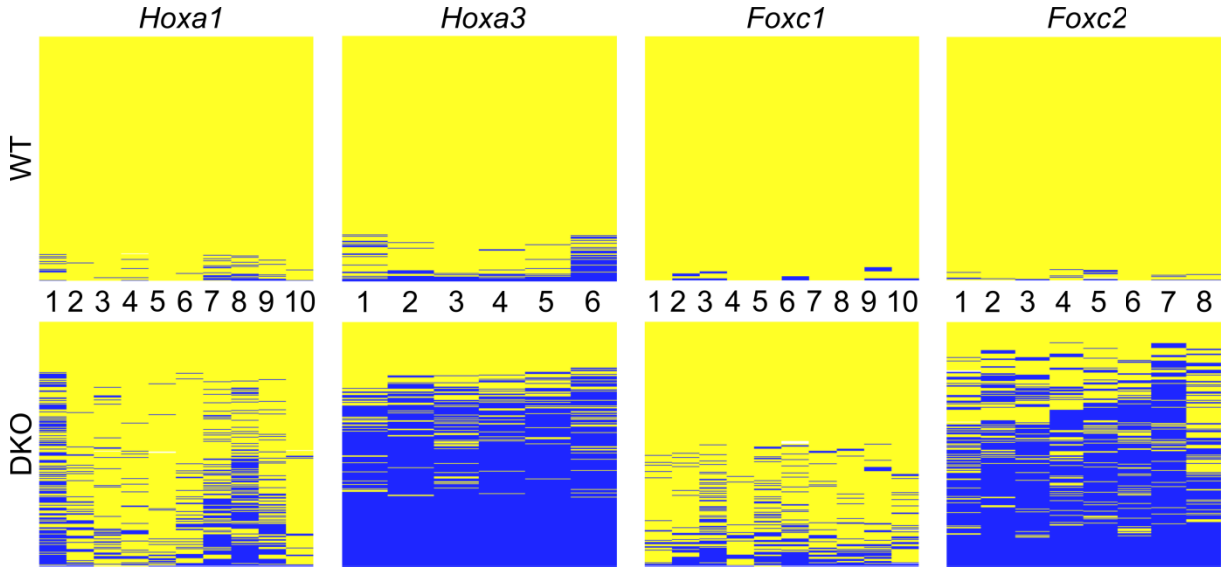

Figure 2.9: Validation of canyon border hypermethylation.

Targeted amplicon bisulfite sequencing at borders of canyons associated with the genes *Hoxa1*, *Hoxa3*, *Foxc1* and *Foxc2*. Data are displayed in heatmaps with each row representing a sequencing read and each column a CpG within this read. Individual CpGs are numbered and blue boxes indicate methylated and yellow boxes unmethylated cytosines. Positions of the analyzed regions are shown in Figure 2.8A and Figure S1.

2.1.3 Gene regulation defects in *Tet1/2*-deficient MEFs

In order to investigate the functional consequences of *Tet1/2*-dependent hypermethylation comprehensively, transcriptomes of WT and DKO MEFs were analyzed using RNA-seq. This identified 301 differentially expressed genes (q-value<0.05) with both up- and downregulated transcripts being present in DKO MEFs (Figure 2.10A). Transcriptional silencing by promoter hypermethylation is thus not the sole mechanism of gene deregulation upon *Tet1/2* KO.

Pathway analysis of deregulated genes revealed strong enrichment for developmental categories (Figure 2.10B). This included “growth and proliferation” as well as “(connective) tissue development” among the top downregulated pathways, which is consistent with the observed reduction in growth rate and differentiation potential in DKO MEFs. These findings are also in accordance with a published study on the involvement of TET1 in cell cycle progression in somatic cells (Huang et al., 2013b).

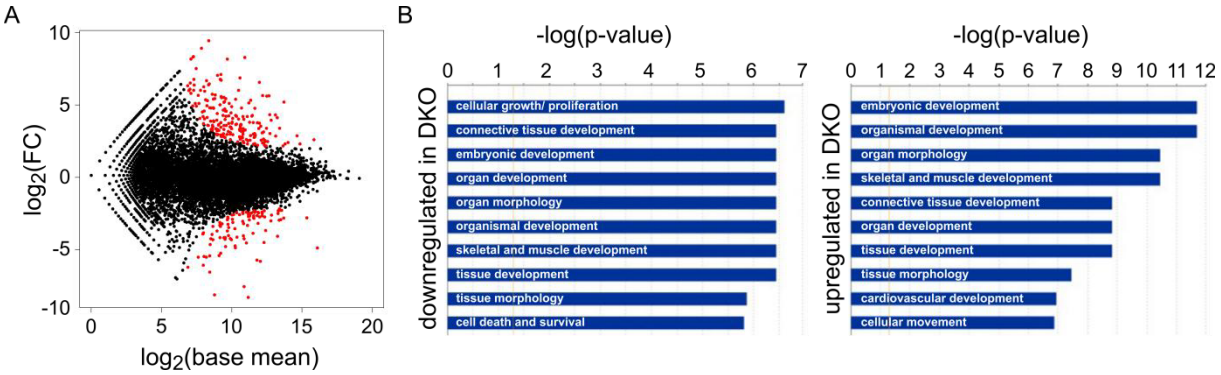

Figure 2.10: *Tet1/2*-deficiency disturbs gene expression in MEFs.
(A) MA plot showing the fold change [log_2(FC)] of each gene as a function of the averaged WT and DKO normalized read counts [log_2(base mean)] per gene. Differentially expressed transcripts based on a q-value<0.05 are shown in red. This analysis was performed by Günter Raddatz. **(B)** Top ten biological pathways significantly enriched among downregulated (left panel) and upregulated (right panel) genes in DKO MEFs.

To validate gene expression changes detected by RNA-seq, qRT-PCR for selected deregulated genes was performed. This confirmed the up- and downregulation of several genes in DKO MEFs observed by RNA-seq (Figure 2.11), of which many are involved in transcriptional regulation (e.g. the *Hox* and *Fox* transcription factors). In summary, these data indicate that *Tet1/2*-deficient MEFs exhibit disturbed gene expression patterns, which potentially alter their developmental state and differentiation capacity.

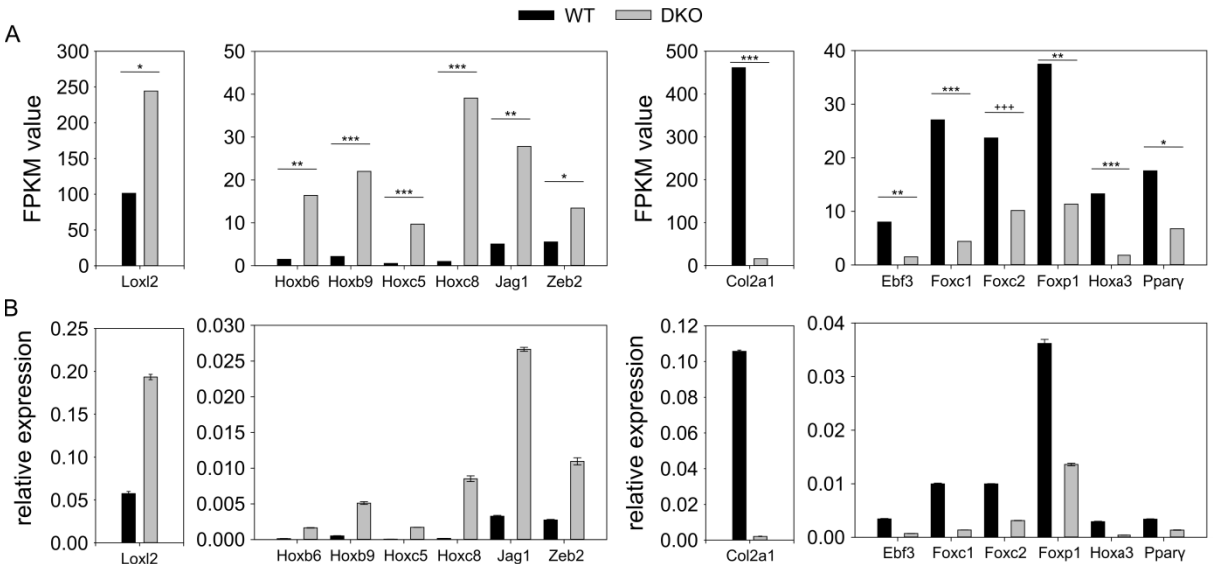

Figure 2.11: Validation of gene expression changes in *Tet1/2*-deficient MEFs.

(A) FPKM values of selected genes extracted from RNA-seq analysis are shown. In the left panels upregulated and in the right panels downregulated genes in DKO MEFs are displayed. **(B)** QRT-PCR analysis of the corresponding genes from (A). *Actb* and *Gapdh* were used as reference genes. Bars represent means and error bars standard deviations of three replicates. All genes except of *Foxc2* were significantly changed in RNA-seq (***: Q<0.001, **: Q<0.01, *: Q<0.05), but the reduction in *Foxc2* expression was found to be significant in qRT-PCR (+++: P<0.001).

2.1.4 Dysregulation of canyons and associated genes during adipogenesis

To examine the correlation between canyon hypermethylation and defective gene expression in MEFs, the overlap between gene deregulation and localization of the respective gene in a hypermethylated canyon was analyzed.

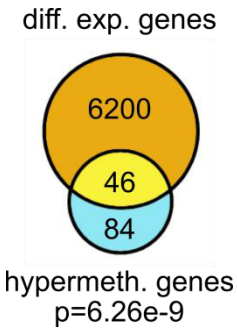

Figure 2.12: Association of gene deregulation and localization in a hypermethylated canyon.

The overlap between all differentially expressed genes (diff. exp. genes) with a |log2(FC)|≥0.5 and all genes associated with a hypermethylated canyon (hypermeth. genes) is shown. Statistical significance was determined using the hypergeometric test (P=7.55 x 10-7). Analysis and statistical testing were performed by Günter Raddatz.

To this end, all differentially expressed genes with a |log(FC)|≥0.5 were determined (6200). Furthermore all genes residing in a hypermethylated canyon were identified (84). More than half of these genes overlapped with the group of differentially expressed genes and this intersection was statistically significant (P=6.26 x 10^{-9}; Figure 2.12), indicating that the presence of a gene promoter in a hypermethylated canyon likely leads to deregulation of the respective gene. However, the majority of deregulated genes was not associated with a hypermethylated canyon, indicating that additional mechanisms lead to gene deregulation upon *Tet1/2*-deficiency.

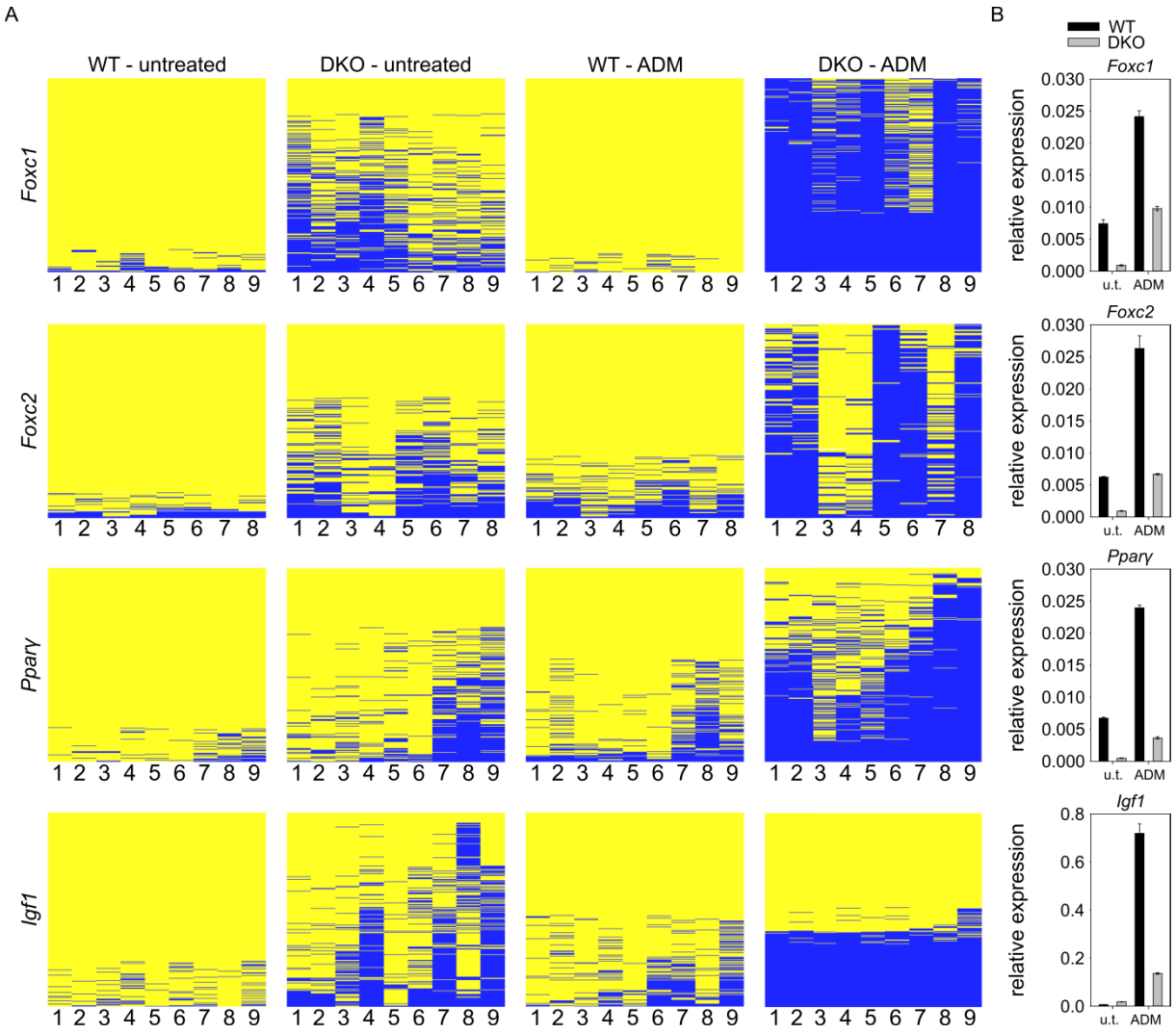

Figure 2.13: Promoter hypermethylation correlates with defective gene induction during adipogenesis in *Tet1/2*-deficient MEFs.

(A) Targeted amplicon bisulfite sequencing of promoters of four canyon-associated adipogenic marker genes during adipogenesis in WT and DKO MEFs. Untreated and cells treated with ADM for 14 days were analyzed. Localization of PCR amplicons in the canyon is shown in Figure S1. **(B)** QRT-PCR of the corresponding genes in untreated (u.t.) and cells treated with ADM for 14 days. *Actb* and *Gapdh*

were used as reference genes. Bars represent means and error bars standard deviations of three replicates.

To investigate the contribution of hypermethylation to the observed differentiation defect, methylation dynamics at DNA methylation canyons during adipogenesis were analyzed. For this purpose, four adipogenic marker genes that were associated with a canyon – *Pparγ, Igf1, Foxc1* and *Foxc2* - were selected. Promoter methylation and expression dynamics of these genes were analyzed by targeted amplicon bisulfite sequencing and qRT-PCR, respectively, during the course of adipogenesis. After 14 days of ADM treatment the initial hypermethylation of these promoters present in undifferentiated DKO cells was greatly exacerbated (Figure 2.13A), leading for example to almost complete methylation of the *Foxc1* promoter in contrast to WT MEFs. In conjunction with this, the expression of the corresponding genes was induced in WT cells, but impaired in DKO MEFs (Figure 2.13B).

In summary, these data indicate that canyon gene dysregulation is significantly associated with hypermethylation of its respective canyon and correlated with promoter hypermethylation during adipogenesis. These observations suggest a protective role for TET1 and TET2 in maintaining canyon-associated genes in a state, in which they can be activated upon differentiation cues.

2.1.5 Cooperative role of TET1 and TET2 in canyon maintenance

Previous studies have suggested distinct genomic target regions for TET1 and TET2 in mESCs based on depletion of hydroxymethylcytosine upon *Tet1* or *Tet2* knockdown (Huang et al., 2014). To dissect the individual contributions of TET1 and TET2 to canyon hypermethylation, single *Tet1* and *Tet2* KO MEFs were analyzed by targeted amplicon bisulfite sequencing of selected canyon borders as described above.

Comparison of single KO MEFs with WT MEFs showed that both, TET1 and TET2, contributed to canyon hypermethylation, as methylation ratios were higher in each single KO than in WT MEFs (Figure 2.14). However, TET2-dependent hypermethylation was always stronger than the TET1-mediated methylation increase, suggesting minor TET1 activity. There was no synergistic effect of *Tet*-double deficiency. These data indicate that TET1 and TET2 cooperate at canyon borders to maintain an adequate methylation state with a major role for TET2. In conclusion, the findings presented in this chapter attribute an essential and cooperative role to TET1 and TET2 in safeguarding DNA methylation canyons from erroneous methylation, thereby preventing the permanent silencing of important developmental genes.

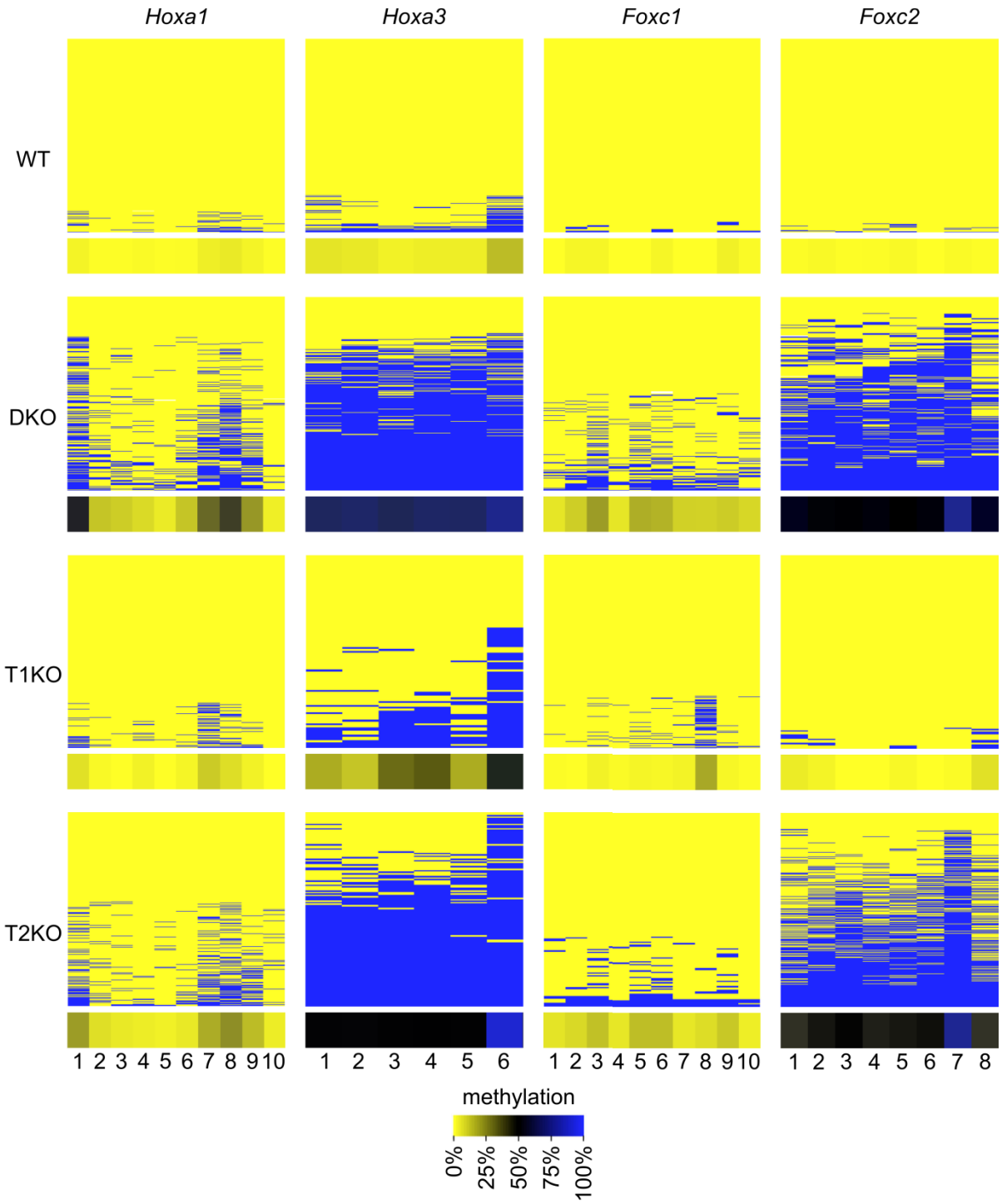

Figure 2.14: TET1 and TET2 cooperate to maintain canyon borders.

Targeted amplicon bisulfite sequencing of canyon borders as in Figure 2.9, comparing WT, *Tet1*$^{-/-}$ (T1KO), *Tet2*$^{-/-}$ (T2KO) and *Tet1/2*$^{-/-}$ MEFs. The small rectangular panels show color-coded average methylation ratio per individual CpG of the amplicon. The color scale is shown below the heatmaps.

2.2 TET function in human hematological disease with *IDH* mutation

Tet2-deficient mice are characterized by impaired differentiation of the hematopoietic lineage, leading to an enlarged stem cell pool and induction of myeloid leukemogenesis over time or together with cooperating mutations (Ko et al., 2011; Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011; Rasmussen et al., 2015). Consequently, *TET2* mutation or downregulation has been recurrently observed in various types of human hematopoietic malignancies (Abdel-Wahab et al., 2009; Delhommeau et al., 2009; Gaidzik et al., 2012; Nibourel et al., 2010; Quivoron et al., 2011; Scopim-Ribeiro et al., 2015; Tefferi et al., 2009). As hypermethylation is a hallmark of many cancers (Jones and Baylin, 2007), it was speculated that TET impairment might be mechanistically related to cancer-specific methylation gain.

Recurrent mutations in *IDH* genes have been associated with genomic hypermethylation or CIMP in different cancer entities (Figueroa et al., 2010; Noushmehr et al., 2010; Wang et al., 2013b). It was then shown that the oncometabolite 2-HG produced by mIDH can inhibit TET enzymes (Koivunen et al., 2012; Xu et al., 2011a). This prompted the hypothesis that *mIDH* causes pathogenic hypermethylation by inhibiting TET-dependent DNA demethylation (Figueroa et al., 2010; Turcan et al., 2012). However, TET activity is only weakly inhibited by 2-HG (Xu et al., 2011a). Moreover, prior analyses of *mIDH*-dependent DNA methylation changes in AML relied on comparisons between unsorted bone marrow aspirates and healthy donor tissues (Akalin et al., 2012; Figueroa et al., 2010). The possible contribution of cell type specific methylation patterns to AML methylomes was neglected, although an mIDH-dependent differentiation block had been repeatedly reported (Figueroa et al., 2010; Losman et al., 2013; Lu et al., 2012; Sasaki et al., 2012b). In order to dissect TET2-, mIDH- and differentiation-dependent methylation changes in AML methylation profiles of AML patients expressing *mIDH* or mutant *TET2* were analyzed and compared to *IDH/TET2* WT AML methylomes. This was supported by methylation profiling of leukemia cell lines expressing *mIDH* and comparisons to various methylomes of normal hematopoietic cell types. The following data are part of a manuscript in preparation by Wiehle et al..

2.2.1 DNA methylation changes in *IDH* mutant AML patients

To investigate the direct effect of *IDH* mutations on the methylome of AML patients, a previously collected patient cohort of adult *de novo* AML was analyzed (Voigt and Reinberg, 2013). The data from this study are publically available and comprise clinical parameters, mutational status, gene expression and DNA methylation profiles of each patient. DNA methylation was assessed using the Illumina Infinium 450K array. This platform monitors the methylation status of 485,577 CpG sites in the human genome and thereby provides

coverage for 99% of annotated RefSeq genes and 96% of CpG islands. The 450K intensity data (IDAT) files containing raw beta values were extracted for 28 *IDH* mutant and 112 *IDH* WT AML patients and normalized, quality-filtered and statistically analyzed. This identified 70,137 significantly (P<0.05) differentially methylated probes between *IDH* mutant and WT AML patients (Figure 2.15A). The vast majority (68,863 probes) of these probes was hypermethylated in *mIDH* carrying patients indicating a pronounced hypermethylation phenotype in AML patients with *mIDH* status. The significantly differentially methylated probes showed a distinct distribution in the two groups of AML patients (Figure 2.15B). While half of the average beta values in *IDH* WT patients ranged from 0.15 to 0.75 with a median of approximately 0.45, this range was increased by 10% in *mIDH* patients with a skewed distribution towards higher beta values (median approximately 0.65). To analyze the genomic distribution of hypermethylation in *mIDH* AML patients, average methylation ratios of significantly changed probes associated with CGI-related epigenomic features were assessed (Figure 2.15C). As expected, CGIs had low average beta values consistent with their largely unmethylated state in mammals. Shores had intermediate methylation levels, whereas shelves and open sea exhibited the highest methylation. *IDH*-mutated methylomes displayed increased average beta values in all the elements, indicating that hypermethylation was widespread and uniformly affecting different genomic regions. However, shores showed the greatest methylation increase compared to the other elements (Figure 2.15C).

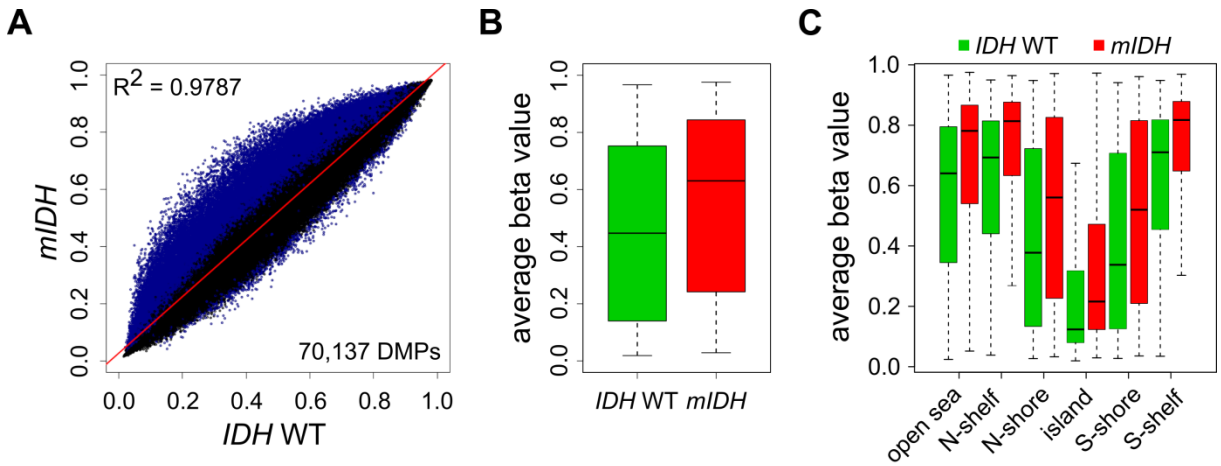

Figure 2.15: AML patients with mutations in *IDH1/2* display genomic hypermethylation.
(A) Scatterplot comparing 28 AML patients with *IDH* mutation to 112 AML patients with *IDH* WT status. Each dot shows the average beta value of an individual CpG probe of the 450K array retained after filtering in the two groups. Blue dots represent significantly (P<0.05) differentially methylated probes (DMPs). **(B)** Boxplot showing the average beta values of the identified significantly differentially methylated probes in the two groups. **(C)** Boxplot showing the average beta values of the significantly differentially methylated probes associated with different epigenomic features in the two patient groups.

To compare methylation profiles of AML patients with mutant and WT *IDH* on a global scale, principal component and clustering analyses were conducted. Standard pre-processing of the array data removed all probes present on sex chromosomes, thereby avoiding the detection of sex-related methylation differences. Principal component analysis based on all 450K probes remaining after pre-processing and quality-filtering revealed a separation of *mIDH* from *IDH* WT samples (Figure 2.16A), indicating the acquisition of characteristic methylation changes upon *IDH* mutation. Hierarchical clustering using the 5000 most significantly differentially methylated probes placed the AML patients into two major groups in a dendrogram (Figure 2.16B). One group comprised mostly *mIDH* patients (red, major left branch), while the other group mainly contained *IDH* WT patients (green, major right branch). The probes used for clustering showed a bimodal distribution. Roughly half of the probes were rather lowly methylated in *IDH* WT patients (reddish boxes; low beta values), whereas the other half displayed high methylation (blue boxes, high beta values), consistent with the typical binary state of methylation observed in mammalia (either methylated or unmethylated). Almost all of these probes gained methylation in the cluster of *mIDH* patients. In summary, these data demonstrate that *mIDH*-associated methylation patterns are sufficient to segregate most AML patients into two classes according to *IDH* mutational status.

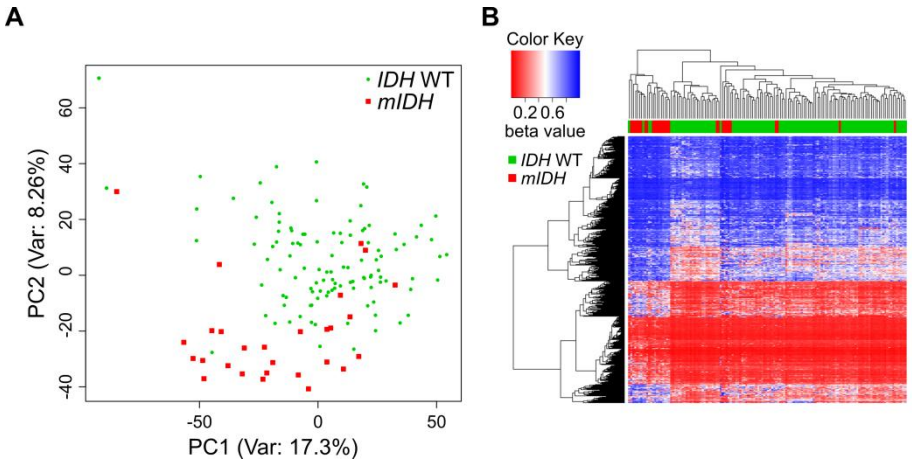

Figure 2.16. *MIDH* confers specific methylation patterns to AML cells.
(A) Principal component analysis of *mIDH* (red) and *IDH* WT (green) AML patients with all 450K CpG probes retained after quality filtering. The first two components covering the dimensions with the highest variance are displayed. **(B**) Heatmap of the 5000 most significantly differentially methylated probes identified between *mIDH* and *IDH* WT AML patients. Dendrograms of patients and probes were obtained using hierarchical clustering by similarity with each column showing one patient and each row one probe. The color scale indicates beta values.

Inspection of the mutational status of *mIDH* patients that did not cluster into their correct group revealed that many of them carried co-occurring mutations in other epigenetic

modifiers such as *DNMT3A*. Of a total of 28 *mIDH* patients 12 had co-occurring mutations in *DNMT3A* with five displaying the R882H substitution that disrupts catalytic methyltransferase activity (Holz-Schietinger et al., 2012). As the lack of the methyltransferase activity of DNMT3A might result in hypomethylation and thus antagonize *mIDH*-associated hypermethylation, the effect of *DNMT3A* mutation on the methylome of AML patients was analyzed. This revealed global hypomethylation in *DNMT3A* mutant AML patients with 23,795 out of 26,334 significantly differentially methylated probes being hypomethylated (Figure S2A), which reduced the median methylation of these patients by about 5% (Figure S2B). The patients with co-occurring *DNMT3A* and *IDH* mutations comprised some of the outliers observed in the PCA in Figure 2.16A and patients with mutations in *DNMT3A* generally appeared to cluster to the right of the area although the group was not well separated (Figure S2C). These findings demonstrate that mutations in epigenetic modifiers do not generally yield genomic hypermethylation. Consequently, patients with mutations in *DNMT3A* were entirely removed from the *IDH* dataset and hierarchical clustering was repeated. This resulted in an improved clustering with most of the patients being assigned to their correct mutational groups (Figure S2D). In conclusion, co-occurring mutations in various epigenetic modifiers can have distinct effects on the AML methylome, which strongly increases the complexity of the observed methylation patterns in AML. In the case of *DNMT3A* and *IDH,* this can explain some of the imperfect clustering by mutational status.

2.2.2 Generation of *IDH* mutant AML cell lines

In order to exclude confounding factors on methylomes of AML patients, such as additional mutations, age-related effects or tumor heterogeneity, and to establish the causality between *mIDH* and hypermethylation, cell-based *(m)IDH*-expressing models were generated. To this end, two leukemia cell lines, NOMO-1 (an acute myeloid leukemia cell line) and HL-60 (an acute promyelocytic leukemia cell line), were lentivirally transduced to stably integrate the WT or mutant *IDH* gene into their genomes. The employed vectors co-expressed GFP or ZsGreen allowing FACS-sorting of successfully transduced and transgene-expressing cell populations. These populations were screened by qRT-PCR for overexpression of the *(m)IDH* genes and 2-HG measurement for production of the oncometabolite D-2-HG.

Expression of *IDH1* WT or *IDH1* R132C controlled by the phosphoglycerate kinase (PGK) promoter (Figure 4.1 for vector details) in HL-60 cells resulted in a modest increase of *IDH1* expression by roughly 2-fold compared to cells transduced with empty vector, while *IDH2* expression levels remained unchanged. Concomitantly, extracellular D-2-HG levels were increased 25-fold to ~10 µM in *mIDH1*-expressing HL-60 cells (Figure 2.17A). Although overexpression of *IDH1* could not be detected by qRT-PCR in NOMO-1 cells transduced with

(m)IDH1, extracellular D-2-HG levels were moderately increased in these cells compared to transduction with *IDH1* WT or empty vector (Figure 2.17B). Conversely, *IDH1* expression dropped upon transduction with WT or *mIDH1* and also endogenous *IDH2* expression levels were reduced. These data indicate that already minor expression of *mIDH1* leads to substantial D-2-HG secretion mimicking the situation in *IDH* mutated AML patients. However, cell lines reacted differently to introduction of *IDH1* enzymes, indicating that transgene expression might be detrimental to certain cell types and thus strategies to downregulate its expression might quickly evolve.

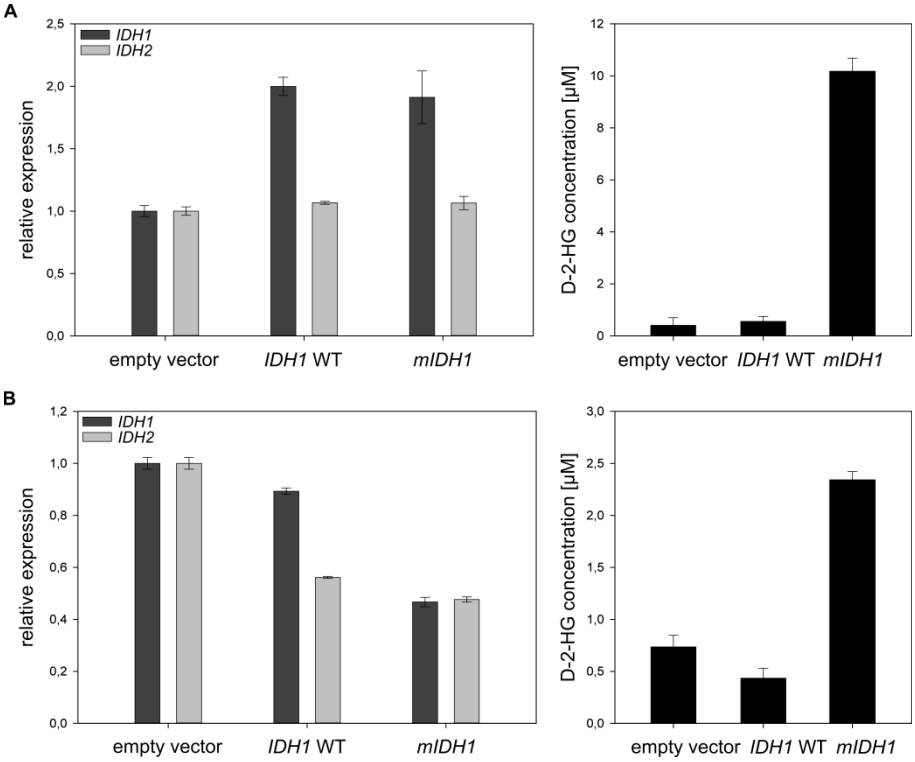

Figure 2.17: Overexpression of *mIDH1* but not WT *IDH1* leads to D-2-HG secretion.
Expression of *IDH1* and *IDH2* transcripts (left panels) and extracellular D-2-HG (right panels) were measured in HL-60 **(A)** and NOMO-1 cells **(B)** transduced with empty vector, *IDH1* WT or *mIDH1*. QRT-PCR was conducted with primers amplifying endogenous as well as exogenous *IDH* transcripts and *ACTB* as reference gene. Bars represent means and error bars standard deviations of three replicates.

The increase of D-2-HG in the medium of NOMO-1 cells transduced with *mIDH1* suggested that mIDH1 was present, although its overexpression was not detectable. In order to confirm this, the base composition at the specific mutated position of the *IDH1* mRNA was examined by cDNA sequencing. The WT base C was detected in NOMO-1 cells transduced with empty vector or *IDH1* WT and a mixture of C and T bases was present in *mIDH1*-transduced cells with the T base comprising the minor fraction (Figure S3A). This

demonstrates the presence of mutated *IDH1* R132C transcripts in NOMO-1 cells. The same was true for HL-60 cells, albeit with roughly equal representation of the WT and mutated base (Figure S3B). This is consistent with the 2-fold increased expression of *IDH1* (Figure 2.17A) and indicates that WT and mutant *IDH1* transcripts were present at similar amounts in HL-60 cells. Since *IDH* mutations are usually heterozygous in AML (Figueroa et al., 2010; Ward et al., 2010), this ratio of expression is likely to be a good approximation of the situation in AML patients.

As primers initially used for qRT-PCR did not distinguish between endogenous and exogenous *IDH* transcripts, potential downregulation of endogenous *IDH* transcription upon introduction of exogenous *IDH1* was neglected. To distinguish these transcripts, qRT-PCR was repeated with primers amplifying specifically the endogenous *IDH1* or the introduced transgene. Endogenous *IDH1* transcript was indeed downregulated upon introduction of *(m)IDH1* compared to empty vector. Expression of exogenous *IDH1* could not be detected in cells treated with empty vector, but was present in cells transduced with WT or *mIDH1* (Figure S3C). Furthermore, determination of intracellular D-2-HG levels in *mIDH1*-expressing NOMO-1 cells revealed a roughly 11-fold increase of 2-HG levels (Figure S3D) compared to empty vector, which strongly exceeded the 4-fold increase in extracellular 2-HG (Figure 2.17B, right panel). These data indicate that NOMO-1 cells expressed introduced transgenes and produced D-2-HG upon expression of *mIDH1*, but responded with downregulation of endogenous *IDH*. Also, they exhibited relatively weak export capacities for 2-HG, leading to modest extracellular D-2-HG concentrations, although substantial 2-HG amounts were intracellularly produced.

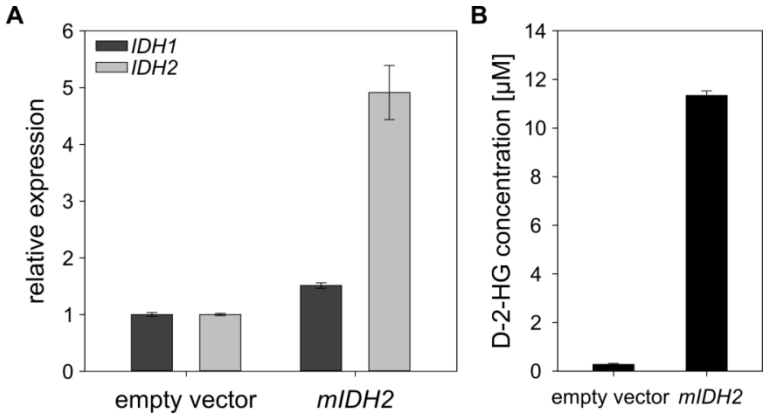

Figure 2.18: Overexpression of *mIDH2* results in robust D-2-HG production.
(A) QRT-PCR of total *IDH1* and *IDH2* in HL-60 cells transduced with empty vector or *mIDH2* (R140Q). *ACTB* was used as reference gene. Error bars show standard deviations (n=3). **(B)** Analysis of D-2-HG concentrations in the medium of HL-60 cells transduced with empty vector or *mIDH2*. Error bars show standard deviations of three replicates.

To similarly express *mIDH2* in leukemia cell lines, *IDH2* R140Q was cloned into a lentiviral construct controlling *IDH* expression by the CMV promoter (Figure 4.1 for vector details). Introduction of *mIDH2* into HL-60 cells increased total *IDH2* levels roughly 5-fold compared to cells transduced with empty vector, while *IDH1* expression remained almost stable (Figure 2.18A). This was accompanied by a more than 30-fold increase in D-2-HG concentrations to approximately 11 µM in the medium of *mIDH2*-expressing cells (Figure 2.18B). This level (~2110 ng/ml) was comparable to the median serum 2-HG concentration found in AML patients with mutant *IDH* (1863 ng/ml; Fathi et al., 2012), indicating that patient-relevant 2-HG concentrations can be obtained with this model. These results furthermore show that both *mIDH1* and *mIDH2* overexpression result in substantial production of D-2-HG in cell-based models.

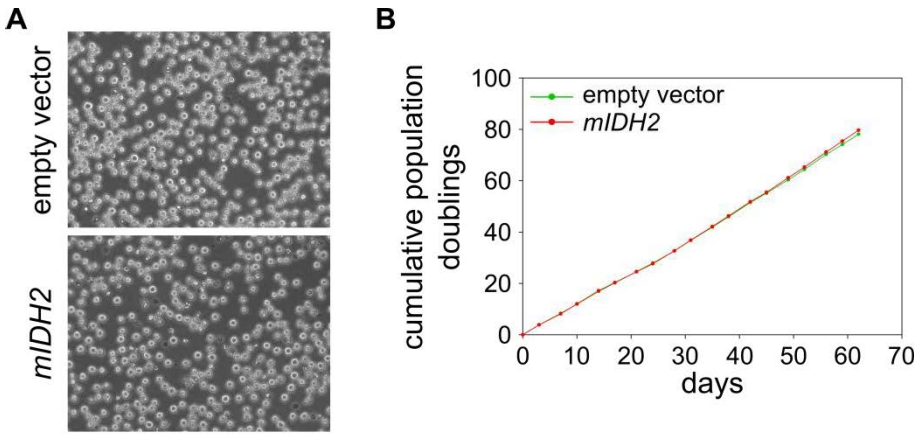

Figure 2.19: *MIDH2*-expressing HL-60 cells show no overt phenotypic alterations.
Bright field microscopy **(A)** and proliferation analysis **(B)** of HL-60 cells transduced with empty vector or *mIDH2*.

In AML *IDH2* mutations are more frequent than *IDH1* mutations and have been shown to produce higher levels of D-2-HG (Ward et al., 2010, 2013). Thus, further analysis focused on *mIDH2*-expressing HL-60 cells. To monitor how *mIDH2* overexpression and accompanying D-2-HG levels would affect cells, phenotypic characterization was carried out. Morphologically, *mIDH2*-expressing cells were indistinguishable from cells transduced with empty vector (Figure 2.19A) and showed similar growth kinetics (Figure 2.19B).

Since mutant IDH enzymes have been reported to inhibit differentiation (Figueroa et al., 2010; Losman et al., 2013; Lu et al., 2012), gene expression of hematopoietic stem cell and myeloid lineage-specific markers were compared between *mIDH2*-expressing cells and empty vector cells. This uncovered a gene expression pattern reminiscent of myeloid progenitor cells. For instance, transcripts specifically expressed in macrophages, monocytes

or granulocytes, such as *CD14* or *CD11B,* were reduced (Figure 2.20A), whereas genes expressed mainly during early stages of hematopoiesis (i.e., in HSCs, MPPs, CMPs and GMPs), such as *MEIS1* or *C-KIT,* were significantly increased in *mIDH2*-expressing cells (Figure 2.20B).

Taken together, in a leukemia cell line model no major phenotypic but moderate gene expression changes are caused by overexpression of *mIDH2,* hinting at an altered differentiation state of the engineered cells.

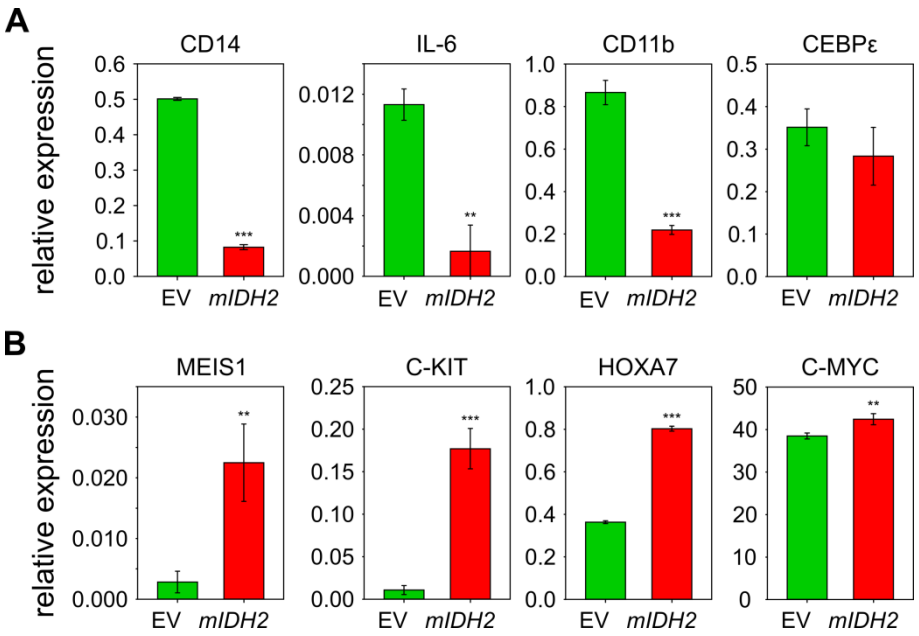

Figure 2.20: *MIDH2*-expressing HL-60 cells display gene expression changes.
QRT-PCR analysis of **(A)** myeloid differentiation and **(B)** hematopoietic stem cell markers in HL-60 cells transduced with empty vector (EV) or *mIDH2*. *ACTB* served as reference gene for qRT-PCR normalization and error bars show standard deviations of three replicates.

2.2.3 DNA methylation changes in an *IDH* mutant AML cell line

The presence of mIDH2 changed gene expression patterns, suggesting that it affects the underlying epigenetic regulatory landscape. To test whether overexpression of *mIDH2* and associated D-2-HG production resulted in changes to the global DNA (hydroxy)methylation level, immunostainings were performed on genomic DNA of transduced HL-60 cells. *IDH2* R140Q-expressing HL-60 cells had increased global 5mC levels compared to cells transduced with empty vector (Figure 2.21A), however 5hmC levels were largely unchanged (Figure 2.21B), indicating that the presence of mIDH2 is associated with a global increase of DNA methylation marks. The stable 5hmC levels suggested that TET activity was not drastically reduced upon *mIDH2* introduction.

To investigate *mIDH2*-associated methylation changes in detail, Infinium methylation analysis of HL-60 cells was conducted. In *mIDH2*-expressing HL-60 cells 59,612 significantly differentially methylated probes were identified compared to cells transduced with empty vector (Figure 2.22B). However, only 65% of these probes were hypermethylated in cells expressing *mIDH2*. Comparison of the two scatterplots from AML patients and HL-60 cells demonstrated that methylation changes in the cellular model were quite distinct from those in AML patients including a wider range of methylation changes and hyper- as well as substantial hypomethylation (compare Figure 2.15A and Figure 2.22A). Calculating the overlap between cells and patients revealed that only a small fraction of all differentially methylated probes was shared between HL-60 cells and AML patients with *IDH* mutations (Figure 2.22B).

Together, these data show that introduction of *mIDH2* into HL-60 cells does not phenocopy the methylation changes in AML patients with *IDH* mutations, indicating that additional mechanisms contribute to the hypermethylation phenotype observed in *IDH*-mutated AML.

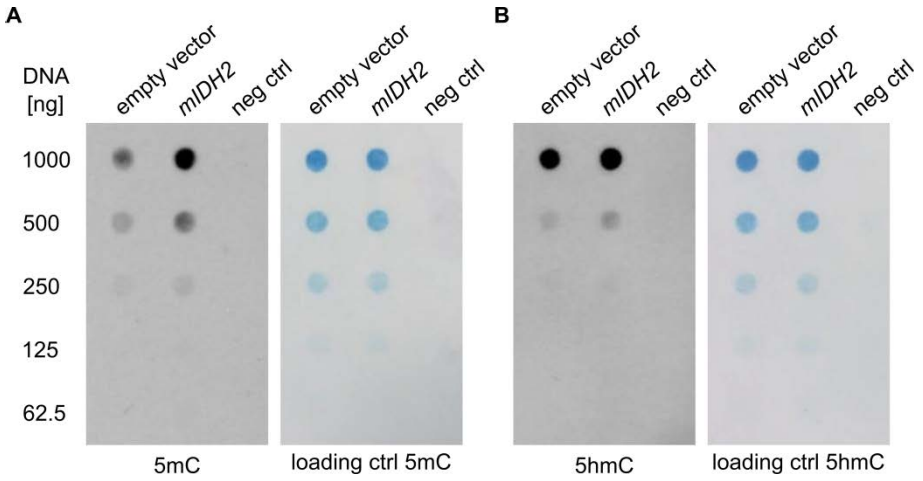

Figure 2.21: *MIDH2*-expressing HL-60 cells display globally increased DNA methylation.
Immunostaining of 5mC **(A)** and 5hmC **(B)** in genomic DNA of HL-60 cells transduced with empty vector or *mIDH2*. Methylene blue staining was used as loading control (loading ctrl). Negative control wells were loaded with buffer (neg ctrl).

PMDs have been reported as large regions of low methylation in human cancers and cell lines, possibly due to inefficient maintenance methylation in quickly proliferating systems (Berman et al., 2011; Gaidatzis et al., 2014; Hon et al., 2012; Lister et al., 2009, 2011). To exclude the possibility that hypomethylation upon introduction of *mIDH2* into HL-60 cells could be caused by the slightly increased proliferation rate of these cells (Figure 2.19B), PMD-associated CpG probes were identified. An approximation of PMD coordinates was

retrieved from publically available nuclear lamina-associated domain profiles (Guelen et al., 2008), which have been shown to coincide with PMDs (Berman et al., 2011). Removal of the PMD-associated probes from the analysis did not affect *mIDH2*-associated methylation changes in HL-60 cells. Despite reducing the total number of significantly differentially methylated probes to 44,015, hypomethylation of a large proportion of probes was retained (Figure S4A). The distribution of average beta values was unchanged compared to the analysis including PMD probes (Figure S4B) and the amount of hypomethylated probes was even increased to 37% (Figure S4C; compared to 35% in the analysis including PMD probes). Hence, loss of methylation from PMDs cannot explain the hypomethylation observed in *mIDH2*-expressing HL-60 cells. This emphasizes the distinct effects of mIDH on the methylomes of AML patients *in vivo* and HL-60 cells *in vitro*.

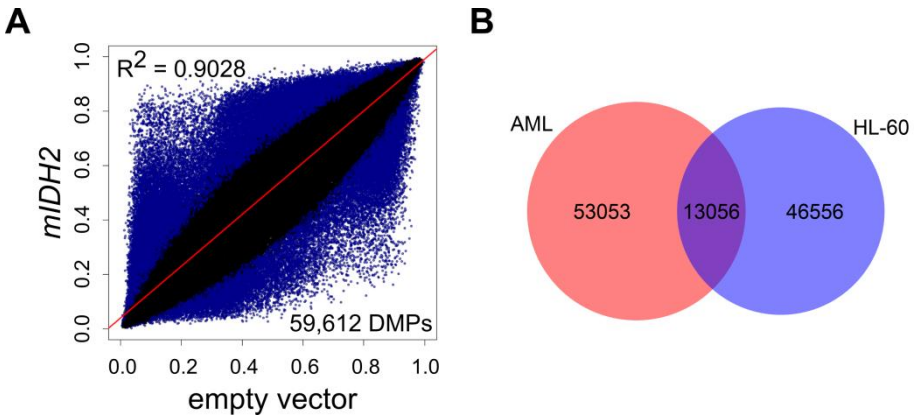

Figure 2.22: DNA methylation changes in AML patients and HL-60 cells with *mIDH* are distinct. **(A)** Comparison of average beta values between HL-60 cells expressing *mIDH2* and empty vector. Each probe is represented in a dot with DMPs (P<0.05) colored in blue. **(B)** Overlap of the significantly differentially methylated probes identified in AML patients and HL-60 cells upon presence of *mIDH*.

2.2.4 DNA methylation changes in AML patients with *TET* mutations

Since neomorphic *IDH* mutations were suggested to inhibit the demethylating activity of TET proteins (Figueroa et al., 2010; Turcan et al., 2012), the loss of *TET* enzymes by inactivating mutations in AML should mimic the methylation changes observed in *mIDH* patients. To address this issue, all patients with mutations in *TET2* were selected from the AML dataset. These 12 patients exhibited exclusively frameshift or nonsense mutations occurring before the catalytic domain, therefore abrogating the dioxygenase activity of TET2. It was previously reported that *TET2* mutations (heterozygous, hemizygous or homozygous) in myeloid cancers correlated with reduced 5hmC levels, suggesting that its hydroxylating activity is impaired even upon heterozygous mutation (Ko et al., 2010). The mutant *TET2* patients were compared to 100 AML patients with *TET2* and *IDH1/2* WT status.

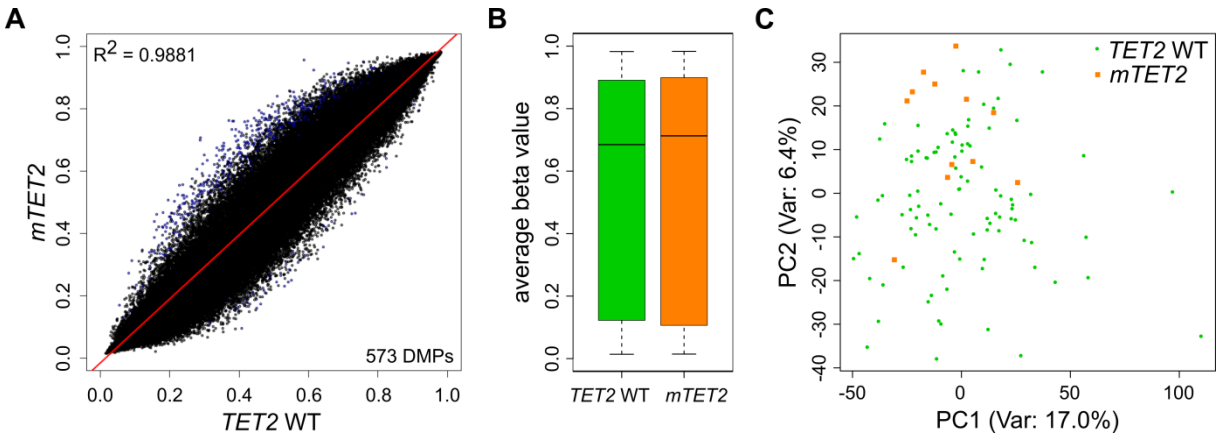

Figure 2.23: Mutations in *TET2* do not recapitulate *mIDH*-associated hypermethylation in AML. **(A)** Comparison of mutant and WT *TET2* AML patients by scatterplot. Each dot shows the average beta value of one probe in the two groups with DMPs (P<0.05) depicted in blue. *IDH* mutant patients were removed from the analysis. **(B)** Average beta values of all probes in the two patient groups. **(C)** Principal component analysis of the methylomes of the two patient groups using all CpG probes left after quality filtering.

This comparison resulted in 573 significantly differentially methylated probes of which 516 were hypermethylated (Figure 2.23A). Consistently, the median of the average beta values of all probes was slightly increased in patients with mutant *TET2* (Figure 2.23B). Principal component analysis based on all CpG probes showed no clear separation of AML patients according to *TET2* mutational status (Figure 2.23C). Taken together, these findings suggest limited hypermethylation upon mutation of *TET2* that does not confer specific discriminatory methylation patterns. Since *mIDH*-associated hypermethylation was much more widespread in AML, mutations in *IDH1/2* and *TET2* do not phenocopy methylation patterns of each other.

2.2.5 Canyon analysis in *IDH* mutant AML patients

In the first part of this thesis, DNA methylation canyons were identified as key targets for demethylation by TET1 and TET2. Since neomorphic *IDH* mutations were reported to inhibit TET function (Figueroa et al., 2010; Turcan et al., 2012), canyon hypermethylation should be an inherent feature of the *mIDH*-associated methylome.

To investigate canyon methylation, a publically available whole genome bisulfite sequencing dataset of an *IDH* WT AML patient was used to map canyons and the corresponding 450K probes. This identified 1711 canyons and 26,117 canyon-associated probes. Average beta values of these probes were only slightly increased in *mIDH* AML patients (Figure 2.24B). When all the canyons were size-normalized and superposed, very

modest hypermethylation was detected in *mIDH* samples. This was evenly distributed over the entire canyon and not enriched at canyon borders, as observed in *Tet*-deficient cells (Figure 2.24C, compare to Figure 2.8B). In conclusion, canyons are not specifically hypermethylated in *mIDH* AML, arguing against a strong inhibition of the demethylating TET activity in the presence of *mIDH*.

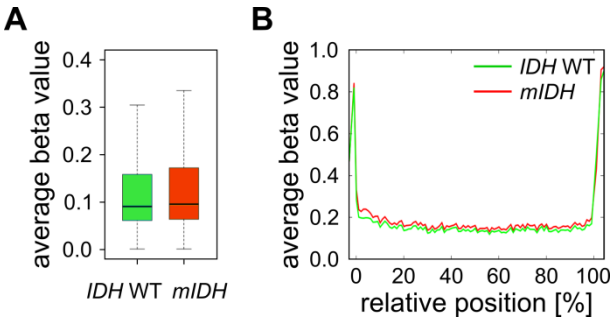

Figure 2.24: TET-dependent DNA methylation canyons are not specifically hypermethylated in *mIDH* AML.

(A) Average beta values of all canyon-associated probes in *IDH* WT and *mIDH* AML patients. **(B)** All canyons identified in whole genome bisulfite sequencing data of an *IDH* WT AML patient were size normalized, superposed and the average beta values of associated probes in *mIDH* and *IDH* WT AML patients were depicted. Canyon analysis was performed by Günter Raddatz.

2.2.6 Effect of D-2-HG on the methylome of an AML cell line

MIDH-dependent TET inhibition was suggested to function by competitive displacement of the essential cofactor α-ketoglutarate by D-2-hydroxyglutarate (Koivunen et al., 2012; Xu et al., 2011a). To reliably inhibit all three TET enzymes and assess resulting DNA methylation changes, WT HL-60 cells were treated with high doses of synthetic D-2-HG. Intracellular D-2-HG quantification confirmed incorporation of the compound into the cells in a concentration-dependent manner, resulting in final concentrations of 10 to 114 pmol/µg protein (Figure 2.25A). However, total intracellular D-2-HG concentrations upon addition of 30 mM dropped over time, possibly due to upregulation of the D-2-HG converting dehydrogenase (D-2HGDH) or cellular transporters that shuttle D-2-HG. Nevertheless, intracellular D-2-HG concentrations at any time point strongly exceeded those obtained by overexpression of *mIDH2* in HL-60 cells (Figure 2.25A, right panel), indicating that this is a valid experiment to model *mIDH*-mediated scenarios in cells and patients.

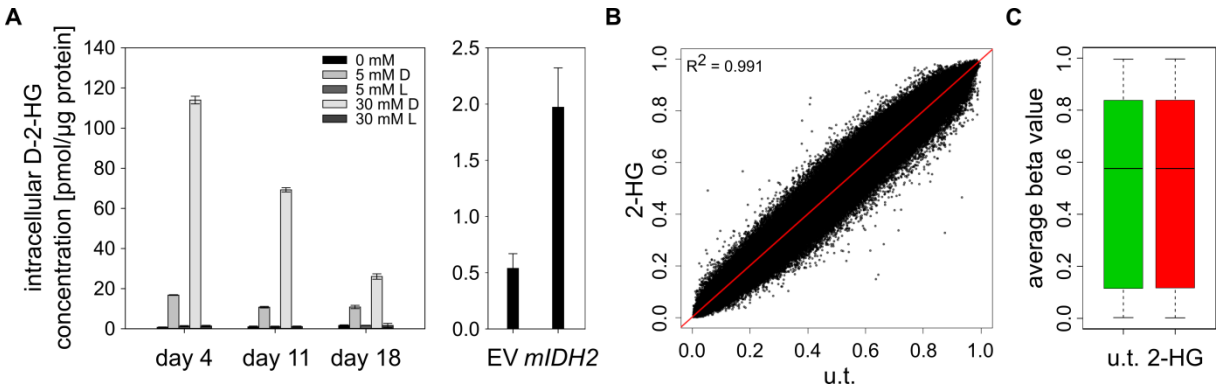

Figure 2.25: D-2-HG does not provoke genomic hypermethylation *in vitro*.
(A) Intracellular D-2-HG quantification in HL-60 cells cultured for 4, 11 and 18 days in medium supplemented with the indicated concentrations of D- or L-2-HG (left panel). On the right the intracellular quantification of D-2-HG in HL-60 cells transduced with empty vector (EV) or *mIDH2* is shown. Error bars indicate standard deviations of three replicates. **(B)** Scatterplot comparing the average beta values of all CpG probes in two replicates of untreated (u.t.) and D-2-HG treated (2-HG) HL-60 cells (30 mM for 21 days). Each dot represents one probe. No significantly changed probes were identified (P<0.05). **(C)** Average beta values of all probes in untreated and D-2-HG treated HL-60 cells.

Analysis of cell proliferation showed that cells grew slower when cultured in D/L-2-HG supplied medium and D-2-HG conferred a bigger growth disadvantage than L-2-HG (Figure S5). However, cells remained viable. After 21 days of treatment with 30 mM D-2-HG genomic DNA was extracted and subjected to DNA methylation analysis by the Infinium chip. Comparison to untreated HL-60 cells demonstrated no significant change in DNA methylation with zero differentially methylated probes upon D-2-HG treatment (P<0.5; Figure 2.25B). Consistently, the distribution of the average beta values of all probes was largely unchanged (Figure 2.25C). Taken together, these data indicate the insufficiency of D-2-HG to induce global DNA methylation changes in cultured cells.

2.2.7 Comparison of *IDH* mutant AML and normal hematopoietic methylomes

Previous studies have shown that *IDH* mutations impair cellular differentiation in hematopoietic and other cell culture systems as well as in hematopoiesis *in vivo* (Figueroa et al., 2010; Losman et al., 2013; Lu et al., 2012; Sasaki et al., 2012b). DNA methylation patterns are dynamic during mammalian differentiation and define cellular identity (Smith and Meissner, 2013; Ziller et al., 2013). Analysis of FACS-purified hematopoietic cell populations with varying differentiation degree revealed that myeloid cells successively reduce their global methylation during commitment and differentiation (Bocker et al., 2011; Farlik et al., 2016; Hodges et al., 2011; Rönnerblad et al., 2014).

To examine the differentiation state of AML cells in patients with and without *IDH* mutation, the distribution of French–American–British (FAB) classes, which were available for each patient and assign the maturation grade to the predominant blast type in AML based on morphological characteristics, were analyzed in the two patient groups. This showed a strong enrichment of FAB category M1 (acute myeloblastic leukemia with minimal maturation) in the group of patients with *IDH* mutations compared to *IDH* WT AML patients (Figure 2.26A). Additionally, undifferentiated acute myeloblastic leukemia (M0) was more frequently observed among *mIDH* AML patients, whereas all the remaining classes occurred more rarely. This indicates a skewed distribution of FAB classes among *mIDH* compared to *IDH* WT AML patients towards less mature cell types. When the assigned FAB category of each patient was marked in the previously conducted principal component analysis (Figure 2.16A), many *mIDH* patients that clustered with the *IDH* WT group were classified as M2 to M5. In contrast, *mIDH* patients clearly segregating from *IDH* WT patients were assigned with M0 or M1 (Figure 2.26B), suggesting that poor maturation determines discriminatory methylation. Consistently, methylomes of blast cells that managed to mature despite carrying *IDH* mutations rather resembled *IDH* WT methylation profiles.

To assess whether *mIDH*-associated hypermethylation in AML cells resembled methylation patterns of normal myeloid cells, comparison analyses between *(m)IDH* AML methylomes and 450K profiles of FACS-sorted human myeloid cell types were performed. Sorted common myeloid progenitors (CMP), granulocyte macrophage progenitors (GMP), promyelocytes (PMC) and polymorphonuclear/terminally differentiated bone marrow granulocytes (PMN) that were previously analyzed by 450K array (Rönnerblad et al., 2014) provided reference methylomes of the major myeloid differentiation stages. Principal component analysis segregated the CD34 positive progenitor cell types (CMP and GMP) from the CD34 negative differentiated cell types (PMC and PMN). When the AML profiles were projected on top of these, *mIDH* AML cases rather clustered with the progenitor cell types, whereas *IDH* WT patients resembled differentiated cell types (Figure 2.26C). Analysis of the 10,000 most differentially methylated probes between CMP/GMP and PMC/PMN populations in hierarchical clustering clearly separated the two groups from each other and placed most of the *mIDH* samples in the same branch as the progenitor cells (Figure 2.26D). Next, the 10,000 most differentially methylated probes between *mIDH* and *IDH* WT patients were extracted and used for hierarchical clustering of myeloid cell types. This ordered the reference cells according to their degree of differentiation with CD34 positive cells separated from CD34 negative cell types (Figure 2.26E). This indicates a pronounced overlap between the *mIDH*-associated methylation signature in AML and the myeloid differentiation-associated methylation program.

In summary, DNA methylation patterns of AML patients with mutations in *IDH* show high similarity to methylomes of myeloid progenitor cell types, identifying an early differentiation state as a source of *mIDH*-associated hypermethylation.

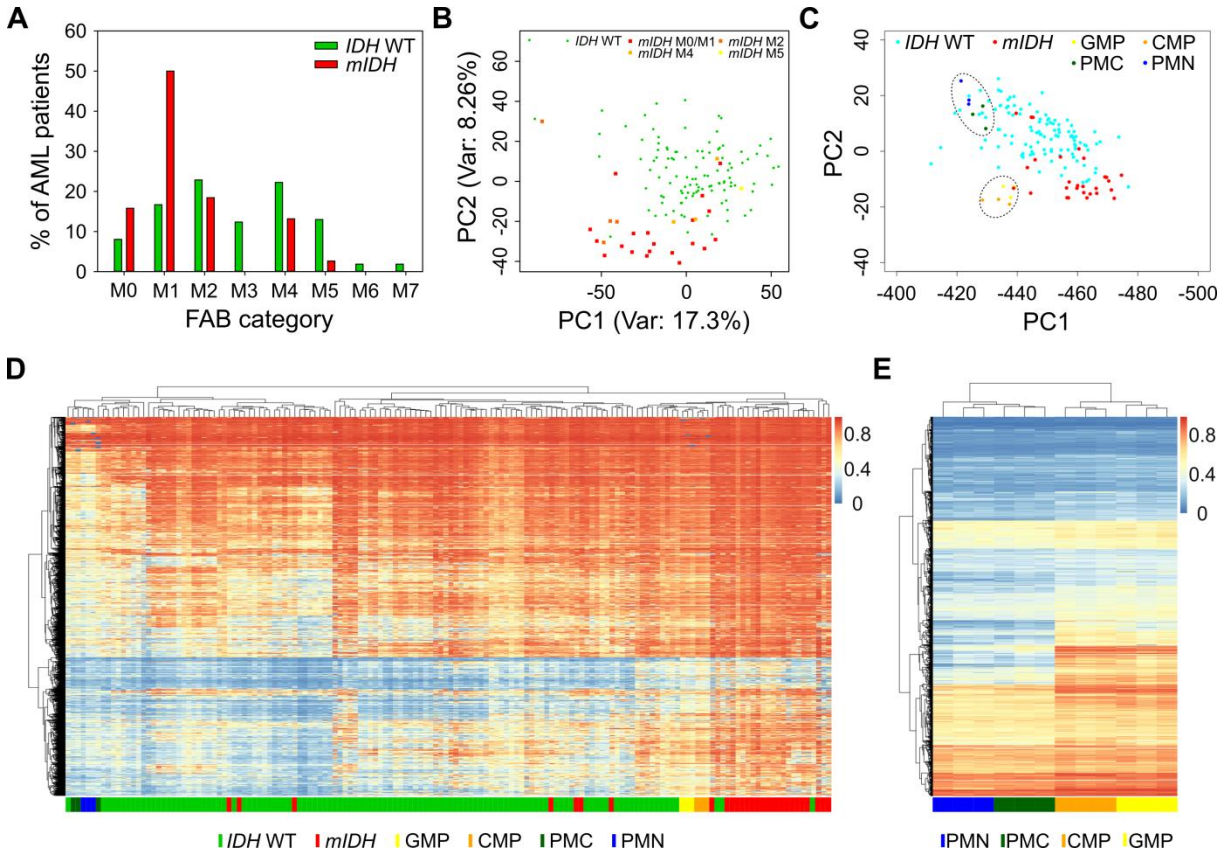

Figure 2.26: *MIDH*-associated hypermethylation in AML resembles methylation patterns of myeloid progenitors.

(A) Frequency distribution of French-American-British (FAB) AML categories in the TCGA cohort divided into *mIDH* and *IDH* WT patients. **(B)** Principal component analysis as in Figure 2.16A with different FAB categories of *mIDH* AML patients marked in orange and yellow hues. **(C)** Principal component analysis using published 450K profiles of four human myeloid cell types from Rönnerblad et al., 2014. AML patient profiles were projected on the cell type clusters. **(D)** Heatmap showing color-coded beta values (color scale to the right) of the 10,000 most differentially methylated probes between the CMP/GMP and PMC/PMN population in AML patients. Each row represents one probe and each column one patient. *IDH* mutational status and reference cell types are indicated below the heatmap by colored boxes. Hierarchical clustering by similarity was used to obtain dendrograms. **(E)** Heatmap of the four hematopoietic reference cell types using the 10,000 most differentially methylated probes between *mIDH* and *IDH* WT AML patients. Each row represents one probe and each column one replicate of the indicated cell type. Analyses for (C) – (E) were performed by Günter Raddatz.

3 Discussion

DNA methylation is dynamically changing during development and differentiation and its deregulation has been associated with cancer pathogenesis (Jones and Baylin, 2007). The interplay of DNMTs, recently identified TET enzymes and their cofactor-providing partners including IDH enzymes defines the cellular methylation landscape. However, the functional contribution of TET enzymes to the methylation profiles of differentiated normal and cancer cells is poorly understood. Here, these two aspects were addressed by characterizing the methylomes of genetically *TET1/2*-deficient primary mouse embryonic fibroblasts and of *IDH* mutated AML patients, in which TET enzymes are chemically inhibited. The findings presented in this thesis identify DNA methylation canyons as key target features of TET1/2-demethylating activity in mouse embryonic fibroblasts and expand on the mechanistic understanding of IDH-associated global hypermethylation in AML.

3.1 Role of TET enzymes in mouse development

TET-mediated DNA demethylation has been implicated in a variety of mammalian developmental processes including the large-scale DNA demethylation waves during PGC and zygotic development and differentiation of ESCs (Wu and Zhang, 2017). A series of recent reports has shown functions of TET enzymes in differentiation and lineage choice of multipotent cells of the adult organism, for example in hematopoietic (Orlanski et al., 2016), neuronal (Perera et al., 2015), hepatic (Ancey et al., 2017) and myogenic (Zhong et al., 2017) cell types. To expand on this knowledge, the biological role of TET enzymes in an embryonic differentiated cell type was analyzed here.

3.1.1 A differentiated model system with impaired DNA demethylation

TET1/2-deficiency resulted in increased 5mC and decreased 5hmC levels in E13.5 mouse embryos, neonates and essentially all analyzed adult DKO tissues (Dawlaty et al., 2013). This indicates impaired DNA demethylation in DKO animals and thus qualifies DKO MEFs – typically isolated at day 13.5 from the embryonic torso – as a convenient differentiated cellular model to investigate epigenomic changes upon 5hmC loss.

TET1 and TET2 are highly expressed in mESCs, the preimplantation blastocyst and the early epiblast (Wu and Zhang, 2014), but they are rapidly downregulated upon *in vitro* differentiation of mESCs, while TET3 emerges (Koh et al., 2011). TET2 and TET3 are the major expressed TET enzymes in a variety of adult tissues (Rasmussen and Helin, 2016). The observed expression pattern of TET enzymes in WT MEFs is in line with these

observations and a previous report (Koh et al., 2011), suggesting that MEFs are in the transition from a stem-like state to an adult fibroblast, where TET1 is still detectable but TET2 and TET3 are likewise expressed. Due to their redundant enzymatic activities TET enzymes in principle can compensate for each other, but TET3 expression was largely unchanged in DKO MEFs. Since global 5hmC in genomic DNA of DKO MEFs was reduced but not absent, TET3 activity most likely accounts for the remaining 5hmC.

3.1.2 Hypermethylated canyons are a key feature of the DKO methylome

To analyze the consequences of impaired DNA demethylation, whole genome bisulfite sequencing of DKO MEFs was performed in this study. Since this approach enables genome-wide detection of methylated cytosines with single-base resolution and high individual CpG coverage, it is currently considered the gold standard of DNA methylation analysis. The resulting average methylation levels were moderately increased in the two DKO replicates, confirming the previously noted genomic hypermethylation upon loss of TET1 and TET2 that was detected by mass spectrometry (Dawlaty et al., 2013).

In order to identify functionally important genomic elements that may be TET-dependently demethylated, it was reasoned that these would accumulate aberrant DNA methylation in *TET1/2*-deficient compared to WT MEFs. An unbiased analysis of the methylation of different genomic features identified enhancers, promoters, transcribed gene bodies, insulators and intragenic regions as most severely hypermethylated. Consistent with this, enhancer hypermethylation has been described in *TET2*-deficient (Hon et al., 2014) and TKO mESCs (Lu et al., 2014). Both studies also provided evidence that promoters, gene bodies and CTCF sites (insulators) are affected by TET-dependent hypermethylation. Another study mapped regions of 5hmC loss upon TET1 or TET2 depletion in mESCs and found 5hmC reduction at TSS and exons of highly expressed genes, respectively (Huang et al., 2014). These findings match the observations presented here and suggest that regions losing 5hmC are overlapping with regions accumulating 5mC upon TET-deficiency, arguing for an active DNA demethylation process at these regions.

The identification of DNA methylation canyons as conserved, hypomethylated features of mammalian genomes whose borders are enriched with 5hmC (Jeong et al., 2013) prompted their analysis in this study. Here, DNA methylation canyons were found as a key feature of hypermethylation upon *TET1/2*-deficiency. While a subset of canyons was lost in DKO cells, remaining canyons displayed hypermethylation. Overall, two thirds of all identified canyons collapsed or diminished in size, strongly suggesting that TET1/2 activities are required to maintain canyons. In particular canyon borders were affected by hypermethylation as confirmed by targeted bisulfite sequencing. Data from (hydroxy)methylated DNA

immunoprecipitation experiments reported in Wiehle et al., 2016 corroborate these findings, as 5mC was increased at canyon borders, while 5hmC was reduced.

Conversely, a quarter of all identified canyons increased in size upon TET1/2-deficiency. It has been reported that canyon borders also depend on the activity of DNMT3A and become eroded upon deletion in mouse HSCs (Jeong et al., 2013). Thus, defective targeting, reduced expression, activity or stability of DNMT3A upon TET1/2 loss may account for expanded canyons. More recently, it was also suggested that TET2 and DNMT3A, which are frequently co-mutated in leukemia, have context-dependent functions on DNA methylation in mouse HSCs, that could be synergistic or competitive (Zhang et al., 2016b). This analysis of TET2 and DNMT3A double deficient HSCs supports the model of opposing activities of DNMT3A and TET2 at the edges of canyons presented in Figure 3.1. Moreover, the authors identified redistribution of global DNA methylation upon double KO and TET2 KO specific hypermethylation in the center of canyons. Similar observations were made in the present study, for example in the anterior HOXA cluster (Figure 2.8A). Although no mechanistic explanation on how synergistic activities may be elicited is provided in that study, expanded canyons could represent such regions of TET and DNMT cooperation or methylation redistribution. Furthermore, ectopically targeted TET3 may exert compensatory demethylation at certain canyons.

The analysis of single KO MEFs indicated that combined cooperative activities of TET1 and TET2 protect the canyon borders from hypermethylation, but with a major contribution of TET2. This is in line with other reports showing that TET1 and TET2 appear together in different cell types and seem to have overlapping and non-redundant functions under certain conditions, for instance in HSCs (Zhao et al., 2015), regulatory T-cells (Yang et al., 2015) and during NANOG-assisted somatic cell reprogramming (Costa et al., 2013). Nevertheless, specific roles for TET1 and TET2 have been reported as well. For instance, in mESCs with *TET2*-deficiency a particular loss of 5hmC from gene bodies and other regulatory elements was reported, whereas *TET1*-deficiency resulted in a general 5hmC reduction with global 5hmC patterns reminiscent of WT ESCs (Hon et al., 2014). Another study showed TET1-dependent 5hmC depletion specifically from TSS (Huang et al., 2014). Similar to the presented findings, 5hmC loss and 5mC gain were generally greater upon *TET2*-deficiency compared to *TET1* loss. Hence, TET enzymes may share their demethylation targets in certain cell types or developmental stages, while having specific roles in others. Moreover, overlapping functions might be a relict of an ancestral single TET gene, while individual TET enzymes may have acquired additional specific functions during evolution. Particularly in unstressed differentiated cells, where transcriptional and methylation patterns are largely established, TET enzymes might serve a broad role in the maintenance of normal methylation patterns rather than directed DNA demethylation of specific features.

The findings presented here further advance on previous publications implicating TET1 in fine-tuning of the physiological DNA methylation landscape. TET1 was found to bind specifically CpG-rich TSS, where 5hmC is also enriched in mESCs. It was thus speculated that TET1 might mediate DNA methylation fidelity by opposing stochastic hypermethylation at CpG-rich gene regulatory sequences (Williams et al., 2011). Similarly, a study that used HEK293T cells proposed that TET1-mediated oxidative 5mC conversion is essential at hypomethylated CGIs to prevent aberrant invasion of DNA methylation (Jin et al., 2014). The present study attributes a shared role for TET1 and TET2 in the maintenance of normal methylation patterns and identifies significantly larger DNA methylation canyons as TET-dependent epigenomic features.

3.1.3 A role for TET1/2-dependent canyon maintenance in differentiation

Canyon hypermethylation upon *TET1/2*-deficiency was associated with severe differentiation and gene regulation defects, suggesting functional relevance for the maintenance of canyon integrity.

Firstly, *in vitro* adipogenesis of DKO MEFs was inhibited as evident by the failure to generate lipid droplets and upregulate key adipogenic transcription factors and effector genes. Secondly, the promoters of canyon-associated genes with reported functions in adipogenesis acquired even more pronounced hypermethylation during *in vitro* adipogenic differentiation of DKO MEFs compared to untreated cells. Finally, TET transcripts were induced during adipogenesis in WT MEFs. Together, these data suggest a functional requirement for TET enzymes in adipogenesis that may be related to the maintenance of canyons associated with adipogenic regulators and their promoters. Indeed, the newest literature confirms these observations showing that TET enzymes are required for adipogenesis (Yoo et al., 2017) and activation of essential genes, such as PPARγ (Dubois-Chevalier et al., 2014; Yoo et al., 2017) or PRDM16 (Yang et al., 2016b). Mechanistically, several explanations have been reported, such as TET-mediated hydroxylation of enhancers directed by an interaction with CTCF (Dubois-Chevalier et al., 2014) or recruitment of TET1 to PPAR response elements by PARylated PPARγ (Fujiki et al., 2013). Although the findings presented herein cannot explain how TET enzymes may be attracted to their target regions, they strongly suggest that the lack of canyon demethylation by TET1 and TET2 is responsible for the adipogenic differentiation defects. The four canyon-associated genes, whose promoter methylation was assessed in this study, have reported roles in adipogenesis. PPARγ is an essential key adipogenic transcription factor (Rosen and MacDougald, 2006), IGF1 initiates IGF-signaling that is involved in staging of the differentiation process (Tang and Lane, 2012) and FOXC1 and 2 are transcription factors that have been associated with the maintenance

of the adipose tissue phenotype (Davis et al., 2004; Omatsu et al., 2014; Rosen and MacDougald, 2006). Especially in the absence of PPARγ adipogenesis cannot take place (Rosen et al., 1999), supporting the notion that defective transcriptional PPARγ activation in DKO MEFs may be causative for the differentiation block.

In this study, deficiency of *TET1* and *TET2* was further accompanied by broad gene deregulation. As 95% of canyons are associated with transcriptional start sites, canyon hypermethylation is likely to influence gene expression through aberrant promoter silencing. Hundreds of differentially expressed genes were present in DKO MEFs and there was a statistically significant overlap between canyon hypermethylation and deregulation of associated genes. This correlation suggests a requirement of canyon maintenance for proper gene expression. Nevertheless, only half of the genes associated with a hypermethylated canyon were deregulated and the majority of deregulated genes were not canyon-associated. These genes might either depend on smaller hypomethylated regions potentially protected by TET enzymes (e.g. reported insulator sites) or other described TET1/2 activities. For example, a study by Hon et al. showed that *TET2* KO in mESCs leads to loss of 5hmC and concomitant gain of 5mC at enhancers. This results in decreased enhancer activity, which correlated with transcriptional delay of lineage-specific genes that are normally induced during neuronal *in vitro* differentiation (Hon et al., 2014). On the other hand, the canyon-associated genes whose expression remained unchanged may be not severely affected by hypermethylation or protected from hypermethylation by active chromatin states. Developmental regulators residing in large hypomethylated domains are usually regulated by chromatin modifications (Jeong et al., 2013; Nakamura et al., 2014; Xie et al., 2013). Indeed, additional bioinformatical analyses documented in Wiehle et al., 2016 suggested that canyons marked by repressed or poised chromatin states are prone to become hypermethylated. In these canyons transcriptional regulation by chromatin states may be disrupted, whereas more active regions may be resistant to hypermethylation. Consistently, the genomic localizations of TET1 and 5hmC appear to significantly overlap with Polycomb-repressed genes (Williams et al., 2011; Wu et al., 2011a, 2011b).

A substantial proportion of genes was also upregulated in DKO MEFs and thus *TET*-deficiency seems to elicit context-dependent transcriptional changes. On the one hand, hypermethylation of transcribed gene bodies, intergenic regions and enhancers observed in DKO MEFs might be responsible for this effect. On the other hand, dual functions for TET1 in transcriptional regulation in mESCs have been reported previously and may also explain the opposing transcriptional effects observed in MEFs. Derepression upon *TET1/2*-deficiency could thus be due to the reported function of TET1 in facilitating PRC2 complex binding to DNA and repressing PRC2 target genes (Wu et al., 2011b). Moreover, TET1 is involved in the

recruitment of the SIN3A co-repressor complex and silencing of its target genes (Williams et al., 2011).

Canyons appear to harbor TSS of many developmentally important regulators including genes involved in transcriptional regulation, embryonic morphogenesis and RNA metabolism (Jeong et al., 2013). In accordance with this, pathway analysis of deregulated genes in DKO MEFs revealed the enrichment of exclusively developmental categories. A prominent example of a hypermethylated canyon is the one at the *HOXA* cluster (Figure 2.8). Consequently, *HOXA3* was significantly and strongly downregulated in DKO MEFs. HOX transcription factors have important roles during early embryonic development and hence their deregulation may account at least in part for the developmental defects observed upon TET DKO *in vivo*. Apart from that, genes with roles in mesenchymal-to-epithelial transition (MET) were deregulated, among them FOXC1 and FOXC2 (Bard et al., 2008; Hader et al., 2010; Thiery et al., 2009). Since these genes are also canyon-associated, their defective regulation may contribute to the reported block in MET during somatic cell reprogramming of TET TKO MEFs (Hu et al., 2014). While the authors considered the defective demethylation and activation of key miRNAs of the miR-200 family responsible for this effect, a potential interplay of these miRNAs with canyon-associated genes that are involved in RNA metabolism is also conceivable.

In summary, the findings from this study strongly suggest that canyon-associated differentiation genes depend on the TET1/2-mediated protection of canyons from hypermethylation. The promoters of canyon-associated developmental genes rely on epigenetic regulation during differentiation. It is proposed that disruption of canyon integrity prevents proper regulation and activation of differentiation genes upon environmental cues. Together with the previously published data on DNMT3A activity at canyons, the balanced antagonizing functions of TET1/2 and DNMT3A may keep DNA methylation canyons in their physiological state (Figure 3.1).

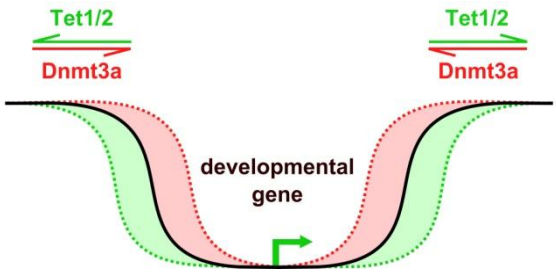

Figure 3.1: Model for the proposed actions of TET1, TET2 and DNMT3A on canyons.

3.1.4 A potential role of canyon hypermethylation in cancer

TET2 and DNMT3A are among the most frequent, recurrently mutated genes in human clonal hematopoiesis and AML (Voigt and Reinberg, 2013; Welch et al., 2012). Canyon-associated genes have been observed to be significantly enriched for genes that are dysregulated in leukemia (Jeong et al., 2013). Hence, cancer cells with mutations in the genes responsible for canyon maintenance may acquire canyon hypermethylation. In turn, this could facilitate the induction of a cancerous gene expression signature because canyon genes appear to have crucial roles in transcriptional regulation. In this study, not only downregulation of *HOXA3,* but also increased expression of various genes of the *HOXB* and *HOXC* clusters was observed in DKO MEFs. *HOX* genes have an appreciated oncogenic role in leukemia (Alharbi et al., 2013). Similarly, the TET-dependent block in MET (Hu et al., 2014) may sustain acquired migratory potential of cancer cells with canyon hypermethylation. In addition, differentiation as a rescue mechanism of aberrantly dividing pre-cancerous stem cells may be impaired by canyon hypermethylation. *TET1* downregulation has been observed in aging (Bormann et al., 2016) and replicative senescence (Sakaki et al., 2017). As such, canyon hypermethylation could contribute to the predisposition of aged cells to cancer development, whereas in senescent cells canyon maintenance may become dispensable. Lastly, it was shown recently that 5hmC marks promoters in matched normal tissue that are resistant to hypermethylation in colorectal cancer. These regions are also occupied by TET2 in colorectal cancer cell lines (Uribe-Lewis et al., 2015). Taken together, these findings suggest that TET1 and TET2 are important to prevent promoter hypermethylation in cancer and TET impairment may participate in the establishment of the cancer epigenome. Further studies will be required to analyze canyon deregulation in cancer.

3.2 Role of IDH-TET axis in human disease

The association between mutations in *IDH* and genomic hypermethylation has been noted first in glioma (Noushmehr et al., 2010) and the paradigm of the causal relationship between the two was established by the study of Turcan and colleagues (Turcan et al., 2012). Mechanistically, TET enzymes became implicated in this pathway soon after their discovery by the analysis of mutational and DNA methylation patterns in AML as well as biochemical studies (Figueroa et al., 2010; Koivunen et al., 2012; Xu et al., 2011a). However, these employed low coverage assays for DNA methylation analysis, indirect detection methods for 5hmC and 5mC in complex overexpression models and partially questionable comparisons. Thus, the nature of mIDH-associated epigenome deregulation remains poorly understood. To advance its understanding, the influence of *IDH* mutations on TET enzymes and the methylome in AML was analyzed in the second part of this thesis.

3.2.1 Genomic hypermethylation is a key feature of AML with mutations in *IDH*

Methylomes of AML patients with mIDH were characterized by a pronounced large-scale genomic hypermethylation. Approximately 70,000 CpG residues were significantly differentially methylated in *mIDH* carrying AML with >98% being hypermethylated. These findings are consistent with earlier reports that found hypermethylation in *IDH*-mutated AML (Figueroa et al., 2010), glioma (Noushmehr et al., 2010; Turcan et al., 2012), chondrosarcoma (Lu et al., 2013), enchondroma (Pansuriya et al., 2011) and cholangiocarcinoma (Wang et al., 2013b) and thus confirm the association between *IDH* mutation and hypermethylation.

Of the differentially methylated probes, 50% had intermediate methylation ratios in *IDH* WT AML (beta values approximately 0.2 – 0.8), indicating that regions of variable methylation are affected by the *mIDH*-associated gain of methylation. This was corroborated by the fact that, although hypermethylation was widespread among all kinds of analyzed epigenomic features, the methylation gain was most pronounced in CGI shores. Consistently, it was reported previously that most cancer-associated methylation changes take place in CGI shores (Irizarry et al., 2009).

CIMP has been defined in colorectal cancer as the coordinated cancer-specific acquisition of hypermethylation at specific gene-associated CGIs in a subset of cancers, instead of a stochastic gain in CGI methylation frequency (Toyota et al., 1999; Weisenberger et al., 2006). In glioblastoma, a distinct subset of the proneural tumors, which are identified by a specific gene expression pattern, exhibits coordinated CGI hypermethylation reminiscent of the changes found in colorectal cancer and thus has been associated with glioma-specific CIMP (G-CIMP; Noushmehr et al., 2010). The extremely high concordance between G-CIMP phenotype and *IDH* mutation has led to the assumption that *mIDH* is causally linked to CIMP establishment (Noushmehr et al., 2010; Turcan et al., 2012). However, it is not clear to which extent mIDH-associated hypermethylation targets exclusively CGIs, due to the use of non-genome-wide methods to call methylation. Further, the influence of cancer type on the methylome upon *IDH* mutation remains elusive. Turcan and colleagues reported that G-CIMP-defining CpGs were located in CGIs and shores, but on the Infinium 450K array the majority of CpG probes are devoted to assess CGI and shore methylation (Sandoval et al., 2011). Other studies made use of (enhanced) reduced representation bisulfite sequencing and found mainly CGIs hypermethylated in the presence of mIDH (Akalin et al., 2012; Lu et al., 2013), however these techniques rely on restriction enzymes that cut more frequently in CpG-rich regions and are thus biased to CGIs as well (Meissner et al., 2005). The results presented here rather suggest no specific hypermethylation of CGIs even though the same experimental strategy (450K array) was used. Consistently, a recent report showed that

mIDH-associated hypermethylation favored non-CGI regions in AML using an alternative method. Interestingly, this study also identified an AML hypermethylation signature that was enriched at CGIs but independent of *IDH* mutations, which was termed A-CIMP (Kelly et al., 2017). These results argue against sufficiency of mIDH to establish CIMP independently of tissue type. In summary, the term CIMP appears to represent a conglomerate of miscellaneous hypermethylation signatures. Thus, it remains to be clarified which are the main targets of IDH-dependent hypermethylation, preferably by using comprehensive genome-wide methods. Ultimately, this should allow discriminating and comparing mIDH-dependent methylation signatures of distinct cancers.

To establish the causality between *mIDH* expression and hypermethylation, a leukemia cell line overexpressing *mIDH2* was generated. Although these HL-60 cells acquired various methylation changes, the AML-specific hypermethylation pattern could not be faithfully recapitulated. Whereas the overall number of differentially methylated probes was comparable between mIDH AML patient samples and HL-60 cells, approximately one third of all significantly changed probes (20,905) displayed hypomethylation in the cellular model. Consistently, a very similar distribution of hyper- and hypomethylation was observed upon introduction of *mIDH2* into TF-1 erythroleukemia cells and *mIDH1* into primary astrocytes, respectively, in previous studies (Kernytsky et al., 2014; Turcan et al., 2012). Hypomethylation of lamina-associated domains reported to occur during repeated passaging of cells *in vitro* (Gaidatzis et al., 2014) could be excluded as a putative cause of methylation loss. Together, these data suggest dual effects of mIDH in cultured cells eliciting hyper- and hypomethylation. It is conceivable that the unique mIDH-associated hypermethylation in AML patients is caused by additional mechanisms such as increased activity of DNMT3A during AML progression, that has been recently reported to result in CGI hypermethylation (Spencer et al., 2017). The influence of the tumor microenvironment on cancer cells, specifically mesenchymal stromal cells in the case of hematological malignancy, can hardly be modeled in cell culture, but has been recognized *in vivo* (Geyh et al., 2013). Recently, D-2-HG was shown to affect surrounding cells in a paracrine fashion (Chaturvedi et al., 2016; Chen et al., 2016) and may thus change the interplay between microenvironment and cancer cells eventually translating into epigenomic changes.

3.2.2 *MIDH*-associated methylomes and their relation to TET inhibition

The inhibition of TET activity by D-2-HG produced by the neomorphic mIDH enzymes has been suggested as the mechanistic basis for the observed hypermethylation. This was due to the observation that IDH and TET2 mutations are mutually exclusive in AML, but display overlapping hypermethylation signatures (Figueroa et al., 2010).

In this study, large-scale *mIDH*-associated hypermethylation did not recapitulate DNA methylation changes associated with *TET2* mutations in AML, nor did it affect specifically TET-dependent DNA methylation canyons. Furthermore, D-2-HG did not induce genome-wide DNA methylation changes in a cell-based model. Together, these findings challenge the currently established paradigm of mIDH-dependent TET inhibition.

Several details from the literature support this notion. Firstly, clinical differences have been reported between *TET2* and *IDH* mutant hematopoietic malignancy, indicating that the oncogenic mechanisms of each mutated enzyme may not be identical. For instance, *TET2* mutations are more frequent in premalignant MDS compared to AML, whereas this is reversed in the case of *mIDH1* (Im et al., 2014), indicating that mIDH may be involved in the progression to more aggressive disease. Furthermore, the prognosis for *TET2* and *IDH* mutated AML appears distinct with adverse outcomes of *TET2* mutated, but favorable prognosis for *IDH2* R140Q carrying AML patients (Patel et al., 2012). Secondly, the phenotypes of *TET2*-deficient and *mIDH*-expressing mice are distinct. *TET2*-deficiency resulted in pronounced expansion of the LSK cell pool, increased repopulation capacity, impaired or skewed differentiation and the development of myeloid malignancies (Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011). In contrast, *mIDH*-expressing cells had an augmented pool of lineage-restricted progenitors, but reduced numbers of long-term HSCs (Inoue et al., 2016b) and normal repopulating activity (Sasaki et al., 2012b). Lastly, the consequences of *TET2* mutation with regard to methylome changes in myeloid malignancies are controversial. Thus, it remains unclear whether reduction of TET activity (be it triggered genetically or chemically) indeed yields hypermethylation. While some studies identified a number of hypermethylated CpGs in *TET2*-mutated AML and CMML (Figueroa et al., 2010; Rasmussen et al., 2015; Yamazaki et al., 2012, 2015), others found hypomethylation upon *TET2* mutation in CMML and a range of myeloid diseases (Ko et al., 2010; Pérez et al., 2012). A recent study depleted erythroid cells of *TET2* or *TET3* but could not detect significant changes in global 5mC levels (Yan et al., 2017). None of these studies reported significant hypermethylation at tens of thousands of CpGs upon TET-deficiency in agreement with the presented findings, indicating that *mIDH* and mutated *TET2* elicit distinct methylome changes.

In vitro enzymatic studies demonstrated that D-2-HG can inhibit TET enzymes (Koivunen et al., 2012; Xu et al., 2011a), albeit only achieving partial inhibition with a 500-fold excess of D-2-HG over α-ketoglutarate. Hence, D-2-HG is a weak TET inhibitor at the most. This would be consistent with the fact that the dotblot immunostaining in HL-60 cells expressing *mIDH2* did not detect changes in 5hmC levels. Either this method is not sensitive enough to detect small changes in 5hmC or the TET inhibition under these conditions is too weak. Arguing for the latter, in the present study extracellularly supplied D-2-HG was not

sufficient to elicit DNA hypermethylation in HL-60 cells, although it penetrated the cell membrane and accumulated to intracellular concentrations exceeding the ones obtained with overexpression of *mIDH2*. Consistent with the observed uptake of D-2-HG into the cells, paracrine activity of D-2-HG was recently reported (Chaturvedi et al., 2016). Consequently, the incorporated D-2-HG should be able to inhibit TET enzymes. However, the absence of significant gains in DNA methylation upon D-2-HG-supplied cell culture contradicts D-2-HG-mediated impairment of TET activity. Recent findings from IDH WT AML mouse models receiving D-2-HG are in support of this notion: Although the metabolite promoted leukemia onset, it was insufficient to induce myeloproliferation or leukemia on its own. Furthermore, mIDH exhibited stronger oncogenic potential and its overexpression changed DNA methylation patterns more drastically than D-2-HG administration, indicating additional D-2-HG-independent roles of mIDH in leukemogenesis and epigenome deregulation (Chaturvedi et al., 2016). Interestingly, increased proliferation, cytokine independence and an *in vitro* differentiation block of TF-1 cells upon expression of *mIDH1* could be mimicked by culturing the cells in D-2-HG, suggesting that 2-HG can confer cancer-promoting features. However, these effects were not elicited by L-2-HG, although it has been shown to inhibit TET enzymes more efficiently than D-2-HG (Koivunen et al., 2012; Losman et al., 2013; Xu et al., 2011a). This argues against a major contribution of TET enzymes and DNA methylation to the cancerous phenotype. Interestingly, recent preclinical tests evaluating small molecule inhibitors specific to mIDH also reported controversial results. Albeit blocking 2-HG production, reducing growth and promoting differentiation of cancer cells, in some studies these inhibitors were unable to restore normal DNA methylation levels of glioma xenografts and primary AML cells (Chaturvedi et al., 2017; Pusch et al., 2017; Rohle et al., 2013), suggesting 2-HG-independent deregulation of DNA methylation by mIDH or the irreversibility of the acquired changes. Nevertheless, another publication claimed the partial reversal of hypermethylation in *mIDH*-overexpressing cells that were treated with an mIDH inhibitor (Kernytsky et al., 2014).

Ultimately, D-2-HG-mediated TET inhibition appears unlikely to be the sole cause of mIDH-associated genomic hypermethylation. Further clarification could be provided by the analysis of patients or models with mIDH expression but low D-2-HG secretion or different catalytic *IDH* mutations that can neither produce α-ketoglutarate nor D-2-HG.

3.2.3 *MIDH*-associated methylomes in view of hematopoietic differentiation states

Mutations in *IDH* have been associated with a differentiation block using *in vitro* and *in vivo* models. For example, expression of *mIDH1* from its endogenous locus in the myeloid lineage of the mouse resulted in increased numbers of early hematopoietic progenitors (Sasaki et al.,

2012b). Furthermore, *in vitro* differentiation of TF-1 erythroleukemia cells and 3T3-L1 cells was blocked by the expression of *mIDH1* or *mIDH2*, respectively (Losman et al., 2013; Lu et al., 2012), and primary mouse bone marrow cells transduced with *mIDH2* showed depletion of myeloid markers and elevation of LSK cell markers (Figueroa et al., 2010). In line with this, the *mIDH2*-expressing HL-60 cells generated here upregulated several hematopoietic stem cell markers and conversely reduced the expression of myeloid genes. Furthermore, the fact that mIDH was specifically enriched in the poorly differentiated M1 FAB category may reflect the inhibition of differentiation *in vivo*. Indeed, methylomes of *mIDH* carrying AML patients displayed high similarity to those of early myeloid progenitor cell types (CMPs and GMPs) in contrast to *IDH* WT AML. Together with the TET independency of mIDH-associated changes discussed in the previous chapter, this study identifies differentiation state rather than TET inhibition as a critical factor responsible for hypermethylation.

The differentiation cascade of myelopoiesis is characterized by an overall loss of methylation compared to hematopoietic stem cells (Figure 1.5; Bock et al., 2012; Farlik et al., 2016; Hodges et al., 2011; Rönnerblad et al., 2014). The analysis of raw unsorted bone marrow aspirates of AML patients may underestimate considerable intra- and inter-individual heterogeneity of samples. Therefore, the comparison of mutational subgroups may in fact be a comparison of distinct differentiation stages. In the case of mIDH, this may result in the detection of hypermethylation as mIDH could arrest hematopoietic cells in an early progenitor state, whereas IDH WT AML cells could mature further. Indeed, the findings presented here suggest the undifferentiated state of the AML cells as a cause for the mIDH-associated hypermethylation (Figure 3.2). In agreement, normal methylation patterns of different B cell maturation states have been discovered in lymphoid malignancy and a vast majority of cancer-specific methylation events actually seem to be physiological (Kulis et al., 2012, 2015; Oakes et al., 2016). Interestingly, G-CIMP was found highly enriched among the so-called proneural subgroup of glioma (Noushmehr et al., 2010), which is characterized by a gene expression signature reminiscent of neural progenitors (Bertrand et al., 2002; Verhaak et al., 2010). Thus, mIDH may likewise block differentiation in tumors of the central nervous system, arresting the cancer cells at the neural progenitor stage. Whether mIDH-associated G-CIMP indeed reflects methylation patterns corresponding to immature neural cells remains to be studied.

Physiological methylation patterns associated with a certain degree of cellular differentiation might also explain some of the observed cancer-associated hypermethylation phenotypes or CIMP that cannot be related to *IDH* or other mutations currently (Kelly et al., 2017; Mack et al., 2014; Toyota et al., 1999; Weisenberger et al., 2006). Likewise, overlapping methylation signatures in *mIDH* and mutant *TET2* leukemia (Figueroa et al., 2010) may be due to the differentiation block rather than similar hypermethylation targets.

Recently, diverging methylation patterns in *DNMT3A* WT and *DNMT3A* R882 AML have been reported, consistent with the data presented here. Hypermethylation was suggested to be a result of cancer progression, that is accompanied by increased DNMT3A activity, rather than a driver of aberrant gene silencing (Spencer et al., 2017). This supports the notion of methylomes reflecting the cellular history proposed here and challenges the concept of cancer-specific hypermethylation driving malignant transformation. Together with the fact that DNMT3A has been implicated in hematopoietic differentiation and silencing of HSC-specific stem cell genes (Challen et al., 2014, 2011), distinct differentiation states could also play a role in the observed methylation differences between mutant and WT *DNMT3A* AML.

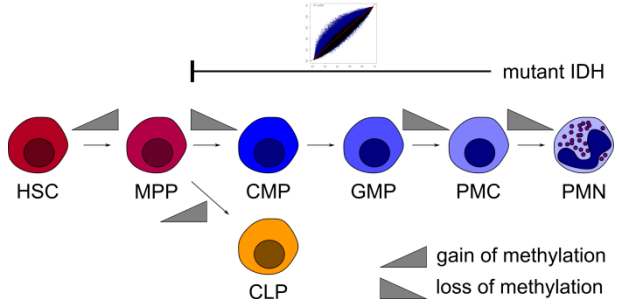

Figure 3.2: Schematic of the proposed action of mIDH in the myeloid lineage.
Whenever *IDH* mutations occur, they arrest cells in their current differentiation state and "lock in" cell-type specific methylation patterns. Since these events often take place early in the differentiation hierarchy, methylomes of myeloid progenitors are associated with *mIDH*.

The genetic mutation model for the evolution of variable cancer phenotypes suggests that the mutational events taking place in any cell with proliferative potential govern its specific phenotype (Visvader, 2011). The hypermethylation phenotype observed in *IDH* mutated cancer could thus be considered a manifestation of this theory. However, due to the enhanced understanding of the distinct epigenomes of specialized cells in healthy tissues (Kundaje et al., 2015; Ziller et al., 2013), it becomes clearer now that many of the "cancer-specific" epigenetic changes may instead reflect the epigenetic make-up of normal cell types. The cell-of-origin model proposes that phenotypically diverging tumor subtypes emerge depending on which cell type in a differentiation hierarchy of a tissue is afflicted by an oncogenic transforming event (Visvader, 2011). The findings presented here rather favor the second model, in which the mutation in *IDH* is one of the early oncogenic events arresting cells in an undifferentiated state and this is reflected in the associated methylation patterns. Remaining outliers in principal component and clustering analysis could thus reflect patients, in which transformation and *IDH* mutation happened later in the differentiation hierarchy. Thus, this study underlines the need to assess homogeneous, comparable cancer cell

populations or ideally cancers with identical cell of origin, to unequivocally detect mutation-associated epigenetic changes.

Although using 450K array data like others before, this study expands on previous analyses. Firstly, comparison between AML patients with *mIDH* and *IDH* WT status allowed the extraction of *mIDH*-dependent instead of general cancer-associated methylation changes. Secondly, unbiased analysis of *mIDH* carrying AML patients and HL-60 cells revealed distinct methylome changes, challenging the concordance of *mIDH*-associated methylation patterns between cancer patients and cell culture systems. Further, in an attempt to dissect mIDH- and TET-dependent methylation changes in AML, methylation profiles of patients with mutant *TET2* and *mIDH* were compared and DNA methylation canyons as specific target features of TET activity were examined in patients with *mIDH*. This uncovered far fewer hypermethylation events in *TET2* mutant compared to *mIDH* harboring patients and poor hypermethylation of canyons in *IDH* mutant samples. Together with the absence of methylation changes upon D-2-HG treatment in HL-60 cells, these observations argue against D-2-HG-mediated TET inhibition as the sole cause for the mIDH-associated hypermethylation. To the best of my knowledge this study is the first to assess DNA methylation changes in 2-HG-treated human cells in order to investigate effects directly mediated by the metabolite. Finally, in contrast to previous reports, the various blood differentiation states potentially present in whole bone marrow aspirates of AML patients were taken into account here. By comparing certain purified hematopoietic cell types to *mIDH* carrying patient samples, a major contribution of methylation patterns characteristic of undifferentiated myeloid progenitors to the *mIDH*-associated hypermethylation was identified.

To finally understand mIDH-associated pathogenesis, it will be important to dissect its neomorphic and catalytic independent effects. It is currently elusive, how mIDH might exactly block differentiation, but one paper suggested that histone hypermethylation due to KDM4C inhibition may be responsible (Lu et al., 2012). In order to precisely determine TET- and mIDH-dependent methylation changes in the hematopoietic lineages a conditional *TET* triple KO mouse model could be compared to the established *mIDH* knock-in model. Furthermore, cancer-specific and differentiation-associated DNA methylation changes resulting from *IDH* mutation may be inferred by in-depth comparison with appropriate reference methylomes.

3.3 Summary and Conclusions

In this thesis a previously unknown, important epigenetic regulatory role of TET enzymes at DNA methylation canyons in differentiated murine cells was identified. Integrity of canyons was found to be essential for *in vitro* adipogenic differentiation and proper gene expression. It hence provides *in vitro* confirmation of the developmental requirement for TET enzymes. Moreover, it underscores the necessity of continuous TET activity beyond preimplantation

development. Future studies will be needed to understand whether and how TET enzymes are specifically recruited to their target features. This should also clarify whether continuous demethylation acts as a default process to antagonize stochastic DNA methylation in differentiated cells or whether TET enzymes act in a more directed way upon certain cues.

Additionally, this study identified the differentiation state of myeloid leukemic cells as a major determinant of *mIDH*-associated hypermethylation in human AML. Conversely, alterations in DNA methylation upon abrogation of TET activity and expression of presumably TET-inhibiting *mIDH* were found to be poorly concordant. It thus challenges the claim of mIDH-mediated inhibition of TET-dependent DNA demethylation and provides an alternative explanation for the associated hypermethylation. More extensive studies will be required to understand how mIDH may disturb differentiation. Nevertheless, these findings reinforce the need to gain a comprehensive understanding of the oncogenic actions of mIDH, which will be vital to fully exploit mIDH as a therapeutic target.

4 Materials and Methods

4.1 Materials

Table 1: Chemicals, reagents and enzymes

Name	Company
(D)-2-hydroxyglutarate dehydrogenase (HGDH)	kind gift from S. Pusch, see Balss et al., 2012
3-Isobutyl-1-methylxanthine	Sigma-Aldrich
Acetic acid (glacial), 100%	Merck Millipore
Agar	Fluka
Agarose	Roth
Ammonium acetate	Fluka
Ammonium sulfate	Roth
Ampicillin sodium salt	Sigma
Antarctic phosphatase + buffer	New England BioLabs
Boric acid	Sigma-Aldrich
Bradford solution (Protein Assay)	Bio-Rad
Chloroform	VWR
Complete protease inhibitor cocktail tablets	Roche
Deoxynucleotide mix (100 mM dNTPs)	Agilent
Dexamethasone	Sigma-Aldrich
Diaphorase	MP Biomedicals
Diethyl pyrocarbonate (DEPC)	Roth
Disodium hydrogen phosphate dihydrate (Na_2HPO_4 x $2H_2O$)	neoLab
Dithiothreitol (DTT)	Gerbu
DNA ladder, O'RangeRuler 100bp+500bp	Thermo Scientific
DNA ladder HyperLadder™ 1kb	Bioline
DNA ladder HyperLadder™ 50bp	Bioline
DNA loading dye	Thermo Scientific
DNAse-free, RNAse-free Water	Gibco
D-α-hydroxyglutaric acid disodium salt	Sigma-Aldrich
ECL substrate	PerkinElmer
Ethanol	Sigma-Aldrich
Ethidium bromide	Roth
Ethylenediaminetetraacetate disodium salt, dihydrate (EDTA)	Gerbu
Fast digest restriction enzyme buffer	Thermo Scientific
Formaldehyde (37%)	J.T.Baker
Glycerol	Sigma-Aldrich
Glyco Blue	Ambion
HEPES	Gerbu
Hydrochloric acid (37% HCl)	Sigma-Aldrich
Indomethacin	Sigma-Aldrich
Insulin human recombinant zinc solution	Invitrogen
Isopropanol	Riedel-de Haën
Kapa HiFi HotStart Uracil+ ReadyMix (2x)	Kapa Biosystems
L-α-hydroxyglutaric acid disodium salt	Sigma-Aldrich
Methanol	Sigma-Aldrich

Methylene blue	New England BioLabs
NAD^+	AppliChem
New England BioLabs restriction enzyme buffers	New England BioLabs
NP40/ Igepal CA-630	Sigma-Aldrich
Oil Red O	Sigma-Aldrich
PBS tablets	Merck Millipore
PBS, RNAse/DNAse free , 1x/10x	Gibco
PBS-Tween tablets	Merck Millipore
Peptone from casein, tryptic digest	Fluka
PhiX Control v3	Illumina
Phusion HF buffer (5x)	New England BioLabs
Phusion High-Fidelity DNA Polymerase	New England BioLabs
Potassium chloride (KCl)	Roth
Potassium dihydrogen phosphate (KH_2PO_4)	Roth
Powdered milk	Gerbu
Proteinase K	Ambion
Resazurin	AppliChem
Restriction enzymes: ClaI	New England BioLabs
Restriction enzymes: EcoRI	New England BioLabs
Restriction enzymes: MscI	New England BioLabs
Restriction enzymes: NcoI	Thermo Scientific
Restriction enzymes: NotI	New England BioLabs
RNAse (DNAse free)	Roche
RNAse A	Sigma-Aldrich
RNAse H	New England BioLabs
Roti-Phenol	Roth
Roti-Phenol/chloroform/isoamylalcohol	Roth
S.O.C. medium	Invitrogen
Sodium acetate (NaOAc)	Roth
Sodium chloride (NaCl)	Sigma
Sodium dodecyl sulfate (SDS)	Roth
Sodium hydroxide (NaOH)	Riedel-de Haën
T4 ligase + buffer	New England BioLabs
Taq polymerase FireTaq Blue + buffer	Steinbrenner
Trisodium citrate dihydrate	Sigma-Aldrich
Triton X-100	Gerbu
Trizma® base	Sigma-Aldrich
TRIzol	Ambion
Trypan blue dye	Bio-Rad
Tween 20	Sigma-Aldrich
Yeast extract	Gerbu/ Sigma-Aldrich
β-Mercaptoethanol	Sigma-Aldrich

Table 2: Consumable material

Name	**Company**
384-well plates	Steinbrenner
5 ml polystyrene round-bottom tubes filter top	BD Falcon
96-well plates black	Brandtech
Amicon filter tubes (Amicon Ultra-4, PLQK Ultracel-PL Membrane, 50 kDa)	Millipore

Cell culture dishes (10 cm)	Greiner
Cell culture flasks (T25, T75, T150)	Greiner
Cell culture plates (12-, 24-well)	Corning
Cell culture plates (6-, 48-, 96-well)	Greiner
Cell strainer (40 µM)	BD Falcon
Centrifugal tubes (15, 50 ml)	Greiner
Counting slides	Bio-Rad
Cryotubes Nalgene (2 ml)	Thermo Scientific
Filter tips (10, 20, 200, 1000 µl)	Steinbrenner
Filters for virus purification (Minisart® NML Syringe Filter 16555, GUK, 0.45 µm Surfactant-free Cellulose Acetate	Sartorius
Filters for virus purification (Millex-HV Filter, 0,45 µm, PVDF)	Millipore
Filtropur S 0.2 sterile filters (0.2 µm)	Sarstedt
Nylon transfer membrane	GE Healthcare
Pasteur pipettes	WU Mainz
PCR tube strips (0.2 ml)	Steinbrenner
PCR tubes (0.2 ml)	Thermo Scientific
Pipette tips (10, 20, 200, 1000 µl)	Sarstedt
Reaction safe-lock tubes (1.5 ml)	Sarstedt
Reaction safe-lock tubes (2 ml)	Eppendorf
Reaction tubes (1.5, 2 ml)	Sarstedt
Serological pipettes (5, 10, 20, 50 ml)	BD Falcon/ Corning
Syringes	Braun
X-ray film	Fujifilm

Table 3: Equipment and devices

Name	**Company**
454 Genome Sequencer FLX Titanium	Roche
Agarose gel electrophoresis chamber	Bio-Rad
Balance EK-200i	A&D
Bioanalyzer Instrument 2100	Agilent Technologies
Bio-Dot® microfiltration apparatus, 96-well	Bio-Rad
Bioruptor sonication system	Diagenode
CCD camera AxioCam ERc 5s	Zeiss
Cell culture hood HERA safe/ HERA safe KS	Thermo Scientific
Cell/Particle counter	Beckman Coulter
Cell Counter TC10	Bio-Rad
Centrifuge 5804	Eppendorf
Centrifuge 5804 R	Eppendorf
Centrifuge benchtop 5415 D	Eppendorf
CO_2 incubator HERA cell 150	Thermo Scientific
CO_2 incubator Sanyo	Sanyo
Covaris S2 Ultrasonicator	Covaris
Cryo 1°C freezing container	Nalgene/Nunc
Developing cassette	Dr. Goos-Suprema
Developing machine COMPACT™ 2	Protec
Fine balance CP64	Sartorius
FLUOstar OPTIMA plate reader	BMG Labtech
Hybridization oven	PeqLab
Illumina HiSeq 2000 system	Illumina

Incubator & Shaker INNOVA 4200	New Brunswick Scientific
Infinium HumanMethylation450 BeadChip	Illumina
Infinium MethylationEPIC BeadChip	Illumina
iScan array scanner	Illumina
LightCycler 480 instrument	Roche
Micropipettes	Gilson
Microscope Axiovert 25	Zeiss
Microscope Axiovert 40 CFL	Zeiss
Microvolume spectrometer NanoDrop 2000	PeqLab
Orbital shaker Unimax 1010	Heidolph
pH meter	Hanna instruments
Power supply	Consort
Real Time PCR System, LightCycler 480	Roche
Reflected-light microscope SZX10	Olympus
Rocker Unimax1010	Heidolph
Thermocycler DNA Engine	Bio-Rad
Thermomixer compact	Eppendorf
UV gel documentation, E.A.S.Y® Doc plus	Herolab
Vortex REAX top	Heidolph
Water Bath TW12/ Julabo 12	Julabo
Wheel REAX 2	Heidolph

Table 4: Kits

Name	**Company**
ABsolute qPCR SYBR green mix	Thermo Scientific
Agilent DNA 1000 Kit	Agilent Technologies
BCA Kit	Pierce
Cell Titer Glo®	Promega
Deproteinization Kit	BioVision
EpiTect Bisulfite Kit	Qiagen
EZ-96 DNA Methylation Kit	Zymo Research
GS FLX Titanium PicoTiterPlate Kit 70 x 75	Roche
GS FLX Titanium Sequencing Kit XLR70	Roche
GS Junior Maintenance Wash Kit V2	Roche
GS Junior Titanium emPCR Kit (Lib-A)	Roche
HiSpeed Plasmid Midi Kit	Qiagen
Mesa green qPCR mastermix PLUS for SYBR assay (no ROX)	Eurogentec
PeqGold Extraction Kit	PeqLab
PyroMark PCR Kit	Qiagen
QIAprep Spin Miniprep Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR purification Kit	Qiagen
QuantiTect cDNA synthesis Kit	Qiagen
Quant-iT^{TM}PicoGreen ®dsDNA Kit	Invitrogen
Qubit Assay Kit Broad Range (BR)	Invitrogen
Qubit Assay Kit High Sensitivity (HS)	Invitrogen
SuperScript® III First-Strand Synthesis System for RT-PCR	Invitrogen
TruSeq DNA Sample Prep v2 Kit	Illumina
TruSeq RNA Library Preparation Kit v2	Illumina
Zymo RNA Clean & Concentrator Kit	Zymo Research

Table 5: Buffers and solutions

Name	Composition
Dotblot blocking solution	5% w/v milk in 1x PBS
Dotblot denaturation solution	2 mM NaOH 50 mM EDTA
Dotblot neutralization solution (pH 7)	2 M ammonium sulfate
LB agar	5 g yeast extract 10 g NaCl 10 g peptone 15 g agar Ad 1 l with ddH_2O
LB medium	5 g yeast extract 10 g NaCl 10 g peptone Ad 1 l with ddH_2O
Methylene blue staining solution (pH 5.2)	Methylene blue 0.04% w/v 500 mM sodium acetate
NP-40 lysis buffer	150 mM NaCl 0.1% v/v NP40 50 mM Tris-HCl (pH 8) + freshly added 1x complete protease inhibitor
Oil Red O stock solution	0.35% w/v Oil Red O in isopropanol
Oil Red O working solution	60% v/v Oil Red O stock solution in ddH_2O
PBS-Tween 20 (PBST)	1x PBS 0,1% v/v Tween20
Phosphate buffered saline (PBS), 10x, (pH 7.5)	1,37 M NaCl 27 mM KCl 100 mM Na2HPO4 18 mM KH_2PO4
Pre-lysis buffer for gDNA isolation (pH 8)	10 mM Tris-HCl pH 8 5 mM EDTA 100 mM NaCl Freshly added: 1.1% v/v SDS 0.1 mg/ml Proteinase K 0.04 mg/ml RNAse A
SSC, 20x, (pH 7)	3 M NaCl 300 mM trisodium citrate
Tris-acetate-EDTA buffer (TAE), 50x, (~pH 8)	2 M Tris 50 mM EDTA 0.95 M acetic acid (to adjust pH)
Tris-borat-EDTA buffer (TBE), 10x	890 mM Trizma® base 890 mM Boric acid 20 mM EDTA
Tris-EDTA buffer (TE)	10 mM Tris-HCl (pH 8) 1 mM EDTA

Table 6: Antibodies

Name	Species	Dilution	Company
5-hydroxymethylcytosine, #39791	rabbit	1:1000	Active Motif

5-methylcytosine, #39649, clone 33D3	mouse	1:1000	Active Motif
HRP-coupled anti mouse IgG (115-035-003)	goat	1:10000	Jackson ImmunoResearch
HRP-coupled anti rabbit IgG (111-035-003)	goat	1:10000	Jackson ImmunoResearch

Table 7: Biological material

Name	**Company/Source**
HEK293T cell line	Hofmann laboratory cell collection
HL-60 cell line	Lyko laboratory cell collection
NOMO-1 cell line	DSMZ (#ACC 542)
Primary MEFs Tet1+/+; Tet1-/-; Tet1/2-/- (DKO)	Rudolf Jaenisch, Whitehead Institute for Biomedical research, Cambridge, US
Primary MEFs Tet2+/+; Tet2-/-	Mingjiang Xu, Herman B Wells Center for Pediatric Research, Indiana University Melvin and Bren Simon Cancer Center, Indiana University School of Medicine, Indianapolis, US
Stbl3™ One Shot™ Chemically Competent *E. coli*	Invitrogen

Table 8: Cell culture reagents

Name	**Composition/ Company**
ADM	1 µM Dexamethasone 1 µg/ml Insulin 500 µM 3-Isobutyl-1-methylxanthine 50 µM Indomethacin in DMEM supplemented with 10% FCS and Pen/Strep
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
DMEM (1x), 4.5 g/l D-Glucose, (+) L-Glutamine, (-) Pyruvate	Gibco
FBS	Gibco
Ficoll-Paque PLUS	GE Healthcare
HEPES buffer solution 1 M	PAN Biotech
L-Glutamine 200 mM (100x)	Gibco
Lipofectamine® 2000	Invitrogen
OptiMEM (1x) + GlutaMAX	Gibco
Penicillin Streptomycin mix, 100x	Gibco
RPMI Medium 1640 (1x) + GlutaMAX	Gibco
Sodium pyruvate 100 mM (100x)	Gibco
Trypsin- EDTA 0.25% (1x)	Gibco

Table 9: Vectors

Name	Company/Source
pHAGE PGK-GFP-IRES-LUC-W	Addgene plasmid # 46793 (gift from Darrell Kotton, see Wilson et al., 2010)
pLVX-IRES-ZsGreen1	kind gift from M. Rodríguez (Division of Epigenetics, Prof. F. Lyko, DKFZ; Clontech plasmid #632187)
pMD2.G	kind gift from M. Rodríguez (Division of Epigenetics, Prof. F. Lyko, DKFZ; Addgene plasmid #12259; gift from Didier Trono)
pMXs-IDH1R132C	kind gift from S. Pusch (Clinical Cooperation Unit Neuropathology, Prof. A. von Deimling, DKFZ)
pMXs-IDH1WT	kind gift from S. Pusch (Clinical Cooperation Unit Neuropathology, Prof. A. von Deimling, DKFZ)
pMXs-IDH2R140Q	kind gift from S. Pusch (Clinical Cooperation Unit Neuropathology, Prof. A. von Deimling, DKFZ)
pMXs-IDH2WT	kind gift from S. Pusch (Clinical Cooperation Unit Neuropathology, Prof. A. von Deimling, DKFZ)
psPAX2	kind gift from M. Rodríguez (Division of Epigenetics, Prof. F. Lyko, DKFZ; Addgene plasmid #12260; gift from Didier Trono)

Table 10: Primers

Application	Name	Sequence
454	m_Hoxa1_up	TGGTTAGGGTTAGAGTTTGTTGTTT
	m_Hoxa1_lo	AAAACCTTAAAATTTCTTATTCCCACT
	m_Hoxa3_up	AATTTTTTGAATAATTAGGGTATAGAATTT
	m_Hoxa3_lo	AAATTTTAATTTTTAAAATTACACTTCTTA
	m_Foxc1_up	TTTGGGGAATTGTAATTTTTTTAG
	m_Foxc1_lo	TAAATACCCTCTTTCTATTATCCCC
	m_Foxc1-prom_up	TTTTTAGTTTTTTTAAGTTTTGGAA
	m_Foxc1-prom_lo	ATTTACTCTCCACAAACAAATTCTC
	m_Foxc2_up	TTTGTTTGATAATTAGGATTTAAGAAGT
	m_Foxc2_lo	CCTAAAATAACACAAAAAACTCTCC
	m_Foxc2-prom_up	GGGTTGGTTGTTTTGTTTTAATTT
	m_Foxc2-prom_lo	ACACCTAAAAAACCATTAATCCTTC
	m_Igf1_up	TTAGAATTGGGGGTTATTTATAAATTGT
	m_Igf1_lo	CATAAACACAAACACAAAAACATACTCTA
	m_Pparγ_up	ATTTAGGGATAGAGTGAGGGGTTT
	m_Pparγ_lo	TCTAAAACAAAAACAACCCTATTCC
qRT-PCR	m_β-actin_up	TGAACCCTAAGGCCAACCGTGAAA
	m_β-actin_lo	CAGGATGGCGTGAGGGAGAGCATAG
	m_C/ebpα_up	TGGACAAGAACAGCAACGAG
	m_C/ebpα_lo	TCACTGGTCAACTCCAGCAC
	m_Col2a1_up	AACTGGCAAACAAGGAGACAGAG
	m_Col2a1_lo	TGCCAGAGGGACCAACAGG

	m_Ebf3_up	CTCCTTCCTGCACCAGCG
	m_Ebf3_lo	GTCTTTTGTAGCATTCAACTGTGG
	m_Foxc1_up	CGAGTGCTTCGTCAAGGTGC
	m_Foxc1_lo	GGTTCTTGGAGGTGCAGCC
	m_Foxc2_up	GCCCAAGGACCTGGTGAAG
	m_Foxc2_lo	GTCGAGCGTCCAGTAGCTGC
	m_Foxp1_up	AACGAGAGTGACAGCAGTCCAG
	m_Foxp1_lo	GTTCGTCTTCGTAATCTCTGTCATG
	m_Gapdh_up	CATGGCCTTCCGTGTTCCTA
	m_Gapdh_lo	TGCTTCACCACCTTCTTGATGT
	m_Hoxa3_up	CAATGGGTTCGCTTACAATGC
	m_Hoxa3_lo	AGGCAGGTCGATGGTACTCAAC
	m_Hoxb6_up	TTCCTATTTCGTGAACTCCACCTT
	m_Hoxb6_lo	CCGCATAGCCAGACGAGTAGA
	m_Hoxb9_up	TGTCCATTTCTGGGACGCTTA
	m_Hoxb9_lo	GAACACCGGCGCTTTGG
	m_Hoxc5_up	ACCCGTGGATGACCAAACTG
	m_Hoxc5_lo	AGGGTCTGGTAGCGCGTGTA
	m_Hoxc8_up	TCTCCCAGCCTCATGTTTCC
	m_Hoxc8_lo	GTCTGATACCGGCTGTAAGTTTGTC
	m_Jag1_up	AGAAGTCAGAGTTCAGAGGCGTC
	m_Jag1_lo	GGAGTGAGTGTGGCTGCTGG
	m_Loxl2_up	CGCATCTGGATGTACAACTGTCA
	m_Loxl2_lo	TGTCCAGAGTGGAAATCTTGTGG
	m_Pparγ_up	CATAAAGTCCTTCCCGCTGA
	m_Pparγ_lo	GAAACTGGCACCCTTGAAAA
	m_Tet1_3-4_up	GCTGGATTGAAGGAACAGGA
	m_Tet1_3-4_lo	GTCTCCATGAGCTCCCTGAC
	m_Tet1_4-5_up	GTCAGGGAGCTCATGGAGAC
	m_Tet1_4-5_lo	CCTGAGAGCTCTTCCCTTCC
	RT-Tet2-m-1_F	GTCAACAGGACATGATCCAGGAG
	RT-Tet2-m-1_R	CCTGTTCCATCAGGCTTGCT
	RT-Tet3-m_F	GGAGTTGGCTGGAGTCACCAC
	RT-Tet3-m_R	CCACCGCATTGCCACTGTAC
	m_Zeb2_up	TACCCAACGGGAGCAGCTAC
	m_Zeb2_lo	CGCAGAAGGGAACTGCTTTTC
	human_IDH1_1_for	TTGTCCAGATGGCAAGACAG
	human_IDH1_1_rev	GCTTTGCTCTGTGGGCTAAC
	human_IDH1_2_for	AGTCTGCAAGACTGGGAGGA
	human_IDH1_2_rev	CCAATTCCACGTAGGGAAAA
	hIDH1_exo_down_for	GAGACAATTGAGGCTGGCTTC
	hIDH1_exo_down_rev	CCAGAGGTTGATTAGGATCTATCG
	human_IDH2_1_for	TGGCTCAGGTCCTCAAGTCT
	human_IDH2_1_rev	CTCAGCCTCAATCGTCTTCC
	hACTB_up	AGCACAGAGCCTCGCCTTT
	hACTB_lo	GAATCCTTCTGACCCATGC
	hGAPDH_up	CGACCACTTTGTCAAGCTCA
	hGAPDH_lo	GGTGGTCCAGGGGTCTTACT
cloning	IDH1_cloning_3_for (to introduce MscI site)	AATCTGTGGCCAACCATGTCCAAAAAAATC AGT
	IDH1_cloning_3_rev (to introduce ClaI site)	AATCTATCGATACCCTAAAGTTTGGCCTGA GC
	IDH2_cloning_3_for	AATCTGTGGCCAACCATGGCCGGCTACCT

	(to introduce MscI site)	GCGG
	IDH2_cloning_3_rev (to introduce ClaI site)	AATCTATCGATACCCTACTGCCTGCCCAGGGC
other	hIDH1_mut_for	CGGTCTTCAGAGAAGCCATT
	hIDH1_mut_rev	TTGGAAGGAACTGTGTGCAA

Table 11: Software

Name	Purpose
BiSQuID	454 data analysis
GIMP	picture manipulation
Image J	particle/pixel quantifications
Ingenuity® Pathway Analysis	calculations of biological pathway enrichments
Inkscape	figure assembly, schemes
Microsoft Office Suite	data analysis, document generation
OPTIMA analysis software	analysis of absorbance or fluorescence measurements (from OPTIMA plate reader)
pDRAW32	vector visualization
R	Infinium data analysis, graphs
SigmaPlot	graphs, statistical testing

4.2 Eukaryotic cells and cell culture

All cells were grown under sterile conditions in a humidified atmosphere with 5% CO_2 and at 37°C. Cells were handled under sterile culture hoods for splitting and manipulation. Standard culture media were supplemented with 10% (v/v) fetal bovine serum (FBS), 100 U/ml penicillin and 100 µg/ml streptomycin. Complement contained in the FBS was heat-inactivated by incubation at 56°C for 30 min and FBS was sterile-filtrated prior to use.

4.2.1 Cell lines and primary cells

4.2.1.1 Culture of adherent cells

Primary MEFs were isolated as described before (Wiehle et al., 2016) and maintained in standard T25 or T75 cell culture flasks in DMEM standard culture medium. Cells were passaged every three to four days at a ratio of 1:3. For subculture cells were washed twice in PBS, dissociated using trypsin, pelleted (130 x g, 5 min) and resuspended in fresh medium. For seeding defined cell numbers MEFs were counted using a particle counter from Beckman Coulter.

4.2.1.2 Culture of suspension cells

HL-60 and NOMO-1 cells were cultured in RPMI standard culture medium in cell culture plates or T25/T75 suspension culture flasks. Cells were maintained at a density of 1 to

10 x 10^5 viable cells/ml and split twice per week by addition or replacement with fresh medium. To seed a defined viable number of cells they were mixed with Trypan blue in a 1:1 ratio to label dead cells and live cells were counted using the TC10 automated cell counter.

4.2.1.3 Thawing and freezing of cells

For long-term storage 1 to 2 x 10^6 cells were resuspended in 1 ml of standard culture medium supplemented with 10 % (v/v) DMSO. The cell suspension was stored in cryotubes and cooled down at a rate of -1°C per minute in freezing containers that were placed at -80°C. Few days later the tubes were transferred to liquid nitrogen (-196°C) for long-term preservation. For thawing cells were placed in a water bath of 37°C and resuspended in 5 ml of medium once the suspension was liquid. After centrifugation at maximally 300 x g for 5 min the freezing medium was quickly removed to prevent cell damage due to toxic DMSO, cells were taken up in pre-warmed standard culture medium and plated.

4.2.2 Generation of stable cell lines

Stable HL-60 and NOMO-1 cell lines expressing either wildtype or mutant IDH enzymes or none of them (empty vector control) were generated by lentiviral transduction. For lentivirus production HEK293T cells grown in RPMI standard culture medium were seeded into 6-well plates. When cells reached 70 – 90% confluency, medium was exchanged with 3 ml of fresh RPMI and cells were transfected with target vector, lentiviral packaging vector (psPAX2) and envelope plasmid (pMD2.G) in a ratio of 5:4:1. Total transfected DNA per well was 2.5 µg.
For each well to be transfected, 140 µl of OptiMEM mixed with 7.5 µl Lipofectamine® 2000, and 140 µl of OptiMEM containing the 2.5 µg of vector DNA, were prepared. The two solutions were mixed, vortexed for 2 seconds, incubated for 5 min at room temperature and added to the HEK293T cells in small droplets. Cells were transferred to the S2 laboratory and 48 h later viral particles were harvested. To this end cell culture supernatants were collected, centrifuged for 5 min at 130 x g and filtered using 0.45 µm filters (Sartorius) to remove remaining HEK293T cells. For infections 200,000 HL-60/NOMO-1 cells were pelleted and resuspended in the freshly harvested and filtered virus-containing cell culture supernatant. Two days later medium was exchanged with standard RPMI and after another three days cells were removed from the S2 laboratory and examined for green fluorescence under the microscope.
In an alternative protocol for infections HEK293T cells were cultured and transfected in DMEM supplemented with 10% (v/v) fetal bovine serum (FBS), 100 U/ml penicillin and 100 µg/ml streptomycin, 2 mM L-Glutamine, 1 mM sodium pyruvate and 10 mM commercial HEPES solution. Virus particles were 0.45 µm filtered (Millipore) and concentrated in Amicon

filter tubes by several 10 minute spins at 1000 x g. After volume reduction to roughly 250 µl the concentrated virus was directly added to the cells kept in 2.5 ml of their appropriate standard medium. Remaining procedures were as described above. Lentiviral transductions were performed in collaboration with Florian Köhler and Bojana Kriznik.

Cells were expanded for a few days and FACS sorted by GFP/ZsGreen1 expression to generate successfully transduced cell pools co-expressing the transgene. The 10% and 30% with the strongest fluorescence signal of the positive population were sorted to a pool of at least 1000 cells. Cell sorting was conducted by Dr. Langlotz at the Flow Cytometry & FACS Core Facility of the ZMBH in Heidelberg.

Sorted cells were pelleted, plated in standard RPMI and expanded for a few days. Ficoll purification was performed after reasonable expansion to remove dead cells. To this end cells were washed once in PBS, resuspended in 700 µl of standard culture medium and 700 µl of Ficoll were overlaid with the cell suspension in Eppendorf tubes. Tubes were subjected to a 15 seconds spin at 16,100 x g, rotated by 180 degree and spun again for 1 minute at 12,000 x g. The viable cells located at the interface between Ficoll and medium were recovered, resuspended in medium, pelleted and plated in fresh standard culture medium.

4.3 Molecular cloning

Two different backbones were used in this study to accommodate transgenes and eventually generate stable cell lines expressing wildtype or mutant *IDH*: pLVX-IRES-ZsGreen1 and pHAGE PGK-GFP-IRES-LUC-W. Empty vectors were transduced as controls. The following scheme depicts the cloning strategy.

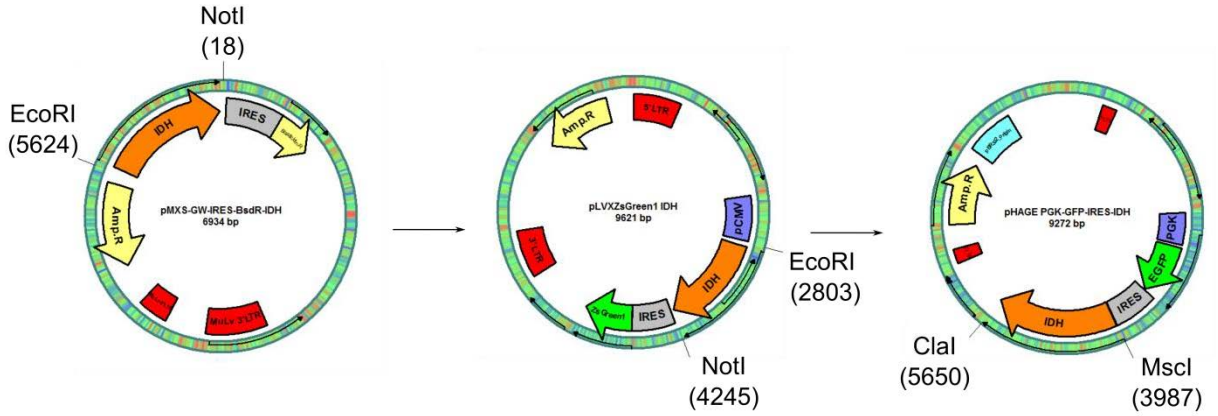

Figure 4.1: Cloning strategy to introduce *IDH* genes into different lentiviral backbones.

PMXs vectors harboring the different wildtype and mutant IDH genes were a kind gift of S. Pusch. Using EcoRI and NotI transgenes were excised and integrated into the pLVX backbone. For cloning them into the pHAGE backbone transgenes were PCR amplified from the pLVX vector using primers

introducing MscI or ClaI restriction sites at the ends of the transgenes that were compatible with the pHAGE backbone.

For all enzymatic digestions in this study the units of restriction enzyme were adapted according to the definition that one unit digests one µg of template in one hour under optimal reaction conditions. To ensure efficient restriction extended incubation times and a volume of at least 1 µl (usually comprising more than the required units) were used. Furthermore volumes of glycerol were kept below 10% in all enzymatic reactions.

4.3.1 pLVX-IRES-ZsGreen1 constructs

4.3.1.1 Preparation of fragments

PMXs vectors harboring wildtype or mutant *IDH1* (R132C) or *IDH2* (R140Q) coding sequences were restriction digested in triplicates to excise the gene of interest as follows:

reagent	**amount**
vector	1 µg
NEB buffer 3.1 (10x)	3 µl
EcoRI	1 µl
NotI	1 µl
ddH_2O	To 30 µl

The pLVX-IRES-ZsGreen1 target vector was treated in the same way to linearize it and generate corresponding sticky ends ensuring correct orientation of the transgene. Restriction was performed overnight at 37°C. To prevent religation of linearized target vector without incorporation of insert its ends were dephosphorylated for 1 hour at 37°C without shaking:

reagent	**amount**
Antarctic phosphatase buffer (10x)	4 µ
Antarctic phosphatase	1 µl
ddH_2O	5 µl
µl target vector reaction (see above)	30 µl

The resulting vector fragments were subjected to gel electrophoresis in a preparative 0.8% TAE agarose gel containing 0.25 µg/ml ethidium bromide (Figure 4.2A, B). Fragments and linearized target vector of the expected molecular size were excised and gel-extracted using a commercial gel extraction kit. To control for the presence of a single pure DNA fragment a 0.8% agarose control gel was run with 75 – 200 ng of DNA loaded per pocket (Figure 4.2C). In this way the measured concentration of the purified DNA was also confirmed or re-estimated if it strongly deviated from Nanodrop quantifications.

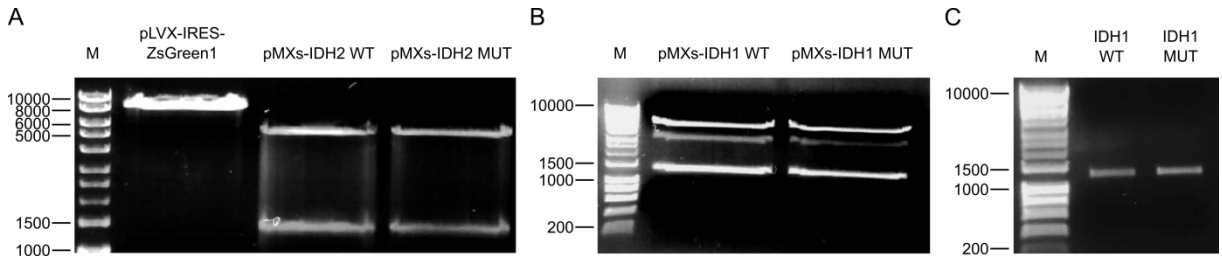

Figure 4.2: Restriction digestion of pLVX-IRES-ZsGreen1.

Target vector pLVX-IRES-ZsGreen1 and source vectors pMXs-*IDH2* WT, pMXs-*IDH2* MUT **(A)**, pMXs-*IDH1* WT and pMXs-*IDH1* MUT **(B)** were digested using EcoRI and NotI. The weak middle band in B either represents supercoiled uncut vector or a contamination of the original plasmid stock with religated vector without insert. **(C)** Examples of expected fragments after gel extraction.

4.3.1.2 Ligation

In order to ligate the pLVX backbone with the *IDH* inserts a ligation reaction with thrice the molarity of fragment DNA compared to target vector DNA was set up. For example, in the case of *IDH2* the insert size was 1442 bp and the linearized target vector was 8179 bp, resulting in the amount X of insert to be used:

$$X\,[ng] = \frac{3 * insert\,[bp]}{1 * target\ vector\,[bp]} * m(target\ vector) = \frac{3 * 1442}{8179} * 25 = 0.53 * 25 = 13.3$$

Routinely, 25 ng of target vector were ligated with the appropriate amount of insert, in this case 13.3 ng. Ligations were set up as follows and incubated for 3 hours to overnight at 4°C.

reagent	**amount**
target vector	25 ng
insert	X ng
ligation buffer (10x)	1 µl
T4 ligase	1 µl
ddH_2O	To 10 µl

A control ligation using only linearized vector was set up in addition to monitor religation frequency and background colony growth on agar plates.

4.3.1.3 Transformation

The ligated plasmids were introduced into Stbl3 chemocompetent E. coli by heat-shock. Bacteria were thawed on ice and 50 – 100 µl bacterial suspension was mixed with the entire ligation reaction. The mix was incubated for 5 min on ice, 2 min at 42°C and again 5 min on ice. Next, 900 µl of S.O.C medium without antibiotics were added and the suspension was incubated for 45 min at 37°C in a thermomixer with shaking at 1000 rpm. Bacteria were

pelleted by high speed centrifugation for 1 minute, resuspended in 100 µl LB medium and seeded onto agar plates containing 100 µg/ml ampicillin which were incubated overnight at 37°C.

4.3.1.4 Plasmid isolation and screening

The next day colonies of different sizes were picked, propagated on a replica plate for potential plasmid midi preparations and expanded for plasmid mini preparations by inoculating one colony in 2 ml of LB with 100 µg/ml ampicillin and shaking them overnight at 37°C. The next day bacteria were harvested by centrifugation and plasmid DNA was prepared following the instructions of the user manual of the Qiagen Miniprep Kit. Control digestions with 1 µg of plasmid DNA using EcoRI and NotI for four hours were performed as described above. The resulting fragments were monitored on a 0.8% agarose gel.

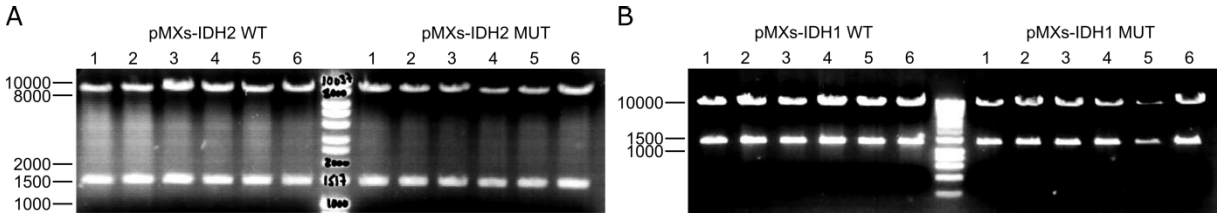

Figure 4.3: Control restriction of pMXs-*IDH* vectors.
Plasmids isolated from six different bacterial clones per construct were digested with EcoRI and NotI and fragments were visualized on a 0.8% agarose gel. **(A)** For *IDH2* constructs clone #6 (*IDH2* WT) and clone #1 (*IDH2* MUT) were sent to sequencing and found to have the correct sequence. **(B)** For *IDH1* constructs clones #1 were sent for sequencing and found to have the expected sequence.

Alternatively, bacterial colonies were picked, resuspended in 20 µl of ddH_2O and monitored for the presence of transgene by colony-PCR (see below). Only positive clones were expanded and plasmids were purified by mini preparation.
In both cases promising plasmids were sequenced to confirm the presence and integrity of the cloned transgene. Sequencing was performed by a commercial sequencing service provider using their universal primers pEGFP-FP and pIRES-RP, which amplified the transgene. Correct clones were propagated from replica plates or bacterial suspensions in 200 ml LB-Ampicillin overnight at 37°C and plasmid DNA was extracted following the instructions of the HiSpeed Plasmid Midi Kit.

4.3.2 pHAGE PGK-GFP-IRES-LUC-W constructs

4.3.2.1 Preparation of fragments

For cloning *IDH* transgenes into pHAGE PGK-GFP-IRES-LUC-W they were PCR-amplified from the previously cloned pLVX constructs with specific primers introducing restriction sites for MscI and ClaI. These enzymes had restriction sites in the pHAGE target vector surrounding the LUC gene. Therefore these nucleases could be used to linearize the target vector, excise the unwanted LUC gene and generate unambiguous DNA ends ensuring correct orientation of the inserts. The target vector was linearized in triplicates as follows:

reagent	**amount**
target vector	~ 800 ng
cutsmart buffer (10x)	3 µl
MscI	1 µl
ClaI	1 µl
ddH_2O	To 30 µl

After overnight incubation at 37°C digestions were pooled and analyzed on a 0.8% agarose gel (Figure 4.4A).

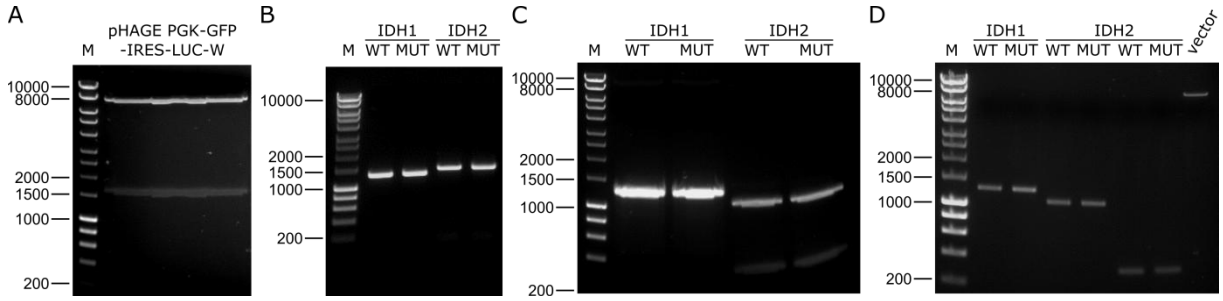

Figure 4.4: Restriction digestion of pHAGE PGK-GFP-IRES-LUC-W and PCR of *IDH* inserts.
(A) Linearized target vector fragment (7609 bp) and excised LUC fragment (1663 bp). **(B)** PCR products of IDH inserts from previously cloned pLVX constructs. **(C)** MscI and ClaI digested PCR products. The *IDH2* transgene harbors an internal MscI site, which results in fragmentation of the insert into two fragments (1059 and 313 bp). **(D)** Gel extracted and purified fragments and target vector.

The vector backbone was excised, gel-extracted into 60 µl of ddH_2O (c = 17.6 ng/µl) and dephosphorylated by addition of 3 µl Antarctic phosphatase and incubation for 1 hour at 37°C. The dephosphorylated vector was then purified by a commercial PCR product purification kit and analyzed on a 0.8% agarose gel (Figure 4.4D, right lane).
PCR amplification of *IDH* transgenes was performed with primers listed in Table 10 in duplicates as follows:

reagent	amount
Vector template	100 ng
Phusion high fidelity buffer (5x)	10 µl
Phusion high fidelity polymerase	0.5 µl
primer mix (10 µM forwards/ reverse each)	2.5 µl
10mM dNTPs	1 µl
ddH_2O	To 50 µl

PCR conditions:

temperature [°C]	time	cycles
98	30 sec	1
98	10 sec	
55	15 sec	30
72	30 sec	
72	7 min	1
4	∞	

PCR products were pooled and 10 µl were loaded to a 0.8% agarose gel (Figure 4.4B). The remaining PCR products were precipitated by addition of 1/10 (v/v) sodium acetate and 3x (v/v) ethanol (100%) and incubation at -80°C for 30 min. After a 10 min centrifugation the pellets were washed with 70% ethanol, air-dried and resuspended in 25 µl of ddH_2O. MscI, ClaI and cutsmart buffer were added as described above and PCR products were digested for 4 hours at 37°C to generate DNA ends corresponding to the target vector. The resulting fragments were separated on a preparative 0.8% agarose gel (Figure 4.4C), excised and gel extracted. Purified fragments were examined for their proper appearance by gel electrophoresis (Figure 4.4D). Cutting of the *IDH2* transgene with MscI generated two fragments because this gene possesses an internal recognition site for MscI. Therefore three fragments had to be religated later on and checked carefully to ensure correct orientation of the smallest fragment.

4.3.2.2 Ligation

Ligations were set up as described above with 25 to 75 ng of target vector and the corresponding amounts of inserts (the total amount of *IDH2* fragment was distributed among the two *IDH2* subfragments proportionally to their molecular weights). Ligations were incubated overnight in an ice-waterbath placed at room temperature resulting in an increasing temperature gradient starting from 4°C.

4.3.2.3 Transformation

Ligation reactions were transformed into Stbl3 bacteria as described before. Bacteria were seeded on LB ampicillin plates and incubated overnight at 37°C.

4.3.2.4 Plasmid isolation and screening

To check the presence of insert several bacterial clones were picked and analyzed by colony-PCR:

reagent	amount
Bacterial suspension/control vectors	1.5 µl/100 ng
FireTaq Blue reaction buffer (10x)	2 µl
FireTaq Blue polymerase	2 µl
primer mix (10 µM forwards/ reverse each)	1 µl
10 mM dNTPs	0.5 µl
ddH_2O	To 20 µl

PCR conditions:

temperature [°C]	time	cycles
95	5 min	1
95	25 sec	
60	25 sec	30
72	30 sec	
72	5 min	1
4	∞	

PCR products were analyzed on a 2.5% agarose gel (Figure 4.5A).

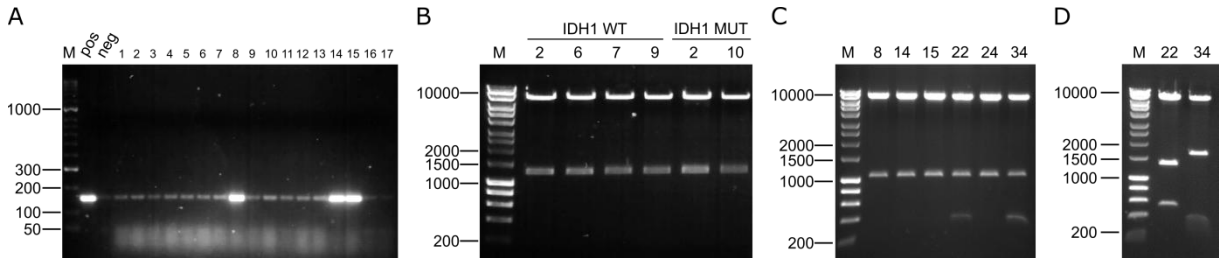

Figure 4.5: Colony-PCR and control restrictions of ligated plasmids.
(A) Example of a colony PCR of *IDH2* conducted on a selection of bacterial clones. The expected PCR product has a size of 155 bp. Weak bands in the negative control and lanes other than #8, #14 and #15 are contaminations. **(B)** Exemplary control digestions with pHAGE PGK-GFP-IRES-IDH1 WT or –IDH1 MUT and **(C)** pHAGE PGK-GFP-IRES-IDH2 WT using MscI and ClaI. Expected fragments were 7609 and 1274 bp for **(B)** and 7609, 1058 and 313 bp for **(C)**. **(D)** Exemplary control digestion of clone #22 and #34 from (C) with NcoI. Expected fragments for the correctly oriented inserts were 7154, 1302 and 523 bp and for the wrong orientation 7154, 1601 and 224 bp. The insert of clone #22 was sequenced and found to have the correct sequence.

Promising clones were amplified and their DNA was isolated as described above using mini preparations. The resulting plasmids were control digested with MscI and ClaI to assess the presence of insert (Figure 4.5B, C). In the case of *IDH2* parallel digestions of 2 x 500 ng with NcoI were performed to control for the correct orientation of small fragment (Figure 4.5D):

reagent	amount
vector	500 ng
Fast digest buffer (10x)	2 µl
NcoI, fast digest	1 µl
ddH_2O	To 20 µl

Promising plasmids were sequenced to confirm the presence and integrity of the cloned transgene with the commercial service provider and pIRES-RP or pEGFP_C2-FP and h_IDH1_1_for/rev or h_IDH2_1_for/rev primers (listed in Table 10), respectively. Bacterial clones corresponding to correct plasmids were propagated and DNA was isolated as described above.

4.4 Nucleic acid analyses

4.4.1 DNA extraction

For genomic DNA isolation freshly prepared cell pellets were washed in PBS and lysed in small falcons by addition of 4.5 ml of pre-lysis buffer. SDS, proteinase K and RNAse A were freshly added as described in

Table 5. The tubes were incubated overnight at 37°C without shaking. The next day 2.5 ml of a 5 M NaCl solution were added, the mixture was shaken vigorously and centrifuged for 15 min at maximal speed. The precipitated proteins should localize to the bottom of the tube in a white pellet-like phase. If a correct separation of phases was not achieved the procedure was repeated in 2 ml Eppendorf tubes. The clear supernatants were then pooled, transferred to a new falcon tube and mixed with 5.6 ml of isopropanol by gentle inversion of the tubes to precipitate the DNA. It was then transferred to a fresh Eppendorf tube with a pipet tip if visible or tubes were centrifuged at maximal speed for 10 min and the supernatant was discarded. The DNA pellet was washed with 70% (v/v) ethanol, collected by centrifugation and the ethanol was removed as completely as possible. Pellets were air-dried for a few minutes and resuspended in 50 to 100 µl of ddH_2O. DNA concentration was determined by absorbance measurement using the Nanodrop device.

4.4.2 Dot blot analyses

For quantification of DNA modifications by dot blot genomic DNA was sonified for 8 cycles at maximal output level with one cycle consisting of 30 sec on and 30 sec off using the Bioruptor device. The DNA concentration of the sheared DNA was measured in triplicates by Qubit broad range assay following the instructions in the user manual. Next, 4 µg of DNA were transferred to an Eppendorf tube and filled up to 80 µl with fresh TE-buffer. 20 µl of denaturation solution (

Table 5) were added and samples were incubated for 10 min at 95°C to denature the DNA. After incubation 100 µl of neutralization solution were added, samples were put on ice immediately and incubated 10 min. 2-fold serial dilutions were prepared in TE ranging from 1000 ng to 31.25 ng. A nylon membrane of 8 x 12 cm was pre-wet in 6 x SSC for 10 min. The dot blot apparatus was washed with a 0.5% SDS/ 0.5 M NaOH solution prior to use to destroy remaining nucleic acids. It then was assembled according to instructions in the user manual and the membrane was vacuum-dried and equilibrated with 100 µl of clean TE. Samples were loaded in duplicates to stain for 5mC and 5hmC in a volume of 48 µl per well. Remaining wells were filled with TE and samples were immobilized on the membrane by gentle application of vacuum. Care was taken to avoid air bubbles in the wells during loading. Afterwards wells were washed with 100 µl of 2 x SSC, the apparatus was disassembled and the membrane was rinsed in 2 x SSC. The membrane was then baked at 80°C for 2 hours. To monitor equal loading the membrane was stained for 5 min with methylene blue solution on a shaker. After documentation the membrane was destained in ddH_2O and shaken in dot blot blocking solution for 1 hour at room temperature. The membranes were incubated likewise in primary antibody solution, washed 3 x 5 min in PBS and stained with secondary antibody for one hour at room temperature. After six more 5 min washes in PBS the membrane was immersed in freshly prepared enhanced chemiluminescence solution for 1 min and the signal was detected by applying light-sensitive X-ray films in the dark. The films were developed according to instructions of the developing machine.

4.4.3 RNA extraction

For total RNA isolation cells were washed in PBS, pelleted, resuspended in 50 µl PBS and mixed with 950 µl of TRIzol. Cells were vortexed until the pellet had dissolved and left at room temperature for 5 min to lyse nuclei. After addition of 200 µl chloroform, tubes were shaken vigorously for 15 sec, incubated for further 3 min at room temperature and centrifuged at 12,000 x g and 4°C for 15 min. The aqueous phase was transferred into a fresh tube and 500 µl of isopropanol (and 1 µl GlycoBlue if small quantities were used) were added to precipitate RNA. Samples were incubated at room temperature for 15 min and spun at 12,000 x g and 4°C for 10 min. The supernatant was discarded and the pellet was washed with 500 µl of 75% ethanol prepared in DEPC water. After vortexing the RNA was pelleted by centrifugation at 7000 x g and 4°C for 5 min. The supernatant was removed as completely as possible, the pellet was air-dried for a few minutes and dissolved in 10 – 100 µl RNAse-free water. The RNA was stored at -80°C.

4.4.4 Reverse transcription

Reverse transcription was performed using the QuantiTect® reverse transcription kit (Qiagen) or the SuperScript™ III kit (Invitrogen).

4.4.4.1 cDNA synthesis by QuantiTect® reverse transcription kit

reagent	amount	incubation
RNA	1 µg	
gDNA Wipeout Buffer (7x)	2 µl	5 min at 42°C, then on ice
RNAse free water	To 14 µl	
Quantiscript RT Buffer (5x)	4 µl	30 min at 42°C
Random hexamer primer mix	1 µl	15 min at 95°C
Quantiscript reverse transcriptase	1 µl	∞ at 4°C

The cDNA was stored at -20°C and for qRT-PCR analysis 1 µl was used.

4.4.4.2 cDNA synthesis by SuperScript™ III kit

reagent	amount	incubation
RNA	2 µg	
Random hexamer primer mix (50 ng/µl)	1 µl	5 min at 65°C, then on ice
dNTP mix (10 mM)	1 µl	for at least 1 min
RNAse free water	To 10 µl	

The following master mix was prepared and 10 µl were added to the RNA:

reagent	amount	incubation
RT Buffer (5x)	4 µl	10 min at 25°C 50 min at 50°C 5 min at 85°C ∞ at 4°C
$MgCl_2$ (25 mM)	4 µl	
DTT (0.1 M)	2 µl	
RNAseOUT (40 U/µl)	1 µl	
SuperScript™ III RT (200 U/µl)	1 µl	

The cDNA was treated with 0.4 µl of RNAse H for 20 min at 37°C and stored at -20°C. For qRT-PCR analysis the cDNA was diluted 1:50 and 2 µl were used per reaction.

4.4.5 Quantitative real time PCR (qRT-PCR)

QRT-PCRs were performed in 384-well plates using the ABsolute qPCR SYBR green mix (Thermo Scientific) and the LightCycler 480 (Roche) to monitor fluorescence. Reactions were set up as follows with qRT-PCR primers listed in Table 10:

reagent	amount
cDNA template	1 – 2 µl
ABsolute qPCR SYBR green mix (2x)	5 µl
primer mix (10 µM forwards/ reverse each)	0.5 µl
ddH_2O	To 10 µl

PCR conditions:

temperature [°C]	time	cycles
95	15 min	1
95 60	15 sec 40 sec	50
Melting: 60 – 95 (0.11°C/sec)	-	1
40	10 min	1

4.4.6 High throughput sequencing

Whole genome bisulfite sequencing and total RNA sequencing was performed by the High Throughput Sequencing Unit of the Genomics and Proteomics Core Facility at the DKFZ.

4.4.6.1 Whole Genome Bisulfite Sequencing (WGBS)

Library preparations for whole genome bisulfite sequencing were conducted using the R&D protocol of the Core Facility and 0.3 to 1.2 µg of genomic DNA as starting material. DNA was fragmented to an average size of 400 bp using the Covaris S2 and quality was assessed using the Agilent DNA 1000 Assay on a Bioanalyzer Instrument 2100. The TruSeq DNA Sample Prep v2 Kit (version 2012) was used to perform end repair with bead clean-up and A-tailing according to the manufacturer´s instructions as well as adapter ligation with double bead clean-up. Quality was again monitored on the Bioanalyzer. Bisulfite treatment was conducted using the EpiTect Kit (Qiagen) and the converted DNA was purified following the Illumina protocol for whole genome bisulfite sequencing for Methylation Analysis (Part #15021861, Rev. B). The libraries were amplified using Kapa HiFi HotStart Uracil+ ReadyMix, TruSeq PCR primer cocktail and 14 PCR cycles following the above mentioned Illumina protocol. Libraries were cleaned-up and quality checked using the Qubit dsDNA HS Assay and the Bioanalyzer DNA 1000 Assay. Average library fragment sizes were around 350 bp. Final library concentrations applied to the flow cell were 8 to 10 pM with 1% PhiX control v3 according to the manufacturer´s instructions. DNA was sequenced in paired-end mode on an Illumina HiSeq 2000 V3 system. Read lengths were 101 base pairs and the average insert size was 230 bp. Two biological replicates per condition were sequenced yielding a combined average coverage of 20 x.

4.4.6.2 Total RNA sequencing

Library preparation for RNA sequencing was performed from total RNA using the TruSeq RNA Library Preparation Kit v2 (version 09-2012; Illumina) according to the manufacturer´s instructions. Briefly, 1 µg of total RNA was mRNA enriched using Oligo dT beads, fragmented and first strand synthesis was followed by second strand synthesis and bead clean-up. Ends

were repaired, bead clean-up was performed and dA-tailing took place followed by adapter ligation and double bead clean-up. Libraries were amplified, cleaned again and quality was monitored using the Qubit dsDNA HS Assay Kit and the Bioanalyzer with DNA 1000 Assay. Average library fragment size was around 290 bp. The libraries were sequenced on an Illumina HiSeq 2000 V3 system in 101 bp paired-end mode with 10 pM final library concentration on the flow cell and 1% PhiX control v3 following Illumina´s instructions. The average insert size was 170 bp. Two biological replicates per condition were sequenced.

4.4.7 454 targeted amplicon bisulfite sequencing

For targeted amplicon bisulfite sequencing 500 ng of genomic DNA were bisulfite treated according to instructions of the EpiTect Bisulfite Kit from Qiagen with the following changes:

step	temperature [°C]	time [min]
denaturation	95	5
incubation	60	25
denaturation	95	5
incubation	60	85
denaturation	95	5
incubation	60	295
hold	20	∞

Regions of interest were PCR amplified using sequence specific primers amplifying the deaminated template and containing cell-type specific barcodes and standard 454 adapters. Ideal annealing temperatures for primers were tested by gradient PCR using the predicted annealing temperature of the primers and 2°C steps of decreasing temperature. PCRs were performed using the PyroMark PCR Kit with the following reaction set up:

reagent	amount
PyroMark PCR Master Mix (2x)	12.5 µl
CoralLoad Concentrate (10x)	2.5 µl
$MgCl_2$ (25 mM)	1 µl
primer mix (10 µM forwards/ reverse each)	2.5 µl
template	1.5 µl/ 10-20 ng bisulfite converted DNA
ddH_2O	5 µl
total volume	25 µl

PCR conditions:

temperature [°C]	time	cycles
95	15 min	1
94	30 sec	30 - 35
gradient/ best annealing	30 sec	
72	45 sec	
72	10 min	1
4	∞	

To obtain maximal yields the parameters were adapted for each primer pair varying $MgCl_2$ concentration (1.5 - 4.0 mM), the amounts of template and Q-solution, the number of PCR cycles (maximal 35), primer annealing temperatures and hotstart PCR protocol (if not using the PyroMark polymerase, which requires heat activation). PCR products were analyzed on 2 - 2.5 % agarose gels, excised and gel-extracted using the PeqGold Extraction Kit (PeqLab) or the QIAquick Gel Extraction Kit (Qiagen). Sample concentrations were determined using the Quant-iT^{TM}PicoGreen ®dsDNA Kit according to instructions and the FLUOstar OPTIMA plate reader. For sequencing amplicons were pooled in an equimolar mix, processed and sequenced according to the manufacturer's instructions on the GS Junior 454 Genome Sequencer (Roche). Sequencing reads were processed, filtered, aligned and displayed in color-coded heatmaps using the BiSQuID tool internally available in the Division of Epigenetics. Furthermore the methylation state of each CpG in each read was assessed and the average methylation per CpG site was computed and displayed in color-coded tables by BiSQuID. This tool was programmed by Cassandra Falckenhayn.

4.4.8 EPIC Methylation Array

For methylation profiling using the Infinium MethylationEPIC BeadChip 1 µg of purified genomic DNA was provided. The analysis of DNA on the array was performed by the microarray unit of the DKFZ Genomics and Proteomics Core Facility.
This platform is an extended version of the Infinium 450K array interrogating 93% of the previous and many additional probes resulting in a total of 866,895 CpG sites. Briefly, per experiment and condition two biological replicates, but no technical replicates, were used. Matched untreated or empty vector transduced cells served as controls. DNA concentrations were measured using the Quant-iT^{TM}PicoGreen ®dsDNA Kit and DNA quality was assessed by agarose gel electrophoresis. Samples had to have an average fragment size of at least 3 kb to pass quality control. 500 ng of DNA were bisulfite converted using the EZ-96 DNA Methylation Kit (Zymo Research) according to the manufacturer´s instructions. After whole genome amplification the DNA was fragmented enzymatically and denatured following the recommendations of the Infinium HD Assay Methylation Protocol Guide (Illumina). The DNA was applied to the BeadChip and hybridization took place for 16 to 24 hours at 48°C. During this process DNA molecules anneal to locus-specific DNA oligomers linked to different bead types. In the next step primers are single-base extended using 2,4-dinitrophenol- (DNP) or Biotin-labelled ddNTPs and the array is fluorescently stained and scanned (iScan array scanner, Illumina) to measure red and green fluorescence intensities. DNA methylation values are recorded for each of the 866,895 probes on the array. Methylation level is expressed as beta value which is the ratio of methylated signal over the sum of the

methylated and unmethylated signals. Thus average beta values represent the percentage of methylation of any given cytosine and can attain any value between 0 and 1, with 0 corresponding to no methylation and 1 to full methylation.
Raw IDAT files containing average beta values for each probe were analyzed using an analysis pipeline available in the Division of Epigenetics. This tool was programmed by Julian Gutekunst. Details can be found in chapter 4.6.3.

4.5 Cellular assays

4.5.1 Adipogenic differentiation

Adipogenic differentiation was performed as described in Wiehle et al., 2016. Briefly, 1 x 10^5 cells/well were seeded into 12-well plates and the next day treatment with adipogenic differentiation medium (ADM; Tuorto et al., 2012) was started. After 7, 14, and 21 days cells were harvested or fixated in order to stain lipid droplets. Fixation was performed for 1 hour in 10% formalin. Wells were washed with 60% isopropanol and filtered (0.2 µm) Oil Red O working solution was added for 10 min. Next, the wells were washed in H_2O to remove dye and air-dried. Five random images per well were acquired at 4 fold magnification using a reflected-light microscope (Olympus) and the particle number per image, representing the amount of lipid droplets, was quantified using ImageJ.

4.5.2 2-HG treatment

To supplement cell culture medium with 2-HG, 8.64 mg of D- or L-2-HG were freshly dissolved in 1.5 ml of RPMI standard culture medium yielding a 30 mM stock solution. Aliquots of powder were stored at 4°C (L-2-HG) or -20°C (D-2-HG) and protected from light until use. The solution was rotated for 30 min at room temperature and sterile-filtrated with 0.2 µm filters prior to use. The solution was then further diluted in RPMI to obtain the required concentration of 5 mM. 600,000 cells/well were seeded into a 24-well plate in 1 ml of 2-HG supplemented medium. Every 3rd and 4th day the medium was exchanged with freshly 2-HG supplemented medium, the cell numbers were counted for proliferation analysis and reduced to 600,000 per well. Remaining cells were quickly washed in PBS for two times, pelleted, liquid was removed as completely as possible and pellets were immediately frozen at -20°C for storage.

4.5.3 D-2-HG measurement

To measure D-2-HG levels in cell culture supernatants (extracellular measurement) or cells (intracellular measurement) the same fluorimetric assay was used. For intracellular

measurements cell lysates were prepared and their protein content was assessed. To this end, a visible cell pellet was collected in safe-lock tubes and washed twice in PBS for less than 5 min. Cells were dissolved in NP-40 lysis buffer with freshly added protease inhibitors and freeze-thawed three times by immersing the tubes in liquid nitrogen and placing them at 37°C until the suspension was liquid again. Vortexing was performed in between. After lysis tubes were spun for 5 min at 13,000 x g to pellet debris. The clear supernatant was transferred to fresh tubes and kept on ice until protein determination. For extracellular measurements 1 ml of cell culture supernatant per condition was collected and directly used for deproteinization or stored at -20°C until quantification.

4.5.3.1 BCA protein assay

In order to normalize measured intracellular D-2-HG levels to total cell numbers the protein concentration of the lysates was measured by BCA protein assay. To this end a 1:5 to 1:25 dilution was prepared in H_2O and several BSA standards were set up to create a standard curve. 25 µl of each sample or standard were mixed with 200 µl of a 1:50 dilution of BCA reagent B in reagent A in a 96-well plate. Pure NP-40 lysis buffer or H_2O was used as blank and treated equally. The mixture was incubated for 30 min at 37°C and the OD at 562 nm was measured using the FLUOstar OPTIMA plate reader. Each sample, standard and blank was at least measured in duplicates. The preset 4-parameter fit function of the OPTIMA analysis software was used to establish a standard curve. Raw sample intensities were background subtracted and protein concentrations were automatically calculated by the analysis software.

4.5.3.2 Deproteinization

To remove proteins from cell lysates or cell culture medium a deproteinizing kit (BioVision) was used. In brief, 100 µl of sample were mixed with 25 µl of perchloric acid (PCA) in 8-strip PCR tubes and incubated on ice for 2 min. To prepare a standard curve 95 µl of corresponding cell culture medium (extracellular measurement) or NP-40 lysis buffer (intracellular measurement) were mixed with 5 µl of each standard and subjected to PCA treatment as described for samples. Standards were prepared in medium (extracellular measurement) or ddH_2O (intracellular measurement). Tubes were centrifuged for 7 min at 14,000 rpm and room temperature before transferring 100 µl of the supernatant to 5 µl chilled neutralization solution in new 8-strip tubes. The mixture was incubated for 2 min on ice and spun again as before to pellet precipitated proteins.

4.5.3.3 D-2-HG assay

Of the supernatant 95 µl were transferred to fresh tubes and spun quickly. 25 µl were transferred into a black 96-well plate in triplicates and mixed with 75 µl of the following master mix:

reagent	µl/well	µl/96-well plate
HEPES pH 8 (1M)	10	1050
NAD+ (10 mM)	1	105
HGDH (0.1 µg/µl)	1	105
Diaphorase (0.1 U/µl)	1	105
Resazurin (125 µM)	4	420
ddH_2O	58	6090

The mixture was incubated in the dark at room temperature for 15 to 60 min and fluorimetric detection was carried out using the FLUOstar OPTIMA plate reader with excitation of 540+/-10 nm and emission of 610+/-10 nm. The standard values were curve fitted using the preset 4-parameter or linear regression fit of the OPTIMA analysis software and D-2-HG values in pmol/25 µl sample were computed by the software. To normalize for protein content the obtained values [pmol/25 µl] were divided by the corresponding protein concentrations [µg/25 µl]. To obtain µM values the obtained concentrations were multiplied by 0.04.

4.5.4 Proliferation assay by Cell Titer Glo

At day 0 cells were seeded at 1000 cells/well into a 96-well plate in triplicates. At day 1, 3 and 6 metabolic activity was assessed by the Cell Titer Glo assay: Cell culture plates were equilibrated to room temperature for 30 min and the Cell titer Glo solution was prepared following the manufacturer´s instructions. Per well 10 µl were added, the plate was shaken for two minutes to lyse cells and left for 10 min in the dark. Luminescence was recorded using the OPTIMA plate reader with a gain of 3000. Triplicates were background-corrected and averaged.

4.5.5 Proliferation assay by assessment of cumulative population doubling

Equal numbers of cells (e.g. 500,000) were seeded in triplicates at each new passage (number of inoculated cells = N_I) and after 3 or 4 days cell numbers in each well were determined. The average was calculated (average number of cells harvested = N_H) and used to compute the average population doubling level (PDL) based on the following formula:

$$PDL = 3.32\ x\ ((\log N_H) - (\log N_I))$$

To obtain the cumulative population doubling level the PDL obtained from each measurement was added to the previous one (or to 0 at the first measurement) and plotted against the time that passed between seeding and harvest.

4.6 High throughput data analysis

4.6.1 Whole genome bisulfite sequencing

Quality control of sequencing reads was performed using FastQC. Reads were trimmed and mapped to the mouse reference genome (NCBI37/mm9) using BSMAP 2.5 (Xi and Li, 2009) and the Picard tool to remove duplicates (http://broadinstitute.github.io/picard). Methylation ratios were calculated with a Python script (methratio.py) from the BSMAP package. PMDs that were previously reported to harbor instable methylation in cell culture models (Gaidatzis et al., 2014), were removed using a sliding window of 100 kb as described before (Raddatz et al., 2012), the *Mus musculus* skin methylome as reference and a 2-state first-order hidden Markov model to identify regions with a strong reduction in methylation (PMDs). Publically available ChIP-seq datasets of different chromatin marks, CTCF and PolII in MEFs (Mikkelsen et al., 2007; Shen et al., 2012) were subjected to ChromHMM analysis (Ernst and Kellis, 2012; http://compbio.mit.edu/ChromHMM/) in order to segment the mouse genome into 15 chromatin states, for which methylation ratios were calculated and compared. DNA methylation canyons were mapped as described in Jeong et al., 2013 and using MOABS (Sun et al., 2014). Canyon-associated genes were determined by the presence of their TSS within the canyon. More details are described in Wiehle et al., 2016. Günter Raddatz conducted the analyses.

4.6.2 RNA sequencing

Quality control of sequencing reads was performed using FastQC. Reads were adapter-trimmed and terminal stretches of bases with a Phred quality score of <30 were removed resulting in a maximal read length of 80 bp. Reads were mapped to the NCBI37/mm9 assembly of the mouse genome using TopHat 2.0.6 (Trapnell, Pachter, and Salzberg, 2009). To identify differentially expressed transcripts DESeq 1.10.1 (Anders and Huber, 2010) and Cuffdiff 2.0 (Trapnell et al., 2013) were applied. To avoid false positives p-values were subjected to multiple testing correction and transcripts with a corrected p-value (q-value) of <0.05 were considered to be differentially expressed. These analyses were performed by Günter Raddatz and are described in Wiehle et al., 2016.

4.6.3 Methylation array

The raw intensity data files were normalized, quality-filtered and statistically analyzed as described before (Geyh et al., 2016). In brief, the minfi package (version 1.20.2; Aryee et al., 2014) was used to load IDAT files into R (version 3.3.2). Data were normalized with the SWAN method (Maksimovic et al., 2012) without prior background correction. Cross-reactive and SNP containing probes (Chen et al., 2013b), probes on the sex chromosomes and probes with a low detection p-value (>0.01) were omitted. Differentially methylated probes were identified by a similar approach as implemented in minfi with the exception of adjusting p-values for false discovery rate using the Benjamini-Hochberg procedure. Probes with an adjusted p-value of <0.05 were considered as differentially methylated. PCA graphs were generated by the plotPCA() function provided by the R affycoretools package (version 1.46.5; MacDonald, 2008) with significantly differentially methylated probes used to compute principal components. Heatmaps were created by the heatmap.2() function of the R gplots package (version 3.0.1, Warnes et al., 2016). For clustering the euclidean distance function was used to compute the distance matrix and the *complete* linkage method to obtain hierarchical clustering. The RColorBrewer package (version 1.1-2) was used to create the assigned colors (Neuwirth, 2014). All above mentioned packages were obtained from Bioconductor. To be able to compare EPIC and 450K data the probes common to both chips were extracted and used for analysis. The bioinformatic pipeline containing all the described functions was programmed by Julian Gutekunst.

4.7 Statistical analysis

Numerical data, such as expression values or concentrations, are presented as arithmetic averages. Standard deviations were computed as a measure of statistical spread. The two-tailed Student´s t-test was used to determine the significance of observed differences between means of gene expression of two groups. The required level of significance was 5% ($p \leq 0.05$). Enrichments of biological pathways among deregulated genes were calculated using the Ingenuity Pathway Analysis tool and figures showing the negative logarithm of the p-values were extracted from the software.

5 Appendix

5.1 Supplemental Figures

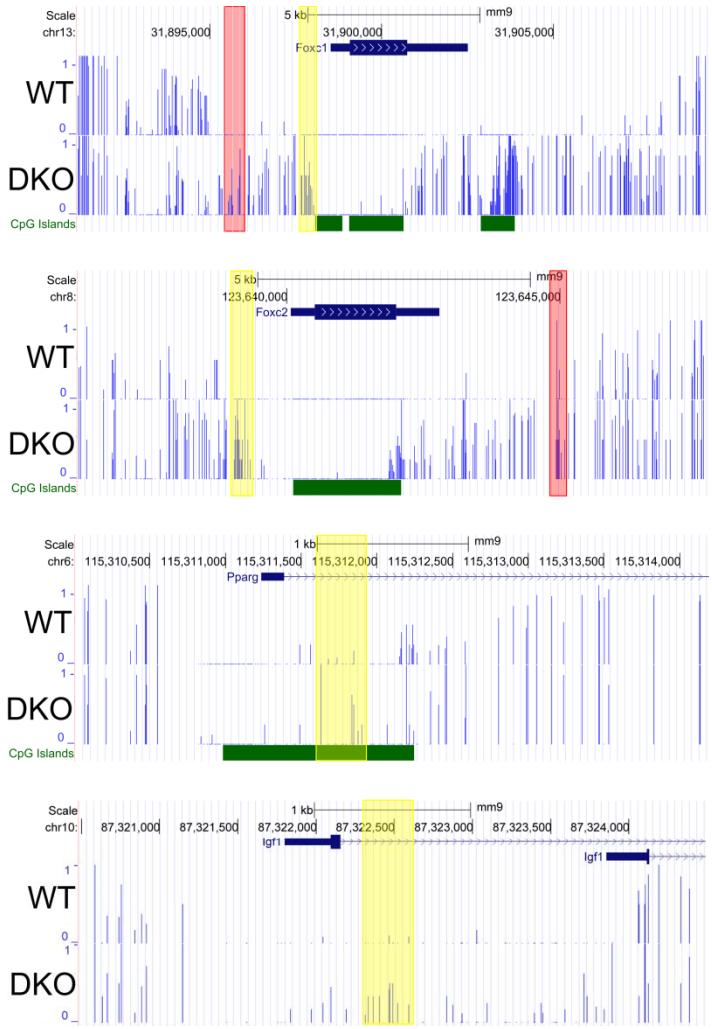

Figure S1: Whole genome bisulfite sequencing tracks of WT and DKO MEFs at selected hypermethylated canyons.

Canyons associated with the genes *Foxc1, Foxc2, Pparγ* and *Igf1* are shown. Each blue bar represents the detected methylation ratio at individual CpGs. Red boxes indicate the loci interrogated by targeted amplicon bisulfite sequencing in Figure 2.8 and Figure 2.14 and yellow boxes show the amplicon localization of Figure 2.13.

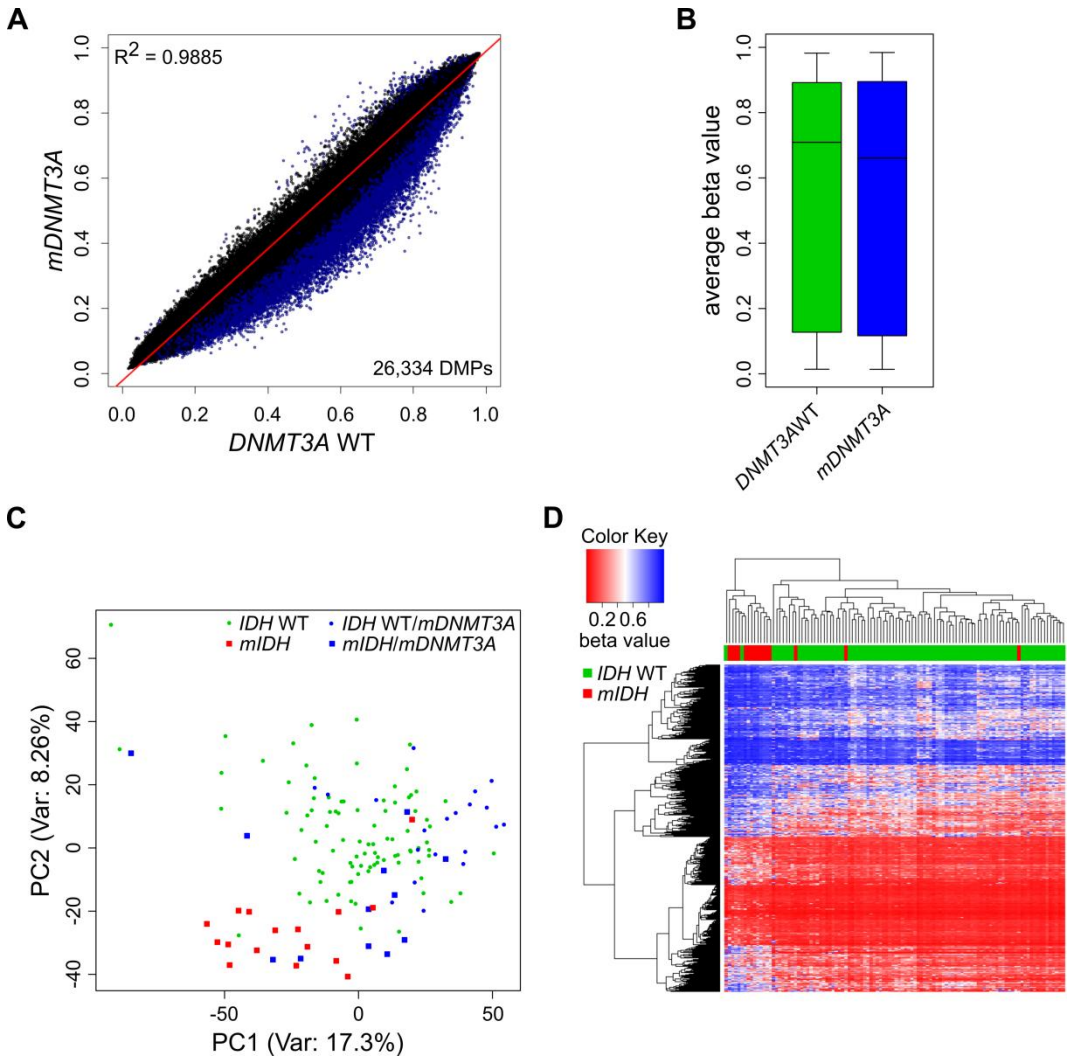

Figure S2: DNA methylation changes in *DNMT3A* mutated AML patients.

(A) Scatterplot comparing 32 AML patients with mutations in *DNMT3A* (*mDNMT3A*) to 108 AML patients with *DNMT3A* WT status. Each dot indicates the beta values of one 450K probe retained after quality-filtering in the two groups and significantly differentially methylated probes (P<0.05) are depicted in blue. **(B)** Distribution of the average beta values of all 450K probes retained after filtering in the group of *mDNMT3A* and *DNMT3A* WT patients. **(C)** Principal component analysis as in Figure 2.16A. Patients with a *DNMT3A* mutation are colored in blue. **(D)** Heatmap and hierarchical cluster dendrogram of *mIDH* and *IDH* WT patients as in Figure 2.16B. Patients with mutated *DNMT3A* were entirely removed from the dataset, resulting in comparison of 92 *IDH* WT to 16 *mIDH* patients.

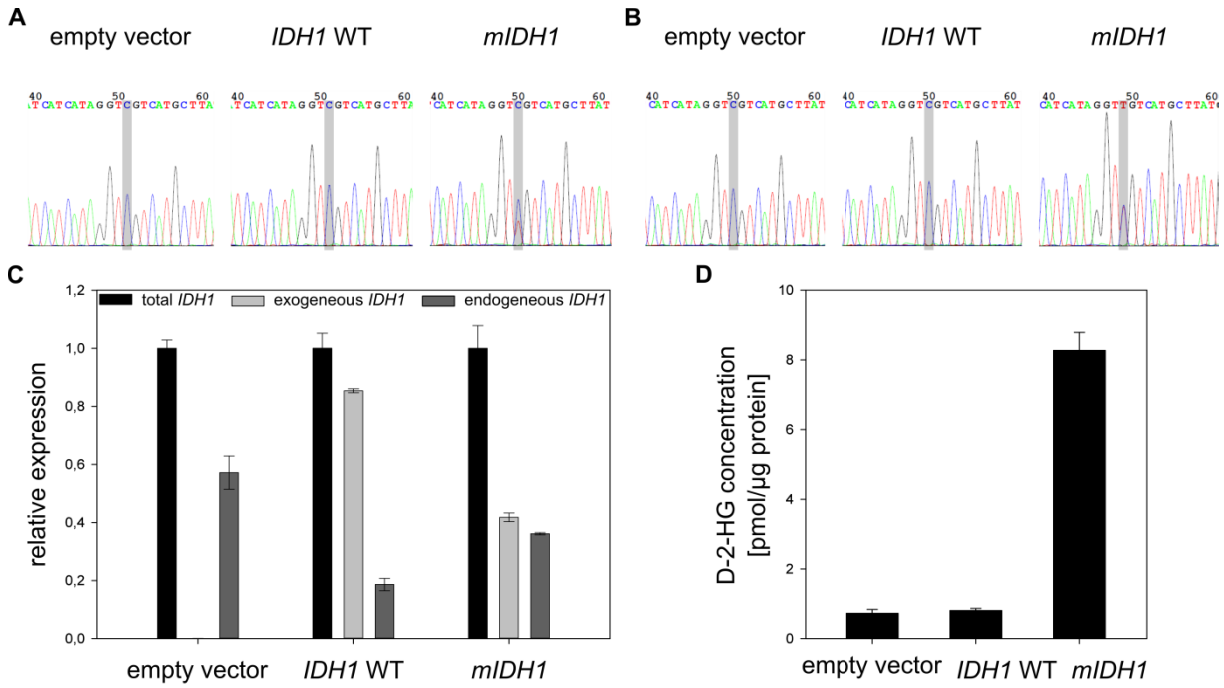

Figure S3: Improved screening methods detect transgene expression and D-2-HG production in NOMO-1 cells.

Electropherograms of the cDNA sequence harboring the *IDH1* R132C mutation in NOMO-1 **(A)** and HL-60 **(B)** cells transduced with empty vector, *IDH1* WT or *IDH1* R132C. Grey boxes indicate the position of the base exchange in the *mIDH* gene. **(C)** QRT-PCR of total, endogenous and exogenous *IDH1* transcripts in transduced NOMO-1 cells using discriminating primers. *ACTB* served as reference gene. Bars represent means and error bars standard deviations of three replicates. **(D)** Intracellular D-2-HG measurement in transduced NOMO-1 cells.

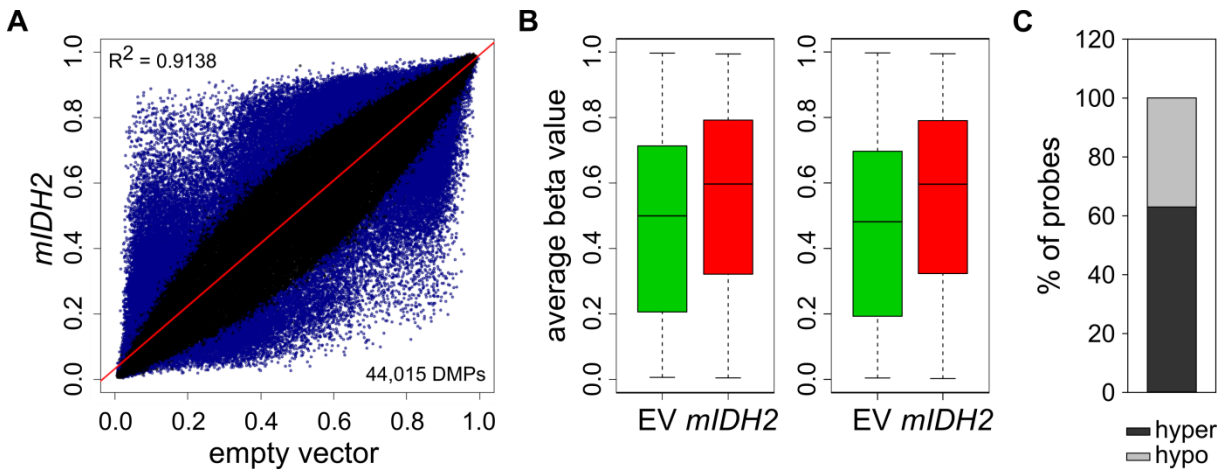

Figure S4: Removal of probes located in PMDs retains *mIDH2*-associated hypomethylation in HL-60 cells.

(A) Scatterplot comparing average beta values of non-PMD CpG probes between HL-60 cells transduced with *mIDH2* and empty vector. Blue dots indicate significantly differentially methylated probes (P<0.05). **(B)** Boxplots showing average beta values for all significantly differentially methylated probes residing outside of PMDs in HL-60 cells transduced with *mIDH2* and empty vector (EV; left panel). In the right panel all significantly differentially methylated probes are shown irrespective of their localization in PMDs. **(C)** Percentage of hyper- and hypomethylated non-PMD probes identified in *mIDH2*-expressing HL-60 cells.

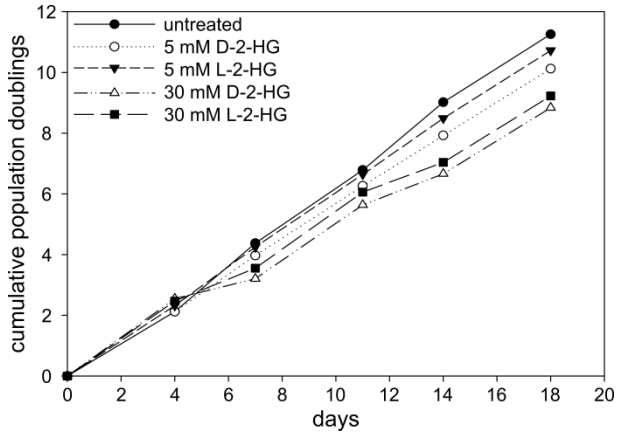

Figure S5: Reduced proliferation of 2-HG treated HL-60 cells.

Cumulative population doublings of HL-60 cells cultured in medium supplied with the indicated concentrations of D- or L-2-HG for 18 days.

5.2 References

Abdel-Wahab, O., Mullally, A., Hedvat, C., Garcia-Manero, G., Patel, J., Wadleigh, M., Malinge, S., Yao, J., Kilpivaara, O., Bhat, R., et al. (2009). Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. Blood *114*, 144–147.

Accomando, W.P., Wiencke, J.K., Houseman, E.A., Nelson, H.H., and Kelsey, K.T. (2014). Quantitative reconstruction of leukocyte subsets using DNA methylation. Genome Biol. *15*, R50.

Achwal, C.W., Iyer, C.A., and Chandra, H.S. (1983). Immunochemical evidence for the presence of 5mC, 6mA and 7mG in human, Drosophila and mealybug DNA. FEBS Lett. *158*, 353–358.

Akalin, A., Garrett-Bakelman, F.E., Kormaksson, M., Busuttil, J., Zhang, L., Khrebtukova, I., Milne, T.A., Huang, Y., Biswas, D., Hess, J.L., et al. (2012). Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. PLoS Genet. *8*.

Alharbi, R.A., Pettengell, R., Pandha, H.S., and Morgan, R. (2013). The role of HOX genes in normal hematopoiesis and acute leukemia. Leukemia *27*, 1000–1008.

Álvarez-Errico, D., Vento-Tormo, R., Sieweke, M., and Ballestar, E. (2014). Epigenetic control of myeloid cell differentiation, identity and function. Nat. Rev. Immunol. *15*, 7–17.

Amouroux, R., Nashun, B., Shirane, K., Nakagawa, S., Hill, P.W.S., D'Souza, Z., Nakayama, M., Matsuda, M., Turp, A., Ndjetehe, E., et al. (2016). De novo DNA methylation drives 5hmC accumulation in mouse zygotes. Nat. Cell Biol. *18*, 225–233.

An, J., González-Avalos, E., Chawla, A., Jeong, M., López-Moyado, I.F., Li, W., Goodell, M. a., Chavez, L., Ko, M., and Rao, A. (2015). Acute loss of TET function results in aggressive myeloid cancer in mice. Nat. Commun. *6*, 10071.

An, J., Rao, A., and Ko, M. (2017). TET family dioxygenases and DNA demethylation in stem cells and cancers. Exp. Mol. Med. *49*, e323.

Ancey, P.-B., Ecsedi, S., Lambert, M.-P., Talukdar, F.R., Cros, M.-P., Glaise, D., Narvaez, D.M., Chauvet, V., Herceg, Z., Corlu, A., et al. (2017). TET-Catalyzed 5-Hydroxymethylation Precedes HNF4A Promoter Choice during Differentiation of Bipotent Liver Progenitors. Stem Cell Reports.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Aravin, A.A., and Bourc'his, D. (2008). Small RNA guides for de novo DNA methylation in mammalian germ cells. Genes Dev. *22*, 970–975.

Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., and Irizarry, R.A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics *30*, 1363–1369.

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., and Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. Nat. Chem. *6*, 1049–1055.

Badal, V., Chuang, L.S.H., Tan, E.H.-H., Badal, S., Villa, L.L., Wheeler, C.M., Li, B.F.L., and Bernard, H.-U. (2003). CpG methylation of human papillomavirus type 16 DNA in cervical

cancer cell lines and in clinical specimens: genomic hypomethylation correlates with carcinogenic progression. J. Virol. *77*, 6227–6234.

Balss, J., Pusch, S., Beck, A.C., Herold-Mende, C., Krämer, A., Thiede, C., Buckel, W., Langhans, C.D., Okun, J.G., and Von Deimling, A. (2012). Enzymatic assay for quantitative analysis of (d)-2-hydroxyglutarate. Acta Neuropathol. *124*, 883–891.

Bard, J.B.L., Lam, M.S., and Aitken, S. (2008). A bioinformatics approach for identifying candidate transcriptional regulators of mesenchyme-to-epithelium transitions in mouse embryos. Dev. Dyn. *237*, 2748–2754.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243–247.

Baylin, S.B., and Jones, P.A. (2016). Epigenetic Determinants of Cancer. Cold Spring Harb. Perspect. Biol. *8*, a019505.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482–485.

Berman, B.P., Weisenberger, D.J., Aman, J.F., Hinoue, T., Ramjan, Z., Liu, Y., Noushmehr, H., Lange, C.P.E., van Dijk, C.M., Tollenaar, R.A.E.M., et al. (2011). Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina–associated domains. Nat. Genet. *44*, 40–46.

Bertrand, N., Castro, D.S., and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. Nat. Rev. Neurosci. *3*, 517–530.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.

Bird, A.P. (1995). Gene number, noise reduction and biological complexity. Trends Genet. *11*, 94–100.

Bird, A.P., and Wolffe, A.P. (1999). Methylation-Induced Repression— Belts, Braces, and Chromatin. Cell *99*, 451–454.

Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martínez, J. a, Goyal, P., Mahapatra, S., Tam, A., Laird, D.J., Hirst, M., Rao, A., et al. (2013). Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature *500*, 222–226.

Bock, C., Beerman, I., Lien, W.H., Smith, Z.D., Gu, H., Boyle, P., Gnirke, A., Fuchs, E., Rossi, D.J., and Meissner, A. (2012). DNA Methylation Dynamics during In Vivo Differentiation of Blood and Skin Stem Cells. Mol. Cell *47*, 633–647.

Bocker, M.T., Hellwig, I., Breiling, A., Eckstein, V., Ho, A.D., and Lyko, F. (2011). Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. *117*, 182–189.

Bocker, M.T., Tuorto, F., Raddatz, G., Musch, T., Yang, F.-C., Xu, M., Lyko, F., and Breiling, A. (2012). Hydroxylation of 5-methylcytosine by TET2 maintains the active state of the mammalian HOXA cluster. Nat. Commun. *3*, 818.

Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W., and Balasubramanian, S. (2012). Quantitative Sequencing of 5-Methylcytosine and 5-Hydroxymethylcytosine at Single-Base Resolution. Science (80-.). *336*, 934–937.

Bormann, F., Rodríguez-Paredes, M., Hagemann, S., Manchanda, H., Kristof, B., Gutekunst,

J., Raddatz, G., Haas, R., Terstegen, L., Wenck, H., et al. (2016). Reduced DNA methylation patterning and transcriptional connectivity define human skin aging. Aging Cell *15*, 563–571.

Bostick, M., Kim, J.K., Estève, P.-O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science *317*, 1760–1764.

Breiling, A., and Lyko, F. (2015). Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. Epigenetics Chromatin *8*, 24.

Bröske, A.-M., Vockentanz, L., Kharazi, S., Huska, M.R., Mancini, E., Scheller, M., Kuhl, C., Enns, A., Prinz, M., Jaenisch, R., et al. (2009). DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. Nat. Genet. *41*, 1207–1215.

Busque, L., Patel, J.P., Figueroa, M.E., Vasanthakumar, A., Provost, S., Hamilou, Z., Mollica, L., Li, J., Viale, A., Heguy, A., et al. (2012). Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. Nat. Genet. *44*, 1179–1181.

de Capoa, A., Musolino, A., Della Rosa, S., Caiafa, P., Mariani, L., Del Nonno, F., Vocaturo, A., Donnorso, R.P., Niveleau, A., and Grappelli, C. DNA demethylation is directly related to tumour progression: evidence in normal, pre-malignant and malignant cells from uterine cervix samples. Oncol. Rep. *10*, 545–549.

Challen, G.A., Sun, D., Mayle, A., Jeong, M., Luo, M., Rodriguez, B., Mallaney, C., Celik, H., Yang, L., Xia, Z., et al. (2014). Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells. Cell Stem Cell *15*, 350–364.

Challen, G. a, Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J.S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., et al. (2011). Dnmt3a is essential for hematopoietic stem cell differentiation. Nat. Genet. *44*, 23–31.

Chan, S.M., Thomas, D., Corces-Zimmerman, M.R., Xavy, S., Rastogi, S., Hong, W.-J., Zhao, F., Medeiros, B.C., Tyvoll, D.A., and Majeti, R. (2015). Isocitrate dehydrogenase 1 and 2 mutations induce BCL-2 dependence in acute myeloid leukemia. Nat. Med. *21*, 178–184.

Chaturvedi, A., Araujo Cruz, M.M., Jyotsana, N., Sharma, A., Goparaju, R., Schwarzer, A., Görlich, K., Schottmann, R., Struys, E.A., Jansen, E.E., et al. (2016). Enantiomer-specific and paracrine leukemogenicity of mutant IDH metabolite 2-hydroxyglutarate. Leukemia 1–8.

Chaturvedi, A., Herbst, L., Pusch, S., Klett, L., Goparaju, R., Stichel, D., Kaulfuss, S., Panknin, O., Zimmermann, K., Toschi, L., et al. (2017). Pan-mutant-IDH1 inhibitor BAY1436032 is highly effective against human IDH1 mutant acute myeloid leukemia in vivo. Leukemia.

Chen, C., Liu, Y., Lu, C., Cross, J.R., Morris, J.P., Shroff, A.S., Ward, P.S., Bradner, J.E., Thompson, C., and Lowe, S.W. (2013a). Cancer-associated IDH2 mutants drive an acute myeloid leukemia that is susceptible to Brd4 inhibition. Genes Dev. *27*, 1974–1985.

Chen, J.-Y., Lai, Y.-S., Tsai, H.-J., Kuo, C.-C., Yen, B.L., Yeh, S.-P., Sun, H.S., and Hung, W.-C. (2016). The oncometabolite R-2-hydroxyglutarate activates NF-κB-dependent tumor-promoting stromal niche for acute myeloid leukemia cells. Sci. Rep. *6*, 32428.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol. Cell. Biol. *23*, 5594–5605.

Chen, Y., Lemire, M., Choufani, S., Butcher, D.T., Grafodatskaya, D., Zanke, B.W., Gallinger, S., Hudson, T.J., and Weksberg, R. (2013b). Discovery of cross-reactive probes and

polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics *8*, 203–209.

Chowdhury, R., Yeoh, K.K., Tian, Y.-M., Hillringhaus, L., Bagg, E.A., Rose, N.R., Leung, I.K.H., Li, X.S., Woon, E.C.Y., Yang, M., et al. (2011). The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. EMBO Rep. *12*, 463–469.

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science *277*, 1996–2000.

Cimmino, L., Dawlaty, M.M., Ndiaye-Lobry, D., Yap, Y.S., Bakogianni, S., Yu, Y., Bhattacharyya, S., Shaknovich, R., Geng, H., Lobry, C., et al. (2015). TET1 is a tumor suppressor of hematopoietic malignancy. Nat. Immunol. *16*, 653–662.

Clark, S.J., Harrison, J., and Molloy, P.L. (1997). Sp1 binding is inhibited by (m)Cp(m)CpG methylation. Gene *195*, 67–71.

Corces-Zimmerman, M.R., Hong, W.-J., Weissman, I.L., Medeiros, B.C., and Majeti, R. (2014). Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proc. Natl. Acad. Sci. *111*, 2548–2553.

Cortázar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A.L., Siegrist, F., et al. (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature *470*, 419–423.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., et al. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67–79.

Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T. a, Shliaha, P. V, Fidalgo, M., Saunders, A., Lawrence, M., Dietmann, S., et al. (2013). NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. Nature.

Coulondre, C., Miller, J.H., Farabaugh, P.J., and Gilbert, W. (1978). Molecular basis of base substitution hotspots in Escherichia coli. Nature *274*, 775–780.

Dai, H.-Q., Wang, B.-A., Yang, L., Chen, J.-J., Zhu, G.-C., Sun, M.-L., Ge, H., Wang, R., Chapman, D.L., Tang, F., et al. (2016). TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling. Nature.

Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C., et al. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature *462*, 739–744.

Dang, L., Yen, K., and Attar, E.C. (2016). IDH mutations in cancer and progress toward development of targeted therapeutics. Ann. Oncol. *27*, 599–608.

Davis, K.E., Moldes, M., and Farmer, S.R. (2004). The forkhead transcription factor FoxC2 inhibits white adipocyte differentiation. J. Biol. Chem. *279*, 42453–42461.

Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.-C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.-W., Page, D.C., et al. (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell *9*, 166–175.

Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M., et al. (2013). Combined deficiency of tet1 and tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev. Cell *24*, 310–323.

Dawlaty, M.M., Breiling, A., Le, T., Barrasa, M.I., Raddatz, G., Gao, Q., Powell, B.E., Cheng, A.W., Faull, K.F., Lyko, F., et al. (2014). Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Dev. Cell *29*, 102–111.

Dawson, M.A. (2017). The cancer epigenome: Concepts, challenges, and therapeutic opportunities. Science (80-.). *355*, 1147–1152.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. *25*, 1010–1022.

Delhommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Massé, A., Kosmider, O., Le Couedic, J.-P., Robert, F., Alberdi, A., et al. (2009). Mutation in TET2 in myeloid cancers. N. Engl. J. Med. *360*, 2289–2301.

DiNardo, C.D., Ravandi, F., Agresta, S., Konopleva, M., Takahashi, K., Kadia, T., Routbort, M., Patel, K.P., Mark Brandt, Pierce, S., et al. (2015). Characteristics, clinical outcome, and prognostic significance of IDH mutations in AML. Am. J. Hematol. *90*, 732–736.

Doege, C.A., Inoue, K., Yamashita, T., Rhee, D.B., Travis, S., Fujita, R., Guarnieri, P., Bhagat, G., Vanti, W.B., Shih, A., et al. (2012). Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. Nature *488*, 652–655.

Döhner, H., Weisdorf, D.J., and Bloomfield, C.D. (2015). Acute Myeloid Leukemia. N. Engl. J. Med. *373*, 1136–1152.

Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., et al. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat. Genet. *41*, 1350–1353.

Domcke, S., Bardet, A.F., Adrian Ginno, P., Hartl, D., Burger, L., and Schübeler, D. (2015). Competition between DNA methylation and transcription factors determines binding of NRF1. Nature *528*, 575–579.

Dubois-Chevalier, J., Oger, F., Dehondt, H., Firmin, F.F., Gheeraert, C., Staels, B., Lefebvre, P., and Eeckhoute, J. (2014). A dynamic CTCF chromatin binding landscape promotes DNA hydroxymethylation and transcriptional induction of adipocyte differentiation. Nucleic Acids Res. *42*, 1–17.

Duncan, C.G., Barwick, B.G., Jin, G., Rago, C., Kapoor-Vazirani, P., Powell, D.R., Chi, J.T., Bigner, D.D., Vertino, P.M., and Yan, H. (2012). A heterozygous IDH1R132H/WT mutation induces genome-wide alterations in DNA methylation. Genome Res. *22*, 2339–2355.

Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V.K., Attwood, J., Burger, M., Burton, J., Cox, T. V, Davies, R., Down, T.A., et al. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. Nat. Genet. *38*, 1378–1385.

Edwards, J.R., Yarychkivska, O., Boulard, M., and Bestor, T.H. (2017). DNA methylation and DNA methyltransferases. Epigenetics Chromatin *10*, 23.

Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res. *10*, 2709–2721.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods *9*, 215–216.

Farlik, M., Halbritter, F., M??ller, F., Choudry, F.A., Ebert, P., Klughammer, J., Farrow, S.,

Santoro, A., Ciaurro, V., Mathur, A., et al. (2016). DNA Methylation Dynamics of Human Hematopoietic Stem Cell Differentiation. Cell Stem Cell *19*, 808–822.

Fathi, A.T., Sadrzadeh, H., Borger, D.R., Ballen, K.K., Amrein, P.C., Attar, E.C., Foster, J., Burke, M., Lopez, H.U., Matulis, C.R., et al. (2012). Prospective serial evaluation of 2-hydroxyglutarate, during treatment of newly diagnosed acute myeloid leukemia, to assess disease activity and therapeutic response. Blood *120*, 4649–4652.

Feinberg, A.P., and Tycko, B. (2004). The history of cancer epigenetics. Nat. Rev. Cancer *4*, 143–153.

Feinberg, A.P., and Vogelstein, B. (1983a). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature *301*, 89–92.

Feinberg, A.P., and Vogelstein, B. (1983b). Hypomethylation of ras oncogenes in primary human cancers. Biochem. Biophys. Res. Commun. *111*, 47–54.

Feldmann, A., Ivanek, R., Murr, R., Gaidatzis, D., Burger, L., and Schübeler, D. (2013). Transcription Factor Occupancy Can Mediate Active Turnover of DNA Methylation at Regulatory Regions. PLoS Genet. *9*, e1003994.

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T. a, Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature *473*, 398–402.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., et al. (2010). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell *18*, 553–567.

Flavahan, W. a., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suvà, M.L., and Bernstein, B.E. (2015). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature 1–16.

Fujiki, K., Shinoda, A., Kano, F., Sato, R., Shirahige, K., and Murata, M. (2013). PPARγ-induced PARylation promotes local DNA demethylation by production of 5-hydroxymethylcytosine. Nat. Commun. *4*, 2262.

Gaidatzis, D., Burger, L., Murr, R., Lerch, A., Dessus-Babus, S., Schübeler, D., and Stadler, M.B. (2014). DNA Sequence Explains Seemingly Disordered Methylation Levels in Partially Methylated Domains of Mammalian Genomes. PLoS Genet. *10*, e1004143.

Gaidzik, V.I., Paschka, P., Späth, D., Habdank, M., Köhne, C.-H., Germing, U., von Lilienfeld-Toal, M., Held, G., Horst, H.-A., Haase, D., et al. (2012). TET2 mutations in acute myeloid leukemia (AML): results from a comprehensive genetic and clinical analysis of the AML study group. J. Clin. Oncol. *30*, 1350–1357.

Gama-Sosa, M.A., Slagel, V.A., Trewyn, R.W., Oxenhandler, R., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983). The 5-methylcytosine content of DNA from human tumors. Nucleic Acids Res. *11*, 6883–6894.

Gao, Y., Chen, J.J.J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y., et al. (2013). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. Cell Stem Cell *12*, 453–469.

Genovese, G., Kähler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., et al. (2014). Clonal hematopoiesis and

blood-cancer risk inferred from blood DNA sequence. N. Engl. J. Med. *371*, 2477–2487.

Geyh, S., Oz, S., Cadeddu, R.-P., Fröbel, J., Brückner, B., Kündgen, A., Fenk, R., Bruns, I., Zilkens, C., Hermsen, D., et al. (2013). Insufficient stromal support in MDS results from molecular and functional deficits of mesenchymal stromal cells. Leukemia *27*, 1841–1851.

Geyh, S., Rodríguez-Paredes, M., Jäger, P., Khandanpour, C., Cadeddu, R.-P., Gutekunst, J., Wilk, C.M., Fenk, R., Zilkens, C., Hermsen, D., et al. (2016). Functional inhibition of mesenchymal stromal cells in acute myeloid leukemia. Leukemia *30*, 683–691.

Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., Brückl, T., Biel, M., and Carell, T. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. PLoS One *5*, e15367.

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. *74*, 481–514.

Graff, J.R., Herman, J.G., Lapidus, R.G., Chopra, H., Xu, R., Jarrard, D.F., Isaacs, W.B., Pitha, P.M., Davidson, N.E., and Baylin, S.B. (1995). E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. Cancer Res. *55*, 5195–5199.

Gross, S., Cairns, R.A., Minden, M.D., Driggers, E.M., Bittinger, M.A., Jang, H.G., Sasaki, M., Jin, S., Schenkein, D.P., Su, S.M., et al. (2010). Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. J. Exp. Med. *207*, 339–344.

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-L.G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X., Jin, S., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477*, 606–610.

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

Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H., Wu, X., Wen, L., Gu, T.-P., Hu, B., et al. (2014). Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. Cell Stem Cell *15*, 447–458.

Hackett, J. a, Sengupta, R., Zylicz, J.J., Murakami, K., Lee, C., Down, T. a, and Surani, M.A. (2013). Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science *339*, 448–452.

Hader, C., Marlier, A., and Cantley, L. (2010). Mesenchymal?epithelial transition in epithelial response to injury: the role of Foxc2. Oncogene *29*, 1031–1040.

Hajkova, P., Jeffries, S.J., Lee, C., Miller, N., Jackson, S.P., and Surani, M.A. (2010). Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science *329*, 78–82.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Hashimoto, H., Liu, Y., Upadhyay, A.K., Chang, Y., Howerton, S.B., Vertino, P.M., Zhang, X., and Cheng, X. (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res. *40*, 4841–4849.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al.

(2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303–1307.

Hellman, A., and Chess, A. (2007). Gene Body-Specific Methylation on the Active X Chromosome. Science (80-.). *315*, 1141–1143.

Herman, J.G., Latif, F., Weng, Y., Lerman, M.I., Zbar, B., Liu, S., Samid, D., Duan, D.S., Gnarra, J.R., and Linehan, W.M. (1994). Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc. Natl. Acad. Sci. U. S. A. *91*, 9700–9704.

Heuser, M., Yun, H., Berg, T., Yung, E., Argiropoulos, B., Kuchenbauer, F., Park, G., Hamwi, I., Palmqvist, L., Lai, C.K., et al. (2011). Cell of origin in AML: susceptibility to MN1-induced transformation is regulated by the MEIS1/AbdB-like HOX protein complex. Cancer Cell *20*, 39–52.

Hodges, E., Molaro, A., Dos Santos, C.O., Thekkat, P., Song, Q., Uren, P.J., Park, J., Butler, J., Rafii, S., McCombie, W.R., et al. (2011). Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. Mol. Cell *44*, 17–28.

Holz-Schietinger, C., Matje, D.M., and Reich, N.O. (2012). Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. J. Biol. Chem. *287*, 30941–30951.

Hon, G.C., Hawkins, R.D., Caballero, O.L., Lo, C., Lister, R., Pelizzola, M., Valsesia, A., Ye, Z., Kuan, S., Edsall, L.E., et al. (2012). Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. Genome Res. *22*, 246–258.

Hon, G.C.C., Song, C.-X., Du, T., Jin, F., Selvaraj, S., Lee, A.Y.Y., Yen, C., Ye, Z., Mao, S., Wang, B., et al. (2014). 5mC Oxidation by Tet2 Modulates Enhancer Activity and Timing of Transcriptome Reprogramming during Differentiation. Mol. Cell *56*, 1–12.

Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. J. Biol. Chem. *175*, 315–332.

Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y.G., Zhu, J., Wang, P., and Xu, Y. (2013). Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. Cell *155*, 1545–1555.

Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., Gong, W., et al. (2015). Structural insight into substrate preference for TET-mediated oxidation. Nature *527*, 118–122.

Hu, X., Zhang, L., Mao, S., Li, Z., Chen, J., Zhang, R.-R., Wu, H.-P., Gao, J., Guo, F., Liu, W., et al. (2014). Tet and TDG Mediate DNA Demethylation Essential for Mesenchymal-to-Epithelial Transition in Somatic Cell Reprogramming. Cell Stem Cell *14*, 1–11.

Huang, H., Jiang, X., Li, Z., Li, Y., Song, C., He, C., Sun, M., and Chen, P. (2013a). TET1 plays an essential oncogenic role in MLL -rearranged leukemia. 1–6.

Huang, S., Zhu, Z., Wang, Y.Y., Xu, L., Chen, X., Xu, Q., Zhang, Q., Zhao, X., Yu, Y., and Wu, D. (2013b). Tet1 is required for Rb phosphorylation during G1/S phase transition. Biochem. Biophys. Res. Commun. *434*, 241–244.

Huang, Y., Chavez, L., Chang, X., Wang, X., Pastor, W. a, Kang, J., Zepeda-Martínez, J. a, Pape, U.J., Jacobsen, S.E., Peters, B., et al. (2014). Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *111*, 1361–1366.

Im, A.P., Sehgal, A.R., Carroll, M.P., Smith, B.D., Tefferi, A., Johnson, D.E., and Boyiadzis, M. (2014). DNMT3A and IDH mutations in acute myeloid leukemia and other myeloid malignancies: associations with prognosis and potential treatment strategies. Leukemia *28*, 1774–1783.

Inoue, A., and Zhang, Y. (2011). Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science *334*, 194.

Inoue, A., Shen, L., Dai, Q., He, C., and Zhang, Y. (2011). Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell Res. *21*, 1670–1676.

Inoue, A., Shen, L., Matoba, S., and Zhang, Y. (2015). Haploinsufficiency, but Not Defective Paternal 5mC Oxidation, Accounts for the Developmental Defects of Maternal Tet3 Knockouts. Cell Rep. *10*, 463–470.

Inoue, S., Lemonnier, F., and Mak, T.W. (2016a). Roles of IDH1/2 and TET2 mutations in myeloid disorders. Int. J. Hematol. *103*, 627–633.

Inoue, S., Li, W.Y., Tseng, A., Nolan, G.P., Cairns, R.A., and Mak, T.W. (2016b). Mutant IDH1 Downregulates ATM and Alters DNA Repair and Sensitivity to DNA Damage Independent Mutant IDH1 Downregulates ATM and Alters DNA Repair and Sensitivity to DNA Damage Independent of TET2. Cancer Cell 1–12.

Iqbal, K., Jin, S.-G., Pfeifer, G.P., and Szabó, P.E. (2011). Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. Proc. Natl. Acad. Sci. U. S. A. *108*, 3642–3647.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet *41*, 178–186.

Ito, S., Alessio, A.C.D., Taranova, O. V, Hong, K., and Lawrence, C. (2010). Role of Tet proteins in 5mC to 5hmC conversion , ES cell self- renewal , and ICM specification. Nature *466*, 1129–1133.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Carolina, N., Hill, C., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300–1303.

Iurlaro, M., Ficz, G., Oxley, D., Raiber, E.-A., Bachman, M., Booth, M.J., Andrews, S., Balasubramanian, S., and Reik, W. (2013). A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. Genome Biol. *14*, R119.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol. Cell. Biol. *24*, 8862–8871.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., et al. (2014). Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. N. Engl. J. Med. *371*, 2488–2498.

Jeong, M., Sun, D., Luo, M., Huang, Y., Challen, G. a, Rodriguez, B., Zhang, X., Chavez, L., Wang, H., Hannah, R., et al. (2013). Large conserved domains of low DNA methylation maintained by Dnmt3a. Nat. Genet. *46*, 17–23.

Ji, H., Ehrlich, L.I.R., Seita, J., Murakami, P., Doi, A., Lindau, P., Lee, H., Aryee, M.J., Irizarry, R.A., Kim, K., et al. (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature *467*, 338–342.

Jin, C., Lu, Y., Jelinek, J., Liang, S., Estecio, M.R.H., Barton, M.C., and Issa, J.-P.J. (2014). TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. Nucleic Acids Res. *42*, 6956–6971.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. *13*, 484–492.

Jones, P.A., and Baylin, S.B. (2007). The epigenomics of cancer. Cell *128*, 683–692.

Kafer, G.R., Li, X., Horii, T., Suetake, I., Tajima, S., Hatada, I., Carlton, P.M., Kafer, G.R., Li, X., Horii, T., et al. (2016). 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability. CellReports 1–10.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., and Saitou, M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. EMBO J. *32*, 340–353.

Kallin, E., Rodr?guez-Ubreva, J., Christensen, J., Cimmino, L., Aifantis, I., Helin, K., Ballestar, E., and Graf, T. (2012). Tet2 Facilitates the Derepression of Myeloid Target Genes during CEBP?-Induced Transdifferentiation of Pre-B Cells. Mol. Cell *48*, 266–276.

Kane, M.F., Loda, M., Gaida, G.M., Lipman, J., Mishra, R., Goldman, H., Jessup, J.M., and Kolodner, R. (1997). Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res. *57*, 808–811.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900–903.

Kats, L.M., Reschke, M., Taulli, R., Pozdnyakova, O., Burgess, K., Bhargava, P., Straley, K., Karnik, R., Meissner, A., Small, D., et al. (2014). Proto-oncogenic role of mutant IDH2 in leukemia initiation and maintenance. Cell Stem Cell *14*, 329–341.

Kelly, A.D., Kroeger, H., Yamazaki, J., Taby, R., Neumann, F., Yu, S., Lee, J.T., Patel, B., Li, Y., He, R., et al. (2017). A CpG island methylator phenotype in acute myeloid leukemia independent of IDH mutations and associated with a favorable outcome. Leukemia.

Kernytsky, A., Wang, F., Hansen, E., Schalm, S., Straley, K., Gliser, C., Yang, H., Travins, J., Murray, S., Dorsch, M., et al. (2014). IDH2 mutation induced histone and DNA hypermethylation is progressively reversed by small molecule inhibition. Blood *125*, 296–304.

Kishikawa, S., Murata, T., Ugai, H., Yamazaki, T., and Yokoyama, K.K. (2003). Control elements of Dnmt1 gene are regulated in cell-cycle dependent manner. Nucleic Acids Res. Suppl. 307–308.

Klug, M., Schmidhofer, S., Gebhard, C., Andreesen, R., and Rehli, M. (2013). 5-Hydroxymethylcytosine is an essential intermediate of active DNA demethylation processes in primary human monocytes. Genome Biol. *14*, R46.

Ko, M., Huang, Y., Jankowska, A.M., Pape, U.J., Tahiliani, M., Bandukwala, H.S., An, J., Lamperti, E.D., Koh, K.P., Ganetzky, R., et al. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature *468*, 839–843.

Ko, M., Bandukwala, H.S., An, J., Lamperti, E.D., Thompson, E.C., and Hastie, R. (2011). Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *2*.

Ko, M., An, J., Bandukwala, H.S., Chavez, L., Aijö, T., Pastor, W.A., Segal, M.F., Li, H., Koh, K.P., Lähdesmäki, H., et al. (2013). Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. Nature *497*, 122–126.

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C. a, Mostoslavsky, G., et al. (2011). Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell *8*, 200–213.

Koivunen, P., Lee, S., Duncan, C.G., Lopez, G., Lu, G., Ramkissoon, S., Losman, J.A., Joensuu, P., Bergmann, U., Gross, S., et al. (2012). Transformation by the (R)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. Nature *483*, 484–488.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929–930.

Kulis, M., Heath, S., Bibikova, M., Queirós, A.C., Navarro, A., Clot, G., Martínez-Trillos, A., Castellano, G., Brun-Heath, I., Pinyol, M., et al. (2012). Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat. Genet. *44*, 1236–1242.

Kulis, M., Merkel, A., Heath, S., Queirós, A.C., Schuyler, R.P., Castellano, G., Beekman, R., Raineri, E., Esteve, A., Clot, G., et al. (2015). Whole-genome fingerprint of the DNA methylome during human B cell differentiation. Nat. Genet. *47*, 746–756.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. Genes Dev. *22*, 1617–1635.

de la Rica, L., Rodr?guez-Ubreva, J., Garc?a, M., Islam, A.B., Urquiza, J.M., Hernando, H., Christensen, J., Helin, K., G?mez-Vaquero, C., and Ballestar, E. (2013). PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. Genome Biol. *14*, R99.

Ladstätter, S., and Tachibana-Konwalski, K. (2016). A Surveillance Mechanism Ensures Repair of DNA Lesions during Zygotic Reprogramming. Cell *167*, 1774–1787.e13.

Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. Genomics *13*, 1095–1107.

Lei, H., Oh, S.P.P., Okano, M., Juttermann, R., Goss, K.A.A., Jaenisch, R., Li, E., Jüttermann, R., Goss, K.A.A., Jaenisch, R., et al. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195–3205.

Ley, T., Ding, L., Walter, M.J., Mclellan, M.D., Lamprecht, T., Larson, D.E., Kandoth, C., Payton, J.E., Baty, J., Welch, J., et al. (2010). DNMT3A Mutations in Acute Myeloid Leukemia. N. Engl. J. Med. *363*, 242–33.

Li, E., and Zhang, Y.Y. (2014). DNA methylation in mammals. Cold Spring Harb Perspect Biol *6*.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.

Li, F., He, X.X., Ye, D., Lin, Y., Yu, H., Yao, C., Huang, L., Zhang, J., Wang, F., Xu, S., et al. (2015). NADP+-IDH Mutations Promote Hypersuccinylation that Impairs Mitochondria Respiration and Induces Apoptosis Resistance. Mol. Cell *60*, 661–675.

Li, T., Yang, D., Li, J., Tang, Y., Yang, J., and Le, W. (2014). Critical Role of Tet3 in Neural Progenitor Cell Maintenance and Terminal Differentiation. Mol. Neurobiol.

Li, Z., Cai, X., Cai, C.-L., Wang, J., Zhang, W., Petersen, B.E., Yang, F.-C., and Xu, M. (2011). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood *118*, 4509–4518.

Lian, C.G., Xu, Y., Ceol, C., Wu, F., Larson, A., Dresser, K., Xu, W., Tan, L., Hu, Y., Zhan, Q., et al. (2012). Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. Cell *150*, 1135–1146.

Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C. a, Donaghey, J., Galonska, C., et al. (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat. Genet. *47*, 469–478.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., et al. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature *471*, 68–73.

Losman, J.-A., Looper, R.E., Koivunen, P., Lee, S., Schneider, R.K., McMahon, C., Cowley, G.S., Root, D.E., Ebert, B.L., and Kaelin, W.G. (2013). (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. Science *339*, 1621–1625.

Lu, C., Ward, P.S.P., Kapoor, G.S.G., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C.R., Khanin, R., Figueroa, M.E., Melnick, A., et al. (2012). IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature *483*, 474–478.

Lu, C., Venneti, S., Akalin, A., Fang, F., Ward, P.S., Dematteo, R.G., Intlekofer, A.M., Chen, C., Ye, J., Hameed, M., et al. (2013). Induction of sarcomas by mutant IDH2. Genes Dev. *27*, 1986–1998.

Lu, F., Liu, Y., Jiang, L., Yamaguchi, S., and Zhang, Y. (2014). Role of Tet proteins in enhancer activity and telomere elongation. Genes Dev.

MacDonald, J.W. (2008). Affycoretools: Functions useful for those doing repetitive analyses with Affymetrix GeneChips.

Mack, S.C., Witt, H., Piro, R.M., Gu, L., Zuyderduyn, S., Stütz, A.M., Wang, X., Gallo, M., Garzia, L., Zayne, K., et al. (2014). Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. Nature *506*, 445–450.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. J. Biol. Chem. *286*, 35334–35338.

Maksimovic, J., Gordon, L., and Oshlack, A. (2012). SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. Genome Biol. *13*, R44.

Mancini, D.N., Singh, S.M., Archer, T.K., and Rodenhiser, D.I. (1999). Site-specific DNA methylation in the neurofibromatosis (NF1) promoter interferes with binding of CREB and SP1 transcription factors. Oncogene *18*, 4108–4119.

Marina, R.J., Sturgill, D., Bailly, M.A., Thenoz, M., Varma, G., Prigge, M.F., Nanan, K.K., Shukla, S., Haque –, N., Oberdoerffer, S., et al. (2016). TET-catalyzed oxidation of intragenic 5-methylcytosine regulates CTCF-dependent alternative splicing. EMBO J. *35*, 335–355.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature *466*, 253–257.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. Nature *403*, 501–502.

Medeiros, B.C., Fathi, A.T., DiNardo, C.D., Pollyea, D.A., Chan, S.M., and Swords, R. (2017). Isocitrate dehydrogenase mutations in myeloid malignancies. Leukemia *31*, 272–281.

Meissner, A., Gnirke, A., Bell, G.W., Ramsahoye, B., Lander, E.S., and Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res. *33*, 5868–5877.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766–770.

Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012). MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell *151*, 1417–1430.

Merlo, A., Herman, J.G., Mao, L., Lee, D.J., Gabrielson, E., Burger, P.C., Baylin, S.B., and Sidransky, D. (1995). 5′ CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. Nat. Med. *1*, 686–692.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M.E., Vasanthakumar, A., Patel, J., Zhao, X., et al. (2011). Tet2 Loss Leads to Increased Hematopoietic Stem Cell Self-Renewal and Myeloid Transformation. Cancer Cell *20*, 11–24.

Nakamura, R., Tsukahara, T., Qu, W., Ichikawa, K., Otsuka, T., Ogoshi, K., Saito, T.L., Matsushima, K., Sugano, S., Hashimoto, S., et al. (2014). Large hypomethylated domains serve as strong repressive machinery for key developmental genes in vertebrates. Development *141*, 2568–2580.

Nakayama, M., Wada, M., Harada, T., Nagayama, J., Kusaba, H., Ohshima, K., Kozuru, M., Komatsu, H., Ueda, R., and Kuwano, M. (1998). Hypomethylation status of CpG sites at the promoter region and overexpression of the human MDR1 gene in acute myeloid leukemias. Blood *92*, 4296–4307.

Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *347*, 1017–1022.

Neri, F., Incarnato, D., Krepelova, A., Rapelli, S., Pagnani, A., Zecchina, R., Parlato, C., and

Oliviero, S. (2013). Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. Genome Biol. *14*, R91.

Neri, F., Incarnato, D., Krepelova, A., Dettori, D., Rapelli, S., Maldotti, M., Parlato, C., Anselmi, F., Galvagni, F., and Oliviero, S. (2015). TET1 is controlled by pluripotency-associated factors in ESCs and downmodulated by PRC2 in differentiated cells and tissues. Nucleic Acids Res. *43*.

Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., and Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. Nature *543*, 72–77.

Nestor, C.E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., Katz, E., Dixon, J.M., Harrison, D.J., and Meehan, R.R. (2012). Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. Genome Res. *22*, 467–477.

Neuwirth, E. (2014). ColorBrewer Palettes [R package RColorBrewer version 1.1-2].

Nibourel, O., Kosmider, O., Cheok, M., Boissel, N., Renneville, A., Philippe, N., Dombret, H., Dreyfus, F., Quesnel, B., Geffroy, S., et al. (2010). Incidence and prognostic value of TET2 alterations in de novo acute myeloid leukemia achieving complete remission. Blood *116*, 1132–1135.

Noushmehr, H., Weisenberger, D.J., Diefes, K., Phillips, H.S., Pujara, K., Berman, B.P., Pan, F., Pelloski, C.E., Sulman, E.P., Bhat, K.P., et al. (2010). Identification of a CpG Island Methylator Phenotype that Defines a Distinct Subgroup of Glioma. Cancer Cell *17*, 510–522.

Oakes, C.C., Seifert, M., Assenov, Y., Gu, L., Przekopowitz, M., Ruppert, A.S., Wang, Q., Imbusch, C.D., Serva, A., Brocks, D., et al. (2016). DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. Nat. Genet. *48*, 253–264.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

Okashita, N., Kumaki, Y., Ebi, K., Nishi, M., Okamoto, Y., Nakayama, M., Hashimoto, S., Nakamura, T., Sugasawa, K., Kojima, N., et al. (2014). PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. Development *141*, 269–280.

Omatsu, Y., Seike, M., Sugiyama, T., Kume, T., and Nagasawa, T. (2014). Foxc1 is a critical regulator of haematopoietic stem/progenitor cell niche formation. Nature *508*, 536–540.

Ong, C.-T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nat. Rev. Genet. *15*, 234–246.

Ono, R., Taki, T., and Taketani, T. (2002). LCX , Leukemia-associated Protein with a CXXC Domain , Is Fused to MLL in Acute Myeloid Leukemia with Trilineage Dysplasia Having t (10 ; 11)(q22 ; q23) LCX , Leukemia-associated Protein with a CXXC Domain , Is Fused to MLL in Acute Myeloid Leukemia w. 4075–4080.

Ooi, S.K.T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C.D., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714–717.

Orlanski, S., Labi, V., Reizel, Y., Spiro, A., Lichtenstein, M., Levin-Klein, R., Koralov, S.B., Skversky, Y., Rajewsky, K., Cedar, H., et al. (2016). Tissue-specific DNA demethylation is

required for proper B-cell differentiation and function. Proc. Natl. Acad. Sci. U. S. A. *113*, 5018–5023.

Oshimo, Y., Nakayama, H., Ito, R., Kitadai, Y., Yoshida, K., Chayama, K., and Yasui, W. (2003). Promoter methylation of cyclin D2 gene in gastric carcinoma. Int. J. Oncol. *23*, 1663–1670.

Pansuriya, T.C., van Eijk, R., d'Adamo, P., van Ruler, M.A.J.H., Kuijjer, M.L., Oosting, J., Cleton-Jansen, A.-M., van Oosterwijk, J.G., Verbeke, S.L.J., Meijer, D., et al. (2011). Somatic mosaic IDH1 and IDH2 mutations are associated with enchondroma and spindle cell hemangioma in Ollier disease and Maffucci syndrome. Nat. Genet. *43*, 1256–1261.

Passegué, E., Jamieson, C.H.M., Ailles, L.E., and Weissman, I.L. (2003). Normal and leukemic hematopoiesis: are leukemias a stem cell disorder or a reacquisition of stem cell characteristics? Proc. Natl. Acad. Sci. U. S. A. *100 Suppl 1*, 11842–11849.

Pastor, W. a, Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., et al. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature *473*, 394–397.

Patel, J.P., G?nen, M., Figueroa, M.E., Fernandez, H., Sun, Z., Racevskis, J., Van Vlierberghe, P., Dolgalev, I., Thomas, S., Aminova, O., et al. (2012). Prognostic Relevance of Integrated Genetic Profiling in Acute Myeloid Leukemia. N. Engl. J. Med. *366*, 1079–1089.

Penn, N.W., Suwalski, R., O'Riley, C., Bojanowski, K., Yura, R., Penn, B.N.W., Suwalski, R., Riley, C.O., Bojanowski, K., Yurat, R., et al. (1972). The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. Biochem. J. *126*, 781–790.

Perera, A., Eisen, D., Wagner, M., Laube, S.K., Künzel, A.F., Koch, S., Steinbacher, J., Schulze, E., Splith, V., Mittermeier, N., et al. (2015). TET3 is recruited by REST for context-specific hydroxymethylation and induction of gene expression. Cell Rep. *11*, 283–294.

Pérez, C., Martínez-Calle, N., Martín-Subero, J.I., Segura, V., Delabesse, E., Fernandez-Mercado, M., Garate, L., Alvarez, S., Rifon, J., Varea, S., et al. (2012). TET2 Mutations Are Associated with Specific 5-Methylcytosine and 5-Hydroxymethylcytosine Profiles in Patients with Chronic Myelomonocytic Leukemia. PLoS One *7*, e31605.

Prendergast, G.C., and Ziff, E.B. (1991). Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. Science *251*, 186–189.

Pusch, S., Krausert, S., Fischer, V., Balss, J., Ott, M., Schrimpf, D., Capper, D., Sahm, F., Eisel, J., Beck, A.-C., et al. (2017). Pan-mutant IDH1 inhibitor BAY 1436032 for effective treatment of IDH1 mutant astrocytoma in vivo. Acta Neuropathol. *133*, 629–644.

Qu, G.Z., Grundy, P.E., Narayan, A., and Ehrlich, M. (1999). Frequent hypomethylation in Wilms tumors of pericentromeric DNA in chromosomes 1 and 16. Cancer Genet. Cytogenet. *109*, 34–39.

Quivoron, C., Couronné, L., Della Valle, V., Lopez, C.K., Plo, I., Wagner-Ballon, O., Do Cruzeiro, M., Delhommeau, F., Arnulf, B., Stern, M.-H., et al. (2011). TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell *20*, 25–38.

Raddatz, G., Gao, Q., Bender, S., Jaenisch, R., and Lyko, F. (2012). Dnmt3a Protects Active Chromosome Domains against Cancer-Associated Hypomethylation. PLoS Genet. *8*, e1003146.

Rampal, R., Alkalin, A., Madzo, J., Vasanthakumar, A., Pronier, E., Patel, J., Li, Y., Ahn, J.,

Abdel-Wahab, O., Shih, A., et al. (2014). DNA Hydroxymethylation Profiling Reveals that WT1 Mutations Result in Loss of TET2 Function in Acute Myeloid Leukemia. Cell Rep. *9*.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc. Natl. Acad. Sci. U. S. A. *97*, 5237–5242.

Rasmussen, K.D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer (Cold Spring Harbor Laboratory Press).

Rasmussen, K.D., Jia, G., Johansen, J. V., Pedersen, M.T., Rapin, N., Bagger, F.O., Porse, B.T., Bernard, O.A., Christensen, J., and Helin, K. (2015). Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes Dev. *29*, 910–922.

Rastegar, F., Shenaq, D., Huang, J., Zhang, W., Zhang, B.-Q., He, B.-C., Chen, L., Zuo, G.-W., Luo, Q., Shi, Q., et al. (2010). Mesenchymal stem cells: Molecular characteristics and clinical applications. World J. Stem Cells *2*, 67–80.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*, 425–432.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089–1093.

Rohle, D., Popovici-Muller, J., Palaskas, N., Turcan, S., Grommes, C., Campos, C., Tsoi, J., Clark, O., Oldrini, B., Komisopoulou, E., et al. (2013). An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. Science *340*, 626–630.

Rondelet, G., Dal Maso, T., Willems, L., and Wouters, J. (2016). Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. J. Struct. Biol. *194*, 357–367.

Rönnerblad, M., Andersson, R., Olofsson, T., Douagi, I., Karimi, M., Lehmann, S., Hoof, I., De Hoon, M., Itoh, M., Nagao-Sato, S., et al. (2014). Analysis of the DNA methylome and transcriptome in granulopoiesis reveals timed changes and dynamic enhancer methylation. Blood *123*, 79–90.

Rosen, E.D., and MacDougald, O. a (2006). Adipocyte differentiation from the inside out. Nat. Rev. Mol. Cell Biol. *7*, 885–896.

Rosen, E.D., Sarraf, P., Troy, A.E., Bradwin, G., Moore, K., Milstone, D.S., Spiegelman, B.M., and Mortensen, R.M. (1999). PPAR gamma is required for the differentiation of adipose tissue in vivo and in vitro. Mol. Cell *4*, 611–617.

Russler-Germain, D.A., Spencer, D.H., Young, M.A., Lamprecht, T.L., Miller, C.A., Fulton, R., Meyer, M.R., Erdmann-Gilmore, P., Townsend, R.R., Wilson, R.K., et al. (2014). The R882H DNMT3A Mutation Associated with AML Dominantly Inhibits Wild-Type DNMT3A by Blocking Its Ability to Form Active Tetramers. Cancer Cell *25*, 442–454.

Russo, V.E.A., Martienssen, R.A., and Riggs, A.D. (1996). Epigenetic Mechanisms of Gene Regulation (Woodbury: Cold Spring Harbor Laboratory Press).

Sakai, T., Toguchida, J., Ohtani, N., Yandell, D.W., Rapaport, J.M., and Dryja, T.P. (1991). Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. Am. J. Hum. Genet. *48*, 880–888.

Sakaki, M., Ebihara, Y., Okamura, K., Nakabayashi, K., Igarashi, A., Matsumoto, K., Hata, K.,

Kobayashi, Y., and Maehara, K. (2017). Potential roles of DNA methylation in the initiation and establishment of replicative senescence revealed by array-based methylome and transcriptome analyses. PLoS One *12*, e0171431.

Sandoval, J., Heyn, H., Moran, S., Serra-Musach, J., Pujana, M.A., Bibikova, M., and Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics *6*, 692–702.

Sasaki, M., Knobbe, C.B., Itsumi, M., Elia, A.J., Harris, I.S., Chio, I.I.C., Cairns, R.A., McCracken, S., Wakeham, A., Haight, J., et al. (2012a). D-2-hydroxyglutarate produced by mutant IDH1 perturbs collagen maturation and basement membrane function. Genes Dev. *26*, 2038–2049.

Sasaki, M., Knobbe, C.B., Munger, J.C., Lind, E.F., Brenner, D., Brüstle, A., Harris, I.S., Holmes, R., Wakeham, A., Haight, J., et al. (2012b). IDH1(R132H) mutation increases murine haematopoietic progenitors and alters epigenetics. Nature *488*, 656–659.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. U. S. A. *103*, 1412–1417.

Scopim-Ribeiro, R., Machado-Neto, J.A., Campos, P. de M., Silva, C.A.M., Favaro, P., Lorand-Metze, I., Costa, F.F., Saad, S.T.O., and Traina, F. (2015). Ten-Eleven-Translocation 2 (TET2) is downregulated in myelodysplastic syndromes. Eur. J. Haematol. *94*, 413–418.

Scourzic, L., Mouly, E., Bernard, O.A., Zhang, S., McPherson, J., and Tao, J. (2015). TET proteins and the control of cytosine demethylation in cancer. Genome Med. *7*, 9.

Scourzic, L., Couronné, L., Pedersen, M.T., Della Valle, V., Diop, M., Mylonas, E., Calvo, J., Mouly, E., Lopez, C.K., Martin, N., et al. (2016). DNMT3AR882H mutant and Tet2 inactivation cooperate in the deregulation of DNA methylation control to induce lymphoid malignancies in mice. Leukemia *30*, 1388–1398.

Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A.C., Fung, H.-L., Zhang, K., and Zhang, Y. (2013). Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell *153*, 692–706.

Shen, L., Inoue, A., He, J., Liu, Y., Lu, F., and Zhang, Y. (2014a). Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. Cell Stem Cell *15*, 459–470.

Shen, L., Song, C.-X., He, C., and Zhang, Y. (2014b). Mechanism and Function of Oxidative Reversal of DNA and RNA Methylation. Annu. Rev. Biochem. *83*, 585–614.

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

Shih, A.H., Meydan, C., Shank, K., Garrett-Bakelman, F.E., Ward, P.S., Intlekofer, A., Nazir, A., Stein, E., Knapp, K., Glass, J., et al. (2017). Combination Targeted Therapy to Disrupt Aberrant Oncogenic Signaling and Reverse Epigenetic Dysfunction in IDH2- and TET2-Mutant Acute Myeloid Leukemia. Cancer Discov. CD-16-1049.

Shlush, L.I., Zandi, S., Mitchell, A., Chen, W.C., Brandwein, J.M., Gupta, V., Kennedy, J.A., Schimmer, A.D., Schuh, A.C., Yee, K.W., et al. (2014). Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature *506*, 328–333.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer,

P., Sandberg, R., and Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature *479*, 74–79.

De Smet, C., De Backer, O., Faraoni, I., Lurquin, C., Brasseur, F., and Boon, T. (1996). The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. Proc. Natl. Acad. Sci. U. S. A. *93*, 7149–7153.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. *14*.

Spencer, D.H., Russler-Germain, D.A., Ketkar, S., Helton, N.M., Lamprecht, T.L., Fulton, R.S., Fronick, C.C., O'Laughlin, M., Heath, S.E., Shinawi, M., et al. (2017). CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence of AML Progression. Cell *168*, 801–816.e13.

Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W.T.C., Bauer, C., Münzel, M., Wagner, M., Müller, M., Khan, F., et al. (2013). Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. Cell *152*, 1146–1159.

Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature *480*, 490–495.

Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., and Jacobsen, S.E. (2011). 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biol. *12*, R54.

Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D.-A., Jones, D.T.W., Konermann, C., Pfaff, E., Tönjes, M., Sill, M., Bender, S., et al. (2012). Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell *22*, 425–437.

Sun, D., Xi, Y., Rodriguez, B., Park, H.J., Tong, P., Meong, M., Goodell, M.A., and Li, W. (2014). MOABS: model based analysis of bisulfite sequencing data. Genome Biol. *15*, R38.

Sun, Z., Dai, N., Borgaro, J.G., Quimby, A., Sun, D., Corrêa, I.R., Zheng, Y., Zhu, Z., and Guan, S. (2015). A sensitive approach to map genome-wide 5-hydroxymethylcytosine and 5-formylcytosine at single-base resolution. Mol. Cell *57*, 750–761.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930–935.

Tamanaha, E., Guan, S., Marks, K., and Saleh, L. (2016). 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. J. Am. Chem. Soc. *138*, 9345–9348.

Tang, Q.Q., and Lane, M.D. (2012). Adipogenesis: From Stem Cell to Adipocyte. Annu. Rev. Biochem. *81*, 715–736.

Tefferi, A., Lim, K.-H., Abdel-Wahab, O., Lasho, T.L., Patel, J., Patnaik, M.M., Hanson, C.A., Pardanani, A., Gilliland, D.G., and Levine, R.L. (2009). Detection of mutant TET2 in myeloid malignancies other than myeloproliferative neoplasms: CMML, MDS, MDS/MPN and AML. Leukemia *23*, 1343–1345.

Terragni, J., Bitinaite, J., Zheng, Y., and Pradhan, S. (2012). Biochemical Characterization of Recombinant β-Glucosyltransferase and Analysis of Global 5-Hydroxymethylcytosine in Unique Genomes. Biochemistry *51*, 1009–1019.

Thienpont, B., Steinbacher, J., Zhao, H., D'Anna, F., Kuchnio, A., Ploumakis, A., Ghesquière, B., Van Dyck, L., Boeckx, B., Schoonjans, L., et al. (2016). Tumour hypoxia causes DNA hypermethylation by reducing TET activity. Nature *537*, 1–25.

Thiery, J.P., Acloque, H., Huang, R.Y.J., Nieto, M.A., Han, J., Niswander, L., Teixeira, J., Donahoe, P.K., Pu, W.T., Roberts, A.B., et al. (2009). Epithelial-Mesenchymal Transitions in Development and Disease. Cell *139*, 871–890.

Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J.G., Baylin, S.B., and Issa, J.P. (1999). CpG island methylator phenotype in colorectal cancer. Proc. Natl. Acad. Sci. U. S. A. *96*, 8681–8686.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2012). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat. Biotechnol. *31*, 46–53.

Trowbridge, J.J., Snow, J.W., Kim, J., and Orkin, S.H. (2009). DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. Cell Stem Cell *5*, 442–449.

Tsagaratou, A., González-Avalos, E., Rautio, S., Scott-Browne, J.P., Togher, S., Pastor, W.A., Rothenberg, E. V, Chavez, L., Lähdesmäki, H., and Rao, A. (2017). TET proteins regulate the lineage specification and TCR-mediated expansion of iNKT cells. Nat. Immunol. *18*, 45–53.

Tuorto, F., Liebers, R., Musch, T., Schaefer, M., Hofmann, S., Kellner, S., Frye, M., Helm, M., Stoecklin, G., and Lyko, F. (2012). RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat. Struct. Mol. Biol. *19*, 900–905.

Turcan, S., Rohle, D., Goenka, A., Walsh, L.A., Fang, F., Yilmaz, E., Campos, C., Fabius, A.W.M., Lu, C., Ward, P.S., et al. (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature *483*, 479–483.

Uribe-Lewis, S., Stark, R., Carroll, T., Dunning, M.J., Bachman, M., Ito, Y., Stojic, L., Halim, S., Vowler, S.L., Lynch, A.G., et al. (2015). 5-hydroxymethylcytosine marks promoters in colon that resist DNA hypermethylation in cancer. Genome Biol. *16*, 69.

Vardimon, L., Kressmann, A., Cedar, H., Maechler, M., and Doerfler, W. (1982). Expression of a cloned adenovirus gene is inhibited by in vitro methylation. Proc. Natl. Acad. Sci. U. S. A. *79*, 1073–1077.

Verhaak, R.G.W., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P., et al. (2010). Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98–110.

Visvader, J.E. (2011). Cells of origin in cancer. Nature *469*, 314–322.

Voigt, P., and Reinberg, D. (2013). Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia The Cancer Genome Atlas Research Network. N. Engl. J. Med. *368*, 2059–2074.

Waddington, C.H. (1942). The epigenotype. Endeavor *1*, 18–20.

Wang, F., Travins, J., Delabarre, B., Penard-lacronique, V., Schalm, S., Hansen, E., Straley,

K., Kernytsky, A., Liu, W., Gliser, C., et al. (2013a). Targeted inhibition of mutant IDH2 in leukemia cells induces cellular differentiation. 622–626.

Wang, L., Zhou, Y., Xu, L., Xiao, R., Lu, X., Chen, L., Chong, J., Li, H., He, C., Fu, X.-D., et al. (2015a). Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. Nature *523*, 621–625.

Wang, P., Dong, Q., Zhang, C., Kuan, P.-F., Liu, Y., Jeck, W.R., Andersen, J.B., Jiang, W., Savich, G.L., Tan, T.-X., et al. (2013b). Mutations in isocitrate dehydrogenase 1 and 2 occur frequently in intrahepatic cholangiocarcinomas and share hypermethylation targets with glioblastomas. Oncogene *32*, 3091–3100.

Wang, Y., Xiao, M., Chen, X., Chen, L., Xu, Y., Lv, L., Wang, P., Yang, H., Ma, S., Lin, H., et al. (2015b). WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. Mol. Cell *57*, 662–673.

Ward, P.S., Patel, J., Wise, D.R., Abdel-Wahab, O., Bennett, B.D., Coller, H.A., Cross, J.R., Fantin, V.R., Hedvat, C. V., Perl, A.E., et al. (2010). The Common Feature of Leukemia-Associated IDH1 and IDH2 Mutations Is a Neomorphic Enzyme Activity Converting ??-Ketoglutarate to 2-Hydroxyglutarate. Cancer Cell *17*, 225–234.

Ward, P.S., Lu, C., Cross, J.R., Abdel-Wahab, O., Levine, R.L., Schwartz, G.K., and Thompson, C.B. (2013). The potential for isocitrate dehydrogenase mutations to produce 2-hydroxyglutarate depends on allele specificity and subcellular compartmentalization. J. Biol. Chem. *288*, 3804–3815.

Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., et al. (2016). gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1.

Weber, A.R., Krawczyk, C., Robertson, A.B., Kuśnierczyk, A., Vågbø, C.B., Schuermann, D., Klungland, A., and Schär, P. (2016). Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. Nat. Commun. *7*, 10806.

Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M.A., Kang, G.H., Widschwendter, M., Weener, D., Buchanan, D., et al. (2006). CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat. Genet. *38*, 787–793.

Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., et al. (2012). The Origin and Evolution of Mutations in Acute Myeloid Leukemia. Cell *150*, 264–278.

Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D.J., Campan, M., Young, J., Jacobs, I., et al. (2007). Epigenetic stem cell signature in cancer. Nat. Genet. *39*, 157–158.

Wiehle, L., Raddatz, G., Musch, T., Dawlaty, M.M., Jaenisch, R., Lyko, F., and Breiling, A. (2015). Tet1 and Tet2 protect DNA methylation canyons against hypermethylation. Mol. Cell. Biol. *36*, MCB.00587-15.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J. V, Cloos, P. a C., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343–348.

Wilson, A.A., Murphy, G.J., Hamakawa, H., Kwok, L.W., Srinivasan, S., Hovav, A.-H., Mulligan, R.C., Amar, S., Suki, B., and Kotton, D.N. (2010). Amelioration of emphysema in

mice through lentiviral transduction of long-lived pulmonary alveolar macrophages. J. Clin. Invest. *120*, 379–389.

Wossidlo, M., Arand, J., Sebastiano, V., Lepikhov, K., Boiani, M., Reinhardt, R., Schöler, H., and Walter, J. (2010). Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. EMBO J. *29*, 1877–1888.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat. Commun. *2*, 241.

Wu, H., and Zhang, Y. (2014). Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. Cell *156*, 45–68.

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nat. Rev. Genet.

Wu, H., D'Alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., Zhang, Y., and Alessio, A.C.D. (2011a). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev. *25*, 679–684.

Wu, H., D'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011b). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature *473*, 389–393.

Wu, H., Wu, X., Shen, L., and Zhang, Y. (2014). Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. Nat. Biotechnol. *32*, 1231–1240.

Wyatt, G.R., and Cohen, S.S. (1952). A new pyrimidine base from bacteriophage nucleic acids. Nature *170*, 1072–1073.

Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics *10*, 232.

Xiao, M., Yang, H., Xu, W., Ma, S., Lin, H., Zhu, H., Liu, L., Liu, Y., Yang, C., Xu, Y., et al. (2012). Inhibition of -KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. Genes Dev. *26*, 1326–1338.

Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D., Leung, D., et al. (2013). Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell *153*, 1134–1148.

Xiong, J., Zhang, Z., Chen, J., Huang, H., Xu, Y., Ding, X., Zheng, Y., Nishinakamura, R., Xu, G.-L., Wang, H., et al. (2016). Cooperative Action between SALL4A and TET Proteins in Stepwise Oxidation of 5-Methylcytosine. Mol. Cell *64*, 913–925.

Xu, W., Yang, H., Liu, Y., Yang, Y., Wang, P., Kim, S.-H., Ito, S., Yang, C., Wang, P., Xiao, M.-T., et al. (2011a). Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer Cell *19*, 17–30.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., et al. (2011b). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol. Cell *42*, 451–464.

Yamaguchi, S., Hong, K., Liu, R., Shen, L., Inoue, A., Diep, D., Zhang, K., and Zhang, Y.

(2012). Tet1 controls meiosis by regulating meiotic gene expression. Nature *492*, 443–447.

Yamaguchi, S., Shen, L., Liu, Y., Sendler, D., and Zhang, Y. (2013). Role of Tet1 in erasure of genomic imprinting. Nature *504*, 460–464.

Yamashita, K., Dai, T., Dai, Y., Yamamoto, F., and Perucho, M. (2003). Genetics supersedes epigenetics in colon cancer phenotype. Cancer Cell *4*, 121–131.

Yamazaki, J., Taby, R., Vasanthakumar, A., Macrae, T., Ostler, K.R., Shen, L., Kantarjian, H.M., Estecio, M.R., Jelinek, J., Godley, L.A., et al. (2012). Effects of *TET2* mutations on DNA methylation in chronic myelomonocytic leukemia. Epigenetics *7*, 201–207.

Yamazaki, J., Jelinek, J., Lu, Y., Cesaroni, M., Madzo, J., Neumann, F., He, R., Taby, R., Vasanthakumar, A., Macrae, T., et al. (2015). TET2 Mutations Affect Non-CpG Island DNA Methylation at Enhancers and Transcription Factor-Binding Sites in Chronic Myelomonocytic Leukemia. Cancer Res. *75*, 2833–2843.

Yan, H., Wang, Y., Qu, X., Li, J., Hale, J., Huang, Y., An, C., Papoin, J., Guo, X., Chen, L., et al. (2017). Distinct roles for TET family proteins in regulating human erythropoiesis. Blood *129*, 2002–2012.

Yang, H., Lin, H., Xu, H., Zhang, L., Cheng, L., Wen, B., Shou, J., Guan, K., Xiong, Y., and Ye, D. (2014). TET-catalyzed 5-methylcytosine hydroxylation is dynamically regulated by metabolites. Cell Res. *24*, 1017–1020.

Yang, J., Guo, R., Wang, H., Ye, X., Zhou, Z., Dan, J., Wang, H., Gong, P., Deng, W., Yin, Y., et al. (2016a). Tet Enzymes Regulate Telomere Maintenance and Chromosomal Stability of Mouse ESCs. Cell Rep. 1–13.

Yang, Q., Liang, X., Sun, X., Zhang, L., Fu, X., Rogers, C.J., Berim, A., Zhang, S., Wang, S., Wang, B., et al. (2016b). AMPK/α-Ketoglutarate Axis Dynamically Mediates DNA Demethylation in the Prdm16 Promoter and Brown Adipogenesis. Cell Metab. *24*, 542–554.

Yang, R., Qu, C., Zhou, Y., Konkel, J.E., Shi, S., Liu, Y., Chen, C., Liu, S., Liu, D., Chen, Y., et al. (2015). Hydrogen Sulfide Promotes Tet1- and Tet2-Mediated Foxp3 Demethylation to Drive Regulatory T Cell Differentiation and Maintain Immune Homeostasis. Immunity *43*, 251–263.

Yen, K., Travins, J., Wang, F., David, M.D., Artin, E., Straley, K., Padyana, A., Gross, S., DeLaBarre, B., Tobin, E., et al. (2017). AG-221, a First-in-Class Therapy Targeting Acute Myeloid Leukemia Harboring Oncogenic IDH2 Mutations. Cancer Discov. *7*, CD-16-1034.

Yildirim, O., Li, R., Hung, J.-H., Chen, P.B., Dong, X., Ee, L.-S., Weng, Z., Rando, O.J., and Fazzio, T.G. (2011). Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell *147*, 1498–1510.

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science *356*, eaaj2239.

Yoder, J.A., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. *13*, 335–340.

Yoo, Y., Park, J.H., Weigel, C., Liesenfeld, D.B., Weichenhan, D., Plass, C., Seo, D.-G., Lindroth, A.M., and Park, Y.J. (2017). TET-mediated hydroxymethylcytosine at the Pparγ locus is required for initiation of adipogenic differentiation. Int. J. Obes. *41*, 652–659.

Yu, M., Hon, G.C.C., Szulwach, K.E.E., Song, C.-X.X., Zhang, L., Kim, A., Li, X., Dai, Q.,

Shen, Y., Park, B., et al. (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell *149*, 1368–1380.

Yue, X., Trifari, S., Äijö, T., Tsagaratou, A., Pastor, W.A., Zepeda-Martínez, J.A., Lio, C.-W.J., Li, X., Huang, Y., Vijayanand, P., et al. (2016). Control of Foxp3 stability through modulation of TET activity. J. Exp. Med. *213*, 377–397.

Zhang, W., Xia, W., Wang, Q., Towers, A.J., Chen, J., Gao, R., Zhang, Y., Yen, C.-A., Lee, A.Y., Li, Y., et al. (2016a). Isoform Switch of TET1 Regulates DNA Demethylation and Mouse Development. Mol. Cell *64*, 1062–1073.

Zhang, X., Su, J., Jeong, M., Ko, M., Huang, Y., Park, H.J., Guzman, A., Lei, Y., Huang, Y.-H., Rao, A., et al. (2016b). DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. Nat. Genet. *48*, 1014–1023.

Zhang, Y.W., Wang, Z., Xie, W., Cai, Y., Xia, L., Easwaran, H., Luo, J., Yen, R.-W.C., Li, Y., and Baylin, S.B. (2017). Acetylation Enhances TET2 Function in Protecting against Abnormal DNA Methylation during Oxidative Stress. Mol. Cell *65*, 323–335.

Zhao, S., Lin, Y., Xu, W., Jiang, W., Zha, Z., Wang, P., Yu, W., Li, Z., Gong, L., Peng, Y., et al. (2009). Glioma-derived mutations in IDH1 dominantly inhibit IDH1 catalytic activity and induce HIF-1alpha. Science *324*, 261–265.

Zhao, Z., Chen, L., Dawlaty, M.M., Pan, F., Weeks, O., Zhou, Y., Cao, Z., Shi, H., Wang, J., Lin, L., et al. (2015). Combined Loss of Tet1 and Tet2 Promotes B Cell, but Not Myeloid Malignancies, in Mice. Cell Rep. *13*, 1692–1704.

Zhong, X., Wang, Q.-Q., Li, J.-W., Zhang, Y.-M., An, X.-R., and Hou, J. (2017). Ten-Eleven Translocation-2 (Tet2) Is Involved in Myogenic Differentiation of Skeletal Myoblast Cells in Vitro. Sci. Rep. *7*, 43539.

Ziller, M.J., Gu, H., Müller, F., Donaghey, J., Tsai, L.T.-Y., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E., et al. (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*.

5.3 List of Publications

Wiehle, L., Raddatz, G., Musch, T., Dawlaty, M.M., Jaenisch, R., Lyko, F., and Breiling, A. (2016). Tet1 and Tet2 protect DNA methylation canyons against hypermethylation. Mol. Cell. Biol. 36(3), 452-61, doi: 10.1128/MCB.00587-15

Wiehle, L., Breiling, A. (2016). Chromatin Immunoprecipitation. Methods Mol. Biol. 1480, 7-21, doi: 10.1007/978-1-4939-6380-5_2

Wiehle, L., Raddatz, G., Pusch, S., Gutekunst, J., Rodríguez-Paredes, M., Lyko, F. (2017). *MIDH*-associated DNA methylation changes in AML reflect undifferentiated cell states rather than inhibition of TET-mediated demethylation. (manuscript in preparation)

Danksagung

Mein herzlicher Dank geht an Prof. Dr. Frank Lyko für die Ermöglichung und Betreuung meiner Doktorarbeit, die sehr guten Arbeitsbedingungen, die stets offene Tür und die Ratschläge zu wissenschaftlichen und anderen Fragen.

Ich danke Prof. Dr. Jan Lohmann, Prof. Dr. Alwin Krämer und Prof. Dr. Gudrun Rappold: Für die Übernahme des Zweitgutachtens, die Begleitung meiner Projekte im Rahmen meines TAC Komitees sowie die Teilnahme in meinem Prüfungskomitee.

Ein besonderer Dank gilt Achim Breiling für die Betreuung, Unterstützung und das Korrekturlesen, sowie Manuel Rodríguez-Paredes für seinen unerschütterlichen Enthusiasmus und Optimismus.

Des Weiteren gebührt ein herzlicher Dank unseren Kollaboratoren Meelad Dawlaty und Rudolf Jaenisch für die Bereitstellung der MEFs und die Unterstützung des Canyon-Projektes.

Stefan Pusch und Jessica Eisel danke ich für die hervorragende Zusammenarbeit im Rahmen des IDH-Projektes, insbesondere die 2-HG-Messungen, die Plasmide, die hilfreichen Ratschläge, Diskussionen und fachliches Know-how.

Ich möchte der DKFZ Genomics Core Facility danken für die Unterstützung durch Illumina-Sequenzierungen und Arrays, sowie die ständige Hilfsbereitschaft.

Ein Riesen-Dankeschön geht an alle derzeitigen und ehemaligen „Epigeneticists“ des Lyko-Labors: Aracely, Bojana, Carine, Cassy, Dagmar, Fanny, Felix, Flo, Francesca, Günter, Imola, Jana, Jessica, Johanna, Julian, Kathi, Lena, Maria, Mark, Matthias M., Matthias S., Nader, Ranja, Reinhard, Sebastian, Sim, Sven, Tanja, Qianchao und Zeljko für tatkräftige Hilfe und Unterstützung im Labor, in der Bioinformatik und in organisatorischen Belangen, für wissenschaftlichen Input, Anregungen und Diskussionen, für das Korrekturlesen dieser Arbeit, für die kollegiale Atmosphäre, Freundschaft sowie für gute Gespräche, gemütliche Feierabend-Bierchen, Parties und amüsante Kaffeepausen!

Bedanken möchte ich außerdem bei den vielen anderen Menschen, die meine PhD-Zeit begleitet haben: Lisa für ihr immer offenes Ohr, den Marburger und Heidelberger Kommilitonen für die herrliche Zeit, den Auftaktlern für gemeinsame Erlebnisse und tolle Musik, Phivos und den Wildtypes für spannende Konzerte und Improvisationen, den Biocontactlern für interessante Events, Erich und Arno für die grandiosen Doppelkopfabende, Elias, Sabrina, Linda, Steffen, Eva und Till für die Steigerung der Work-Life-Balance!

Zu guter Letzt gilt mein aufrichtiger Dank meiner Familie, meinen Eltern und besonders Stephen für die Liebe, Wärme, geduldige Unterstützung und dafür, dass sie immer für mich da sind.