

Universität Stuttgart

**Application and engineering of targeted DNA methylation editing
tools for the modulation and characterization of the epigenome net-
work**

Von der Fakultät 4: Energie-, Verfahrens- und Biotechnik der Universität Stuttgart zur
Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigte Abhandlung

Vorgelegt von

Julian Broche

Geboren am 20.02.1988 in Freiburg im Breisgau

Hauptberichter: Prof. Dr. Albert Jeltsch

Mitberichter: Prof. Dr. Stefan Legewie

Prüfungsvorsitzender: Prof. Dr. Stephan Nussberger

Tag der mündlichen Prüfung: 27.01.2022

Institut für Biochemie und Technische Biochemie der Universität Stuttgart

2022

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel

„Application and engineering of targeted DNA methylation editing tools for the modulation and characterization of the epigenome network“

selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Aus fremden Quellen entnommene Passagen und Gedanken sind als solche kenntlich gemacht.

Stuttgart, 19.08.2021

Julian Broche

Table of contents

Acknowledgements	V
List of publications	VI
Zusammenfassung	VII
Abstract	X
List of abbreviations	XII
1. Introduction.....	1
1.1 Epigenetics and the structure of chromatin	1
1.2 DNA methylation	4
1.2.1 Occurrence of DNA methylation in different organisms	4
1.2.2 CpG dinucleotides and CpG islands.....	5
1.2.3 Regulatory functions of DNA methylation	6
1.2.4 DNA methylation in development	9
1.3 Mechanisms of DNA maintenance methylation.....	13
1.4 Mammalian DNA methyltransferases.....	14
1.4.1 DNMT1	16
1.4.2 DNMT3 family.....	18
1.5 Active demethylation by TET family members	22
1.6 Post-translational modifications at histone tails and their connection to DNA methylation.....	25
1.6.1 Histone 3 lysine 4 trimethylation	27
1.6.2 Histone 3 lysine 9 trimethylation	29
1.6.3 Histone 3 lysine 27 acetylation	32
1.6.4 Histone 3 lysine 27 trimethylation	34
1.6.5 Histone 3 lysine 36 trimethylation	38
1.7 Targeted epigenome editing	40

2. Principal aims of the study	46
3. Materials and methods	48
3.1 Molecular cloning	48
3.2 Cell culture experiments.....	50
3.2.1 Cell maintenance.....	50
3.2.2 Transient transfection	50
3.2.3 Generation of stable cell lines	51
3.2.4 Cell culture experiments for genome-wide methylation studies.....	54
3.2.5 Knockdown of TET enzymes in DNA methylation studies	54
3.2.6 Application of enzyme inhibitors in cell culture	55
3.3 SDS PAGE and western blot	56
3.4 Nucleic acid analysis.....	57
3.4.1 MBD2-pulldown	57
3.4.2 Cross-linked ChIP (XChIP).....	58
3.4.3 Native ChIP (NChIP)	60
3.4.4 Quantitative PCR (qPCR).....	61
3.4.5 Reverse transcription quantitative PCR (RT-qPCR).....	64
3.4.6 Library preparation of MBD2-pulldown, NChIP and XChIP samples for next-generation sequencing	66
3.4.7 Library preparation for RNA-seq.....	66
3.4.8 Bisulfite sequencing.....	67
4. Results.....	68
4.1 Dynamics of the epigenome network after global DNA methylation editing	68
4.1.1 Analysis of the ZnF binding properties	70
4.1.2 Genome-wide methylation dynamics after induced ZnF-3AC overexpression	72
4.1.3 Knockdown and inhibition of TET enzymes.....	78

4.1.4 Implications of global <i>de novo</i> methylation on gene expression.....	81
4.1.5 Changes in the epigenetic landscape after global DNA methylation editing.....	84
4.1.6 The implications of bivalency or H3K27me3 on DNA methylation stability	86
4.2 Engineering of an EpiEditor with improved specificity for targeted DNA methylation editing.....	89
4.3 Reprogramming of the <i>H19/IGF2</i> imprinting control region by targeted demethylation.....	95
5. Discussion	100
5.1 Dynamics of the epigenome network after global DNA methylation editing ...	101
5.2 Engineering of an EpiEditor with improved specificity for targeted DNA methylation editing.....	112
5.3 Reprogramming of the <i>H19/IGF2</i> imprinting control region by targeted demethylation.....	114
6. Outlook	118
References	120
Author's contributions	147
Appendix.....	149

Acknowledgements

I am very thankful to my supervisor Prof. Dr. Albert Jeltsch for always being very supportive and offering me the opportunity to work on various exciting projects in this great working group. It has always been a great pleasure for me to work in this inspiring and innovative environment.

I would also like to thank Prof. Dr. Stefan Legewie who has kindly agreed to be the second co-referee, and to Prof. Dr. Stephan Nussberger for being the head of my thesis committee.

Moreover, I want to express my deepest gratitude to my fabulous current and former colleagues from the IBC, who gave me a wonderful and joyful time over the past six years. Thank you very much, Max, Rebekka, Stefan, Micha, Alex, Mike, Sara, Sabine, Philipp R., Pavel, Cristiana, Aga, Sabrina, Tabea, T.C., Nivethika, Franzi, Anja, Mina, Philipp S., Jannis, Gizem, Goran, Maren, Denis, Nicole, Katrin, Johannes, Ingo, Karl, Regina, Dragica, Branka, Elisabeth, Hans, Ruth, Phil, Peter, Tomek, Miru, Renata, Srikanth and Rustem. I will never forget the special moments we experienced together! And also great thanks to all those people associated to the institute who worked with me for only a short time or were supervised by me in their projects.

I would particularly like to thank my wife Karolina, the love of my life, who always supported me and gave me a lot of strength. The same applies for my family, who always provided me a lot of support and showed a lot of interest into my scientific work.

Thank you all - you all kept me motivated and contributed to my personal development as a scientist!

List of publications

Research articles:

Lungu, C., Pinter, S., **Broche, J.**, Rathert, P. & Albert Jeltsch (2017). **Modular fluorescence complementation sensors for live cell detection of epigenetic signals at endogenous genomic sites.** *Nature Communications*, 8(1), 649.

Rajavelu, A., Lungu, C., Emperle, M., Dukatz, M., Bröhm, A., **Broche, J.**, Hanelt, I., Parsa, E., Schiffers, S., Karnik R., Meissner, A., Carell, T., Rathert, P., Jurkowska, R. Z. & Jeltsch, A. (2018). **Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2.** *Nucleic Acids Research*, 6(17), 9044-9056.

Hofacker, D.* , **Broche, J.***, Laistner L., Adam, S., Bashtrykov, P. & Jeltsch, A. (2020). **Engineering of Effector Domains for Targeted DNA Methylation with Reduced Off-Target Effects.** *International Journal of Molecular Sciences*, 21(2), 502.

*co-shared first authors

Broche, J., Kungulovski, G., Bashtrykov, P., Rathert P. & Jeltsch, A. (2021). **Genome-wide investigation of the dynamic changes of epigenome modifications after global DNA methylation editing.** *Nucleic Acids Research*, 49(1), 158-176.

Reviews:

Jeltsch, A., **Broche, J.** & Pavel Bashtrykov (2018). **Molecular Processes Connecting DNA Methylation Patterns with DNA Methyltransferases and Histone Modifications in Mammalian Genomes.** *Genes (Basel)*, 9(11), 566.

Jeltsch, A., **Broche, J.**, Lungu, C. & Pavel Bashtrykov (2019). **Biotechnological Applications of MBD Domain Proteins for DNA Methylation Analysis.** *Journal of Molecular Biology*, 432(6), 1816-1823.

Zusammenfassung

Trotz der großen Vielfalt unterschiedlicher Zelltypen in Menschen besitzen alle Zellen die gleiche genetische Information. Um ein zelltypspezifisches Expressionsprofil zu ermöglichen, muss eine Vielzahl epigenetischer Signale durch verschiedene epigenetische Mechanismen etabliert und aufrechterhalten werden. Posttranslationale Modifikation von Histonproteinen und DNA-Methylierung sind zwei wichtige epigenetische Signale und Bestandteile des epigenomischen Netzwerks, welche die Aktivierung oder Unterdrückung der Genexpression regulieren, indem sie die Zugänglichkeit des Chromatins modulieren. Bei Säugetieren werden DNA-Methylierungsmuster durch DNA-Methyltransferasen etabliert, und fehlerhafte DNA-Methylierungsmuster werden mit verschiedenen Krankheiten, darunter Krebs, in Verbindung gebracht. Daher ist es für die Grundlagenforschung und medizinische Anwendungen von besonderem Interesse, diese Modifizierung besser zu verstehen und den DNA-Methylierungszustand in bestimmten Regionen künstlich zu manipulieren. Dies kann durch den Einsatz von epigenommodifizierenden Werkzeugen erreicht werden, die als „EpiEditoren“ bezeichnet werden, und es ermöglichen, DNA-Methylierung mit variabler Spezifität zu setzen oder zu entfernen.

Ziel des Hauptteils dieser Doktorarbeit war, die Stabilität von künstlich eingefügter DNA-Methylierung in CpG-Inseln (CGIs) zu charakterisieren. Zur Adressierung der artifiziellen DNA-Methylierung wurde ein Zinkfinger (ZnF) eingesetzt, der über ein kurzes DNA-Erkennungsmotiv verfügt und somit zahlreiche genomische Regionen gleichzeitig ansteuern kann. Durch ChIP-Experimente wurden mehr als 15.000 Bindungsstellen des ZnF im Genom identifiziert, was die relativ unspezifische DNA-Interaktion des ZnFs dokumentiert. Eine Fusion der katalytischen Domäne von DNMT3A (3AC) mit dem ZnF ermöglichte es, an Tausenden von CGIs DNA-Methylierung einzubauen. Anschließend wurde die Dynamik dieser Methylierung für bis zu 11 Tage durch MBD2-seq verfolgt. Bemerkenswerterweise war die gesetzte DNA-Methylierung bei etwa 90 % aller zuvor unmethylierten CGIs nur vorübergehend vorhanden. Bei 10 % der CGIs wurde jedoch ein hohes Maß an restlicher Methylierung festgestellt werden. Interessanterweise waren diese stabil methylierten CGIs stark mit H3K27me3 angereichert, das mit Polycomb-reprimiertem Chromatin in Verbindung gebracht wird. Das Setzen von DNA-Methylierung führte zu einer deutlichen Reduktion der aktivierenden Markie-

rungen H3K4me3 und H3K27ac, welche nach dem Verlust der DNA-Methylierung teilweise zu ihrem Ursprungsniveau zurückkehrten. Überraschenderweise wurden keine direkten Auswirkungen der DNA-Methylierung auf die H3K9me3- und H3K36me3-Level beobachtet. Jedoch wurde die Expression von mehr als 900 Genen nach dem Einfügen von DNA-Methylierung in den zugehörigen Promoter mindestens zweifach reduziert, was das Repressionspotential von Promotermethylierung verdeutlicht. Dabei waren die Expressionsänderungen eng mit den zeitlichen Änderungen der Promotor-Methylierungsniveaus korreliert.

Im zweiten Projekt bestand das Ziel darin, einen verbesserten EpiEditor zu entwickeln, der in Lage ist, DNA-Methylierung mit hoher Effizienz am Zielort zu setzen und dabei in anderen Regionen nur geringe Aktivität zu zeigen. Dazu wurden zwei dCas9-basierte Rekrutierungsstrategien mit zwei unterschiedlichen Effektoren verglichen, zum einen 3AC, zum anderen ein chimäres Konstrukt von 3AC, welches mit der C-terminale Domäne von DNMT3L fusioniert wurde (3AC-3L). Die direkte Fusion der Effektor-domänen an dCas9 führte zu einer geringeren Spezifität als die Verbindung von dCas9 und Effektor-domänen in dem dCas9-SunTag-System, das auf der Rekrutierung mehrerer Antikörper-fusionierter Effektor-domänen an eine Aneinanderreihung von GCN4-Peptiden basiert, die an dCas9 fusioniert sind. Obwohl eine hohe Spezifität des dCas9-SunTag-Systems in Kombination mit 3AC beobachtet wurde, war die DNA-Methylierungseffizienz eher gering. Im Gegensatz dazu war dCas9-SunTag zusammen mit 3AC-3L effizienter, zeigte aber auf genomweiter Ebene eine sehr hohe unspezifische Aktivität. Durch das Einführen von ladungsumkehrenden Mutationen in basische Aminosäuren, die für die Interaktion von 3AC-3L mit DNA erforderlich sind, konnte die unspezifische DNA-Methylierung des EpiEditors aufgehoben werden, ohne die zielgerichtete Aktivität wesentlich zu beeinträchtigen. Am Ende übertraf der neuartige EpiEditor das bisher verfügbare dCas9-SunTag-Konstrukt in Kombination mit 3AC.

Das dritte Projekt dieser Arbeit zielte darauf ab, die H19/IGF2-Imprinting-Kontrollregion (ICR) durch gezielte DNA-Demethylierung umzuprogrammieren. Zu diesem Zweck wurde das dCas9-SunTag-System in Kombination mit der katalytischen Domäne von TET1 verwendet. Mehrere CpG-reiche Motive innerhalb der ICR konnten unter Verwendung eines multi-sgRNA-Plasmids adressiert werden. Dabei konnte eine starke Verringerung der DNA-Methylierung festgestellt werden. Anhand von CTCF-ChIP-Ex-

perimenten konnte gezeigt werden, dass die DNA-Demethylierung mit der Rekrutierung des methylierungssensitiven Isolatorproteins CTCF einherging. Die reduzierte DNA-Methylierung und die erhöhte CTCF-Anreicherung wurden bis zu vier Wochen aufrechterhalten, was die große Robustheit des umprogrammierten Zustands illustriert.

Insgesamt liefern die gewonnenen Daten wertvolle Informationen über die Dynamik der gezielten DNA-Methylierung und ihre regulatorischen Wirkungen. In Kombination mit der Entwicklung besserer EpiEditoren kann diese Doktorarbeit den Weg zu sichereren und effizienteren klinischen Anwendungen von Epigenom-Editierung ebnen.

Abstract

Despite of the large variety of different cell types in humans, all cells carry the same genetic information. In order to enable a cell type-specific expression profile, a multitude of epigenetic signals has to be established and maintained by various epigenetic mechanisms. Histone post-translational modifications and DNA methylation are two major epigenetic signals and parts of the epigenome network, which regulate the activation or silencing of gene expression by modulating the chromatin accessibility. In mammals, DNA methylation patterns are established by DNA methyltransferases, and aberrant methylation patterns have been connected to various disorders including cancer. Thus, it is of particular interest for basic research and medical applications to gain a better understanding of this mark, and to artificially manipulate the DNA methylation state at a certain locus. This can be achieved by the employment of diverse epigenome editing tools, termed 'EpiEditors', allowing to install or remove DNA methylation with variable specificity.

Aim of the main part of this PhD work was to characterize the stability of DNA methylation at CpG islands (CGIs) after artificial introduction. For installation of DNA methylation, a zinc finger (ZnF) was used as targeting device, which has a short DNA binding motif allowing to target numerous genomic loci simultaneously. By conducting ChIP experiments, more than 15,000 ZnF peaks were identified, demonstrating its promiscuous binding in the genome. The fusion of the DNMT3A catalytic domain (3AC) to the ZnF enabled to deposit DNA methylation at thousands of CGIs. Afterwards, the dynamics of this mark were tracked for up to 11 days by MBD2-seq. Strikingly, the installed DNA methylation was only transient at around 90 % of all previously unmethylated CGIs. However, high levels of residual DNA methylation were observed at 10 % of the CGIs. Intriguingly, these stably methylated CGIs were strongly enriched in H3K27me3, a mark associated with Polycomb group chromatin. The deposition of DNA methylation resulted in a marked reduction of the activating marks H3K4me3 and H3K27ac, which were partially recovered upon DNA methylation loss. Surprisingly, no direct effects of DNA methylation were observed on H3K9me3 and H3K36me3 levels. However, the expression of more than 900 genes was downregulated at least two-fold after DNA methylation editing of the corresponding promoter, demonstrating the high

silencing capacity of the mark. Thereby, the expression changes were tightly correlated to the temporal changes promoter methylation levels.

In the second project of this thesis, the goal was to develop an improved EpiEditor, capable to introduce DNA methylation with high efficiency at the target locus but with low off-target activity. For this, two dCas9-based targeting strategies in combination with two different effector domains, 3AC and a chimeric construct of 3AC fused to the C-terminal domain of DNMT3L (3AC-3L), were compared. The direct fusion of the effector domains to dCas9 resulted in a lower specificity than a connection of dCas9 and effector domain with a dCas9-SunTag system, which is based on the recruitment of multiple antibody-fused effector domains to an array of GCN4 peptides fused to dCas9. Although a high specificity of the dCas9-SunTag system in combination with 3AC was observed, the DNA methylation efficiency was rather low. In contrast, dCas9-SunTag together with 3AC-3L was more efficient, but on genome-wide scale, the off-target activity was very high. By introducing charge-reversal mutations into basic residues required for the interaction of 3AC-3L with DNA, most of the off-target activity of the EpiEditor could be removed without compromising its on-target activity strongly. In the end, the novel EpiEditor outperformed the previously available dCas9-SunTag construct in combination with 3AC.

The third project of this work aimed to reprogram the H19/IGF2 imprinting control region (ICR) by targeted DNA demethylation. To this end, the dCas9-SunTag system in combination with the catalytic domain of TET1 was employed. Multiple CpG-rich motifs within the ICR were targeted by using a multi-sgRNA plasmid. Strikingly, a strong reduction in DNA methylation was obtained. By conducting CTCF-ChIP, it could be demonstrated that the DNA demethylation was accompanied by the recruitment of the methylation-sensitive insulator protein CTCF. The reduced DNA methylation and the increased CTCF occupancy were maintained for almost four weeks, indicating the robustness of the reprogrammed state.

All in all, the obtained data provide valuable information on the dynamics of targeted DNA methylation and its regulatory effects. In combination with the development of better EpiEditors, this PhD work may pave the way towards safer and more efficient clinical applications of EpiEditing.

List of abbreviations

3AC	Catalytic domain of DNMT3A
3BC	Catalytic domain of DNMT3B
3L/3LC	C-terminal part of DNMT3L
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
6mA	N6-methyladenine
aa	amino acid
Ab	Antibody
Ac	Acetyl
ADD domain	ATRX-Dnmt3a-Dnmt3l domain
AdoHcy	S-Adenosyl-L-homocysteine
AdoMet	S-Adenosyl-L-methionine
α KG	α -Ketoglutarate
bp	Base pair
BWS	Beckwith-Wiedemann Syndrome
CGI	CpG island
ChIP	Chromatin immunoprecipitation
COMPASS	Complex of Proteins Associated with Set1
CpG	Cytosine-guanine dinucleoside (linked by phosphate)
CRISPR	Clustered Regulatory Interspaced Short Palindromic Repeats
CTCF	CCCTC-binding factor

CTS	CTCF target site
DBD	DNA binding domain
dCas9	Deactivated CRISPR-associated protein 9
ddH ₂ O	Double-distilled water
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethylsulfoxid
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
DNMT	DNA (cytosine-5)-methyltransferase
dNTP	2'-Desoxynucleotid-5'-triphosphate
dox	Doxycycline
DPBS	Dulbecco's phosphate-buffered saline
ENCODE	Encyclopedia of DNA elements
EpiEditor	Epigenome editing tool
EPZ-6438	EZH2 inhibitor (also known as 'Tazemetostat')
EV	Empty vector
FACS	Fluorescence-activated cell sorting
FF interface	Hydrophobic Phe/Phe interface of DNMT3 enzymes
FPKM	Fragments per kilobase of exon model per million reads mapped
GCN4	Epitope tag derived from the yeast Gcn4p protein
GFP	Green fluorescent protein
gDNA	Genomic DNA
GST (tag)	Glutathione S-transferase (tag)
H3K27ac	Histone H3 acetylated at lysine 27

H3K27me3	Histone H3 trimethylated at lysine 27
H3K36me3	Histone H3 trimethylated at lysine 36
H3K4me3	Histone H3 trimethylated at lysine 4
H3K9me3	Histone H3 trimethylated at lysine 9
HA-tag	Hemagglutinin tag
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HEK293	Human Embryonic Kidney 293 cell line
HEK293R	HEK293 cell line expressing the ecotropic receptor
HMTase	Histone methyltransferase
ICR	Imprinting control region
IRES	Internal ribosomal entry site
kb	Kilobases
KD	Knockdown
KDM	Lysine demethylase
KO	Knockout
LINE	Long Interspersed Nuclear Elements
LOI	Loss of imprinting
MBD	Methyl-CpG-binding domain
me1	Monomethyl
me2	Dimethyl
me3	Trimethyl
MPGC	Max Planck-Genome-centre Cologne
mRNA	Messenger RNA
XIV	

MTase	Methyltransferase
N-OG	N-Oxalylglycine
NChIP	Native chromatin immunoprecipitation
NGS	Next-Generation Sequencing
NLS	Nuclear localization signal
nt	Nucleotide
PAM	Protospacer adjacent motif
PcG	Polcomb group
PCR	Polymerase Chain Reaction
Plat-E	Platinum-E virus packaging cell line
PRC1	Polycomb Repressive Complex 1
PRC2	Polycomb Repressive Complex 2
PTM	Post-translational Modification
PWWP domain	Pro-Trp-Trp-Pro motif-containing domain
qPCR	Quantitative polymerase chain reaction
RD interface	Polar Arg/Asp interface of DNMT3 enzymes
RNA	Ribonucleic acid
RNAPII	RNA polymerase II
RT	Room temperature
RT-qPCR	Reverse transcription quantitative PCR
rtTA3	Reverse tetracycline-controlled transactivator 3
SDS PAGE	Sodium dodecyl sulfate Polyacrylamide Gel Electrophoresis
sgRNA	Single guide RNA
shRNA	Short hairpin RNA

SunTag	SUPerNova tag
TALE	Transcription Activator-Like Effector
TET	Ten-Eleven Translocation
TET1(CD)	Catalytic domain of TET1
TF	Transcription factor
TRD	Target-recognition domain
TSS	Transcriptional start site
Ub	Ubiquitin
XChIP	Cross-linked chromatin immunoprecipitation
XCI	X-chromosome inactivation
ZnF protein	Zinc finger protein

1. Introduction

1.1 Epigenetics and the structure of chromatin

In the year 1942, the developmental biologist Conrad H. Waddington coined the term 'epigenetics' in order to describe the developmental mechanisms that connect the genotype and phenotype of an organism. The prefix 'epi-' comes from the Greek and literally translates into 'in addition to', thus the meaning of epigenetics can be described as "in addition to the genetic sequence." Waddington argued that an adaptive response to an environmental stimulus does not necessarily require a mutation within the genome, and instead depends on an adequately controlled genetic reaction (Waddington 1942, 2012; Weinhold 2006).

In fact, the human body is composed of around 200 different cell types, which all contain the same genetic information, and originate from the same totipotent cells. Waddington proposed the 'epigenetic landscape' (Figure 1A) as a metaphor to illustrate the cellular decision-making during development. Thereby, the cells reach a more and more differentiated state that at final stages can hardly be altered by perturbations (Baedke 2013; Hatano et al. 2011). With the scientific progress in studying epigenetic mechanisms arising over several decades, Wu and Morris developed a refined definition of epigenetics in 2001. They described epigenetics as "the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence" (Wu and Morris 2001).

In order to implement an epigenetic state that can be stably inherited, three types of signals have to consecutively cooperate with each other. First, an environmental signal (e.g. hormones, temperature, nutrition), the so-called 'Epigenator', acts on the cell and triggers a downstream signal cascade (Figure 1B). This cascade is then executed by the 'Epigenetic Initiator', which translates and transmits the signal from the Epigenator to a specific genomic locus. Finally, the 'Epigenetic Maintainer' modulates the epigenetic state of the underlying locus, and establishes the heritable epigenetic phenotype (Berger et al. 2009; Li et al. 2014).

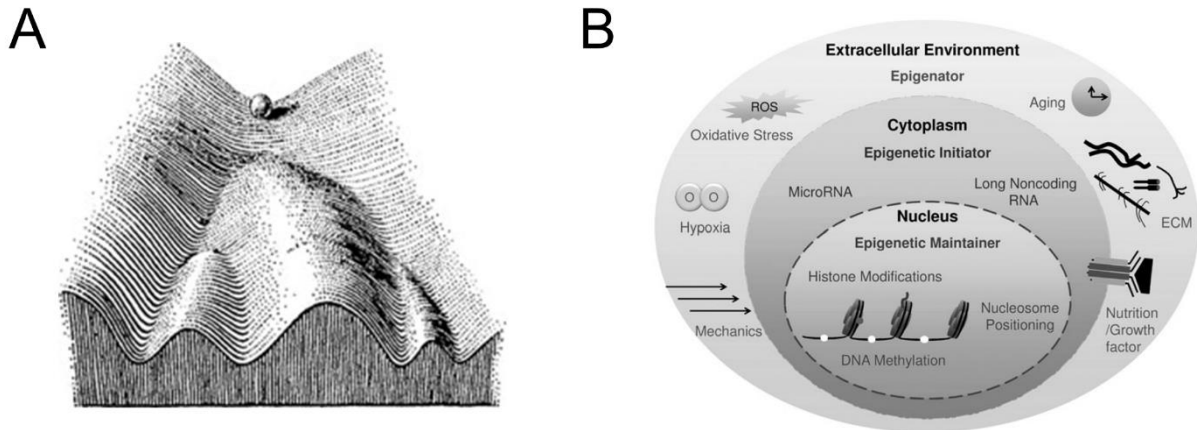


Figure 1: A) Artistic depiction of the epigenetic landscape model proposed by Waddington. The ball rolling down the landscape in differential paths represents the cellular differentiation during development. Taken from Baedke (2013). B) Illustration of the three epigenetic signals involved in the establishment of an epigenetic state. The epigenator induces a downstream process, which is transmitted to the nucleus by the epigenetic initiator. Finally, the epigenetic maintainer alters the epigenetic state. Taken from Li et al. (2014).

Three epigenetic signals have been described to shape the epigenome of a cell, which are non-coding RNAs (ncRNAs), as well as covalent chemical modifications such as DNA methylation and histone post-translational modifications (PTMs) (Figure 2) (Blake and Watson 2016). To gain a better understanding of how these epigenetic signals regulate the cellular processes, it is necessary to understand the principal organization of the genetic information in the cell. The DNA of a human diploid cell spans around two meters in total, and has to fit into a nucleus which is usually around 10 - 20 μm in diameter. For this reason, the DNA has to be condensed into a structure named 'chromatin' by association with histones and non-histone proteins. The smallest unit of chromatin is the nucleosome, in which 147 bp of DNA are wrapped around a histone octamer 1.7-times. A histone octamer in turn is composed of two copies of the core histone proteins H2A, H2B, H3, and H4. The tight wrapping of the DNA in a nucleosome is facilitated by the attraction of the negatively charged sugar-phosphate backbone of DNA to the positively charged surface of the histones. These core histone proteins all have a globular structure at their C-terminus, and contain a large unstructured N-terminal tail, which is target for various post-translational modifications (Gordon et al. 2013; Kouzarides 2007).

Two nucleosomes are separated by linker DNA with a length between 20 - 90 bp, forming the 11 nm-wide 'beads-on-a-string' structure. By association of linker histone H1, or other chromatin-associated proteins, the formation of a more condensed 30 nm

chromatin fiber is induced, presumably adopting various conformations *in vivo*. In addition, the 30 nm fiber can further be organized in three-dimensional structures, and reach the highest condensed state in the form of a chromosome (Fyodorov et al. 2018; Gordon et al. 2013; Szerlong and Hansen 2011).

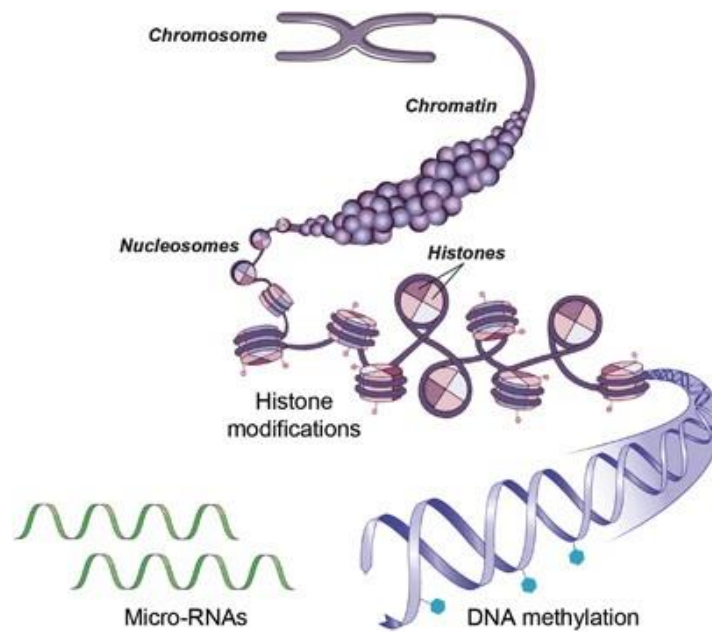


Figure 2: Schematic depiction of the chromatin organization in the cell. The DNA is wrapped around histone octamers, together forming the nucleosome, and a beads-on-a-string structure. Further compaction of chromatin leads to the formation of the 30 nm chromatin fiber, and the highest condensed chromatin structure, the chromosome. Epigenetic mechanisms like histone modifications, DNA methylation, and non-coding RNAs control gene expression and affect the chromatin states. Figure taken from Weber-Stadlbauer (2017).

In eukaryotic genomes, chromatin appears in different states of compaction. Among them, euchromatin represents the most relaxed state, which for instance enables transcription factors or RNA polymerase II to access the DNA, and induce gene expression. In contrast, heterochromatin is highly compacted, and hence is connected to gene repression as it prevents the association of gene-activating factors to DNA. It can be subdivided into two states, constitutive and facultative heterochromatin. Constitutive heterochromatin is rather static and, *inter alia*, localized at telomeres and centromeres to keep these chromatin regions transcriptionally inert. In comparison, facultative heterochromatin is particularly involved in gene repression during development and can be dynamically switched towards euchromatin or constitutive heterochromatin. The presence of histone PTMs can directly or indirectly affect the compaction of the chromatin, and define the chromatin state. On the one hand, this can be achieved directly due to structural perturbation of the local chromatin, e.g. by reduction of the

positive charge of the histones. On the other hand, histone PTMs can recruit chromatin factors that modify the chromatin structure in distinct modes (Bannister and Kouzarides 2011; Saksouk et al. 2015; Trojer and Reinberg 2007). The combinations of specific histone tail PTMs create the so-called ‘histone code’, and thereby also define the chromatin state (Jenuwein and Allis 2001).

1.2 DNA methylation

1.2.1 Occurrence of DNA methylation in different organisms

DNA methylation is an epigenetic modification which has already been discovered in 1925 in the prokaryote *Mycobacterium tuberculosis* (Johnson and Coghill 1925). The existence of modified cytosines in mammals has first been proven by Hotchkiss (1948) who separated nucleosides originating from calf thymus by paper chromatography (Hotchkiss 1948). Since DNA methylation has also been detected in archaea and plants, its presence has thus been demonstrated in all kingdoms of life (Blow et al. 2016). So far, three different variants of methylated nucleobases have been described which are N6-methyladenine (6mA), 5-methylcytosine (5mC), and N4-methylcytosine (4mC) and which are all found in bacteria (Figure 3) (Beaulaurier et al. 2019). The transfer of a methyl group from S-aden

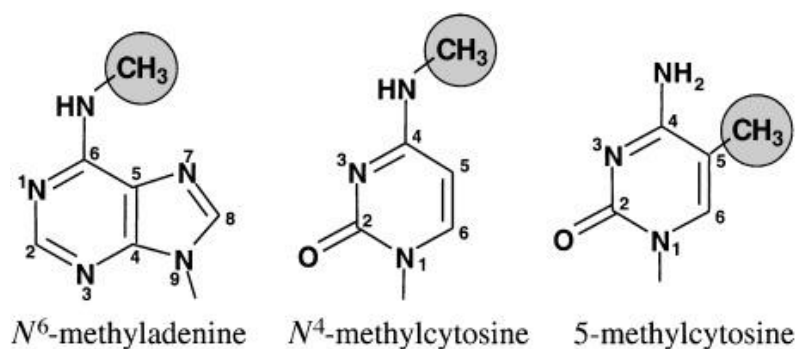


Figure 3: Chemical structures of methylated nucleobases, introduced by DNA methyltransferases. The inserted methyl group is highlighted by a grey circle. The image is taken from Jeltsch (2002).

osyl-L-methionine (SAM, or AdoMet) to a nucleobase is catalyzed by so-called DNA methyltransferases (MTase), which will be described in the following chapter in more

detail. Methylation can occur at the exocyclic amino group of adenine (6mA) or cytosine (4mC). In the case of 5mC, the methyl group is added at C5 position of cytosine (Jeltsch 2002). 6mA has also been found in various genomes of multicellular organisms (Luo et al. 2015), and recently also at very low levels in mammals (Wu et al. 2016; Xiao et al. 2018). The best characterized DNA modification, however, is 5mC which is highly conserved across animals, fungi, and plants, and only few species like *C. elegans* or *S. cerevisiae* lost DNA methylation in the course of evolution (Zemach and Zilberman 2010). In mammals, cytosine methylation usually occurs in palindromic CpG sequences (Jones 2012), but has also been detected in CpH (H = A, C or T) sequence context in human neuronal tissues and embryonic cells (Jeong et al. 2021; Kozlenkov et al. 2016).

1.2.2 CpG dinucleotides and CpG islands

Between 1.9 and 4.5 % of all cytosines are methylated in different human tissues and cell lines (Globisch et al. 2010; Lisanti et al. 2013). However, CpG dinucleotides are almost five times underrepresented in the genomes of humans and mice (Gentles and Karlin 2001; Simmen 2008). Considering this sequence bias, between 70 and 80 % of all CpGs are methylated in somatic cells (Bird 2002). The low proportion of CpGs can be accounted to the spontaneous hydrolytic deamination of methylated cytosines to thymine, resulting in C-to-T mutations on one of the DNA strands, which can then be propagated by replication (Fryxell and Moon 2005).

Despite the overall low abundance of CpG dinucleotides, there are genomic regions with high CpG density, named CpG islands (CGIs). A well accepted definition for the properties of CGIs is based on their length, which ranges from 500 bp to several kilobases, on their observed over expected CpG ratio larger than 0.65, and on their GC content, which has to be equal or larger than 55 % (Takai and Jones 2002). CGIs are often flanked by CGI shores which can be up to 2 kb in length and have a lower CpG content. In addition, regions which are 2 - 4 kb away from CGIs are known as 'CGI shelves' and the regions beyond this distance are termed as 'open sea' (Marzese and Hoon 2015). In the human genome, around 30,000 CGIs were identified, which are mostly located at gene promoters (49 %) or within gene bodies (32 %) (Jeziorska et al. 2017). Around 72 % of all gene promoters are containing CGIs, and classified as

high-CpG promoters (HCPs), whilst 28 % are non-CGI promoters and, hence, classified as low-CpG promoters (LCPs). Overall, promoters display a bimodal distribution in regard to their CpG content, with a small population of intermediate-CpG promoters (ICPs). In somatic cells, LCPs are usually hypermethylated, which does not influence the expression of the respective gene. ICPs are hypermethylated in somatic cell as well, however, their activity is negatively affected by methylation. In contrast, HCPs are mostly unmethylated, regardless of their activity, and the CpG content within this group does not correlate with gene expression (Saxonov et al. 2006; Weber et al. 2007). An overall low CGI methylation was also observed in hematopoietic stem cells (HSC) (Figure 4) (Jeong and Goodell 2014). Moreover, a CpG island methylator phenotype (CIMP) has been described in cancer cell lines, and some subtypes of colorectal cancer, showing aberrant methylation, and downregulation of numerous tumor-suppressor genes (Smiraglia et al. 2001; Toyota et al. 1999).

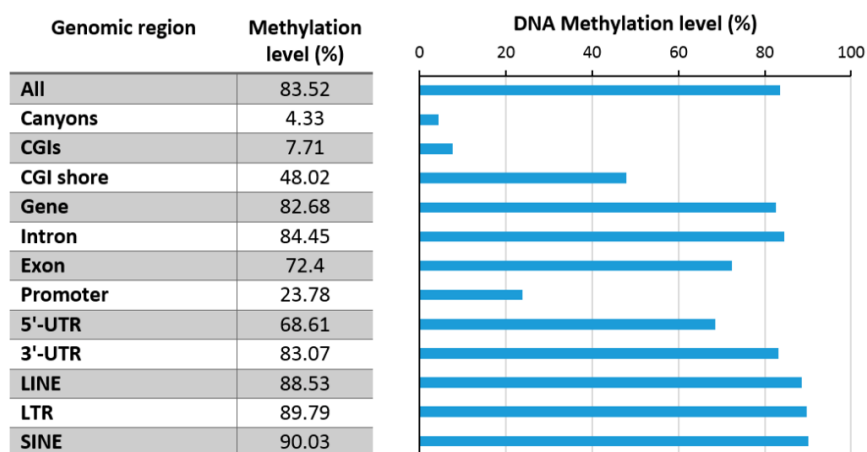


Figure 4: Relative methylation levels of different genomic elements in mouse hematopoietic stem cells obtained by whole genome bisulfite sequencing. The figure is adapted from Jeltsch et al. (2019).

1.2.3 Regulatory functions of DNA methylation

In general, DNA methylation is connected to the repression of genes by various mechanisms, which include the induction of heterochromatin or the silencing of enhancers (Figure 5B) (Jones 2012). The effect of promoter methylation on gene expression can be explained by a differential sensitivity of DNA-binding factors in the presence of this modification (Figure 5E). It was shown that the binding of several transcription factors (TFs) of the bHLH-, bZIP-, and ETS-family, but also many others, was reduced by DNA

methylation. *Vice versa*, the binding of some other TFs, e.g. Klf4, NFAT- or POU family members, to DNA was promoted by cytosine methylation (Gaston and Fried 1995; Spruijt et al. 2013; Yin et al. 2017). TFs can also recruit DNA methyltransferases to a specific genomic locus, and hence can induce repression of an associated gene (Brenner et al. 2005; Suzuki et al. 2006).

A special class of 5mC readers, the methyl-CpG-binding domain (MBD) protein family, has been identified, and consists of seven members, which are MBD1-6 and MeCP2 (Du et al. 2015). *In vitro* studies showed that MBD1, MBD2, and MBD4 display a high specificity for 5mC. In addition, MBD3 and MeCP2 were also capable of binding to 5-hydroxymethylcytosine (5hmC), an oxidative derivative of 5mC (Mellén et al. 2012). Another group reported that MBD3 even has a strong preference for 5hmC over 5mC (Yildirim et al. 2011). MBD5 and MBD6 are not capable to bind methylated DNA *in vitro*, despite of their classification as MBD proteins (Laget et al. 2010). The readout of DNA methylation by MBD proteins is entailed by multiple downstream effects, such as epigenetic remodeling or gene silencing. MBD1, for instance, can recruit the histone lysine methyltransferases (HKMTs) SETDB1 and the SUV39H1-HP1 complex to methylated loci, and hence induce the formation of heterochromatin (Fujita et al. 2003; Ichimura et al. 2005). MBD2 and MBD3 are known interaction partners of the NuRD/Mi-2 complex, which in turn is capable to erase active enhancers or promoter marks via histone deacetylase (HDACs) or H3K4 demethylase activity. In contrast, binding of MBD4 to 5mC is mainly connected to DNA repair (Du et al. 2015).

MeCP2 was shown to be part of the mSin3A, CoREST, and the NCoR/SMRT corepressor complexes, which are both involved in transcriptional repression (Lunyak et al. 2002; Lyst et al. 2013; Suzuki et al. 2003). Further roles of MeCP2 in long-range chromatin interactions, and binding to HP1 in the context of heterochromatin formation have also been demonstrated (Agarwal et al. 2007; Fuks et al. 2003a; Kernohan et al. 2014). In addition, MeCP2 was found to directly interact with DNMT3 enzymes and recruit them to pericentromeric heterochromatin (Rajavelu et al. 2018). These MBD-mediated downstream processes resulting from DNA methylation highlight the key regulatory role of this modification.

Elevated levels of DNA methylation have been shown at the gene bodies of active genes in dividing cells or cell lines, while methylation was reduced at the transcriptional start site (TSS) of these genes at the same time (Aran et al. 2011; Li et al. 2010). In H1 human embryonic stem cells, CpG methylation in gene bodies did not correlate with gene expression, whereas CHH and CHG methylation in exons, introns and at the 3'UTR was strongly increased in highly expressed genes (Lister et al. 2009). It was demonstrated that CpG methylation has a role in pre-mRNA alternative splicing (Figure 5C), as exons are hypermethylated compared to their neighboring introns, thus demarcating the exon-intron boundaries of a gene (Lev Maor et al. 2015). On the other hand, methylation of alternative intragenic CGI promoters is accompanied by their silencing (Greenberg and Bourc'his 2019).

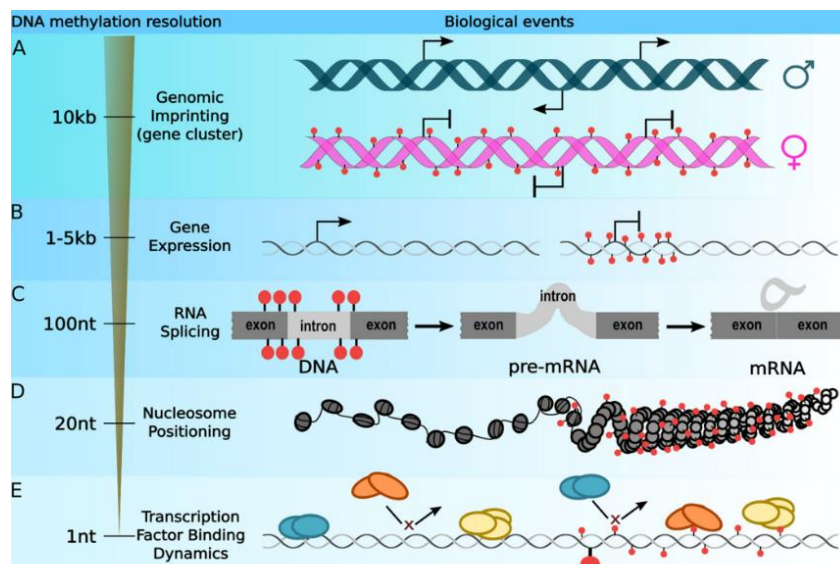


Figure 5: The diverse regulatory functions of DNA methylation in the genome on different genomic scales. DNA methylation can directly influence A) genomic imprinting, B) gene expression, C) RNA splicing, D) nucleosome positioning, and E) transcription factor binding. The image was taken from Tirado-Magallanes et al. (2017).

More than half of the human genome consist of repetitive elements, which include Long Interspersed Nuclear Elements (LINEs), Short Interspersed Nuclear Elements (SINEs), Long Terminal Repeats (LTRs), RNA repeats, and DNA transposons. These genomic elements are all depleted three to five times in CpG dinucleotides and are methylated by approximately 80 – 90 % in human embryonic stem cells, and in mouse HSC. The high levels of methylation (Figure 4) are of great relevance, as they support the transcriptional repression of repeats, and thus safeguard genome stability (Jeong

and Goodell 2014; Su et al. 2012). When retrotransposon elements become hypomethylated, they are de-repressed, randomly incorporate into the genome, and thus compromise genome integrity. (Carnell and Goodman 2003). Hypomethylation of LINEs has been documented for different types of cancer and human cell lines (Dante et al. 1992). Moreover, an increased telomeric recombination rate was described in cells displaying hypomethylation at subtelomeric regions (Gonzalo et al. 2006). DNA methylation also participates at the transcriptional silencing of repetitive sequences at pericentric satellite repeats (Lehnertz et al. 2003). Reduced methylation levels of repeats can also lead to gene deletions or genomic rearrangement in consequence of mitotic recombination (Chen et al. 1998).

A further important regulatory function of DNA methylation in mammals is its involvement in a phenomenon called X-chromosome inactivation (XCI), leading to the formation of the Barr Body in female cells. DNA methylation contributes to the induction of a repressive chromatin state of the inactivated X-chromosome, which is critical for dosage compensation in gene expression between both sexes (Heard and Distèche 2006).

1.2.4 DNA methylation in development

In the course of mammalian development, reprogramming of the DNA methylation pattern occurs twice. The first wave of reprogramming occurs during gametogenesis, when primordial germ cells (PGCs), which are derived from posterior epiblast cells, experience genome-wide demethylation (Figure 6). This process enables the recovery of germ cell identity, followed by the re-establishment of a sex-specific methylation pattern. (Greenberg and Bourc'his 2019; Hanna et al. 2018; Zeng and Chen 2019). Following fertilization, the paternal and maternal genomes are initially located in separate pronuclei. During pre-implantation development, the zygote experiences the second wave of reprogramming to attain totipotency. Thereby, the paternal genome first undergoes active, then passive demethylation upon blastocyst stage. In parallel, the maternal genome is mainly demethylated passively by multiple rounds of cell division. During the post-implantation period, the paternal and maternal genome become rapidly re-methylated, and global methylation levels reach around 80 % upon epiblast stage. (Greenberg and Bourc'his 2019; Hanna et al. 2018; Lee et al. 2014).

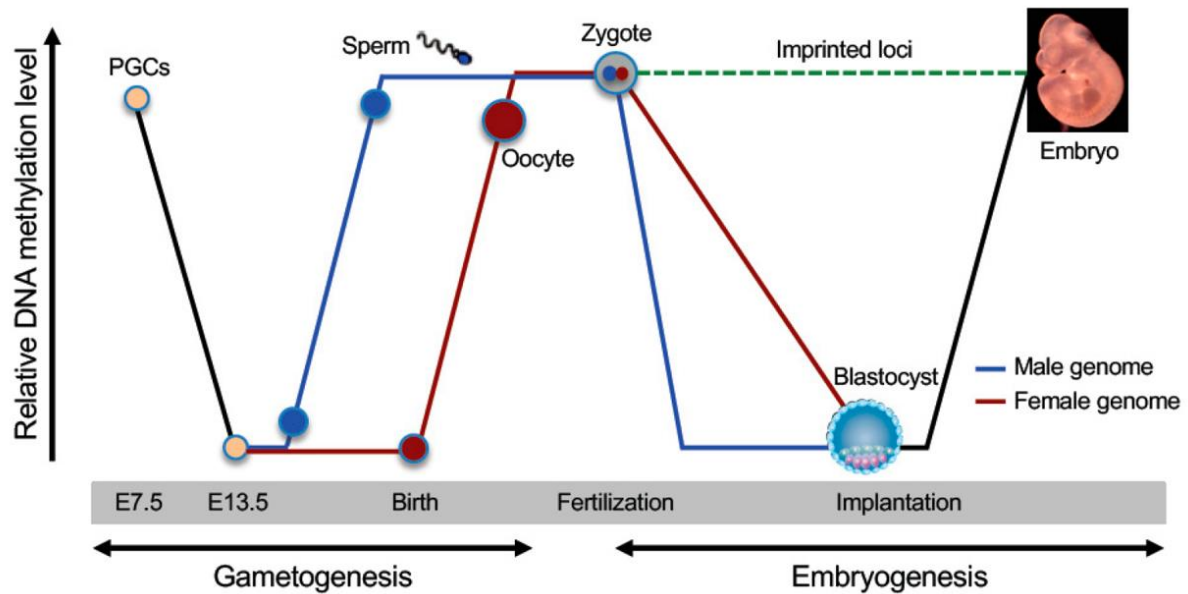


Figure 6: Schematic depiction of the genome-wide DNA methylation changes throughout mammalian development. During gametogenesis and embryogenesis, the maternal and paternal genomes undergo two waves of methylation reprogramming. Thereby, the primordial germ cells (PGCs) and fertilized zygote transiently lose the majority of their DNA methylation by active or passive demethylation. The methylation states of imprinted loci are maintained during embryogenesis. Picture taken from Zeng and Chen (2019).

In the first wave of remethylation in sperm and oocyte development, so-called ‘genomic imprints’ are established as well. Several hundred germline differentially methylated regions (DMRs) on the maternal alleles could be identified in humans and mice which were mostly located at CGIs and promoters (Sanchez-Delgado et al. 2016; Smith et al. 2014). Moreover, around 170 genes in humans or 200 genes in mice are known to show monoallelic expression either from the paternal or the maternal allele (Tucci et al. 2019). Many of the imprinted genes are involved in the regulation of growth during embryogenesis, as well as of the placenta and of the newborn (Barlow and Bartolomei 2014).

Imprinted genes often form large clusters that all contain an imprinting control region (ICR). The ICRs are differentially methylated on the paternal or maternal allele. Around 20 ICRs have been discovered in mouse and human, of which only three are methylated on the paternal allele and localized in intergenic regions. They are protected against demethylation in the second wave of reprogramming, and preserved in somatic tissues. Based on the methylation state of the ICRs, the expression of the surrounding gene cluster is regulated. Dysregulation of imprinted genes is associated with various

congenital diseases, and different types of cancer (Greenberg and Bourc'his 2019; Zeng and Chen 2019).

One well-characterized example for genomic imprinting is the allele-specific regulation of the insulin-like growth factor 2 (*IGF2*) and the 2.3-kb lncRNA *H19* at the 11p15.5 locus on chromosome 11 in human (7qF5 in mouse). While in healthy cells, *IGF2* is solely expressed from the paternal allele, *H19* expression is reciprocal, and only takes place from the maternal allele (Figure 7). An ICR which is located 2 - 4 kb upstream of the *H19* controls the allelic-specific expression of both genes. It displays a high CpG density and is methylated on the paternal, but not on the maternal allele (Baulina et al. 2021; Nordin et al. 2014). Loss of imprinting (LOI) was observed when this locus was deleted in mouse. Within the *H19/IGF2* ICR, multiple binding sites (five in mouse, seven in human) of the CCCTC-binding factor (CTCF) could be identified (Bell and Felsenfeld 2000; Hark et al. 2000).

CTCF is a vertebrate insulator protein which contains a tandem array of eleven C2H2 zinc fingers, recognizing a 15-bp core motif. Binding of CTCF to the consensus motif is impaired when the second cytosine in the motif is methylated (Hashimoto et al. 2017; Kim et al. 2007). The unmethylated ICR of the maternal allele allows CTCF to bind. As a consequence, the accessibility of an enhancer, which is located downstream of the *H19* gene, is blocked towards the *IGF2* promoter (Figure 7). The CTCF-bound ICR

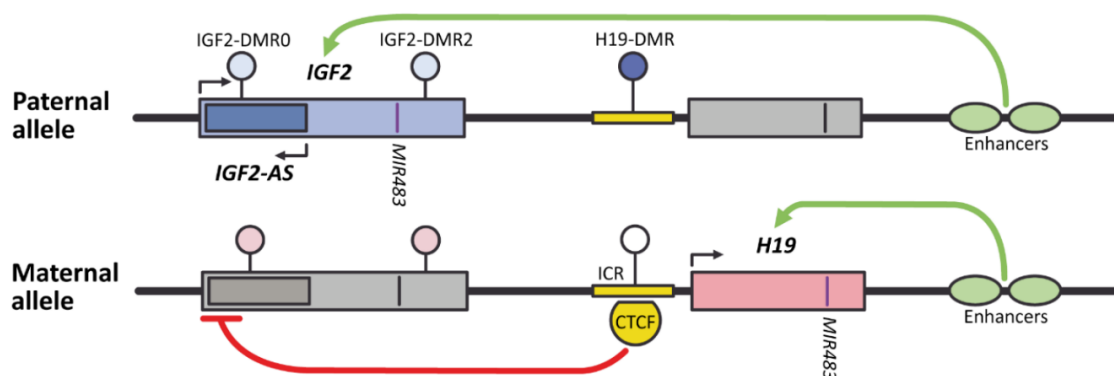


Figure 7: Epigenetic regulation of the *H19/IGF2* imprinting control region (ICR). The maternal allele is unmethylated (white lollipop) at the ICR, which enables CTCF to bind. CTCF then blocks the access of the enhancers to the *IGF2* locus and instead promotes *H19* expression. On the paternal allele, the *H19* ICR locus is methylated (dark blue lollipop), which impairs CTCF binding. Thus, the enhancers can contact the *IGF2* promoter and induce gene expression. The scheme was adapted from Baulina et al. (2021).

thus functions as an insulator, and the enhancer promotes *H19* expression instead. On the paternal allele, CTCF binding is impeded due to methylation of the ICR, enabling the enhancer to activate *IGF2* expression, whereas *H19* is downregulated (Bell and Felsenfeld 2000; Hark et al. 2000). At the ICR, higher-order chromatin structures are formed due to CTCF binding. This additionally requires the association of cohesin, leading to chromatin loops which connect two genomic loci in *cis* (Kurukuti et al. 2006; Nativio et al. 2009).

In mouse experiments, the *H19* lncRNA was shown to function as a negative regulator of *IGF2* expression as well as on the imprinting gene network in *trans*. The body weight of the animals was anti-correlated to *H19* transcript levels, thus highlighting its impact on growth (Gabory et al. 2009). In contrast, the *IGF2* gene encodes a pre-cursor hormone which is post-translationally processed to the final 67-aa peptide hormone. IGF2 promotes multiple processes such as cell proliferation, growth or differentiation (Bergman et al. 2013). An imbalance of *IGF2* and *H19* expression, caused by epimutations in the ICR is associated with several congenital diseases. In up to 64 % of all patients diagnosed with Silver-Russell Syndrome (SRS), the ICR is hypomethylated at the paternal allele, followed by downregulation of the *IGF2* gene. The phenotype of SRS patients includes intrauterine and postnatal growth retardation, a triangular face morphology, and body asymmetry (Gicquel et al. 2005; Netchine et al. 2007).

Opposite to SRS, in Beckwith-Wiedemann Syndrome (BWS), the maternal allele of the *H19/IGF2* ICR is hypermethylated in 5-10 % of all cases. This in turn prevents CTCF binding, which is accompanied with LOI, and expression of *IGF2* from both alleles. Patients with BWS usually display a phenotype of pre- and postnatal overgrowth, macroglossia and visceromegaly (Brown et al. 1996; Murrell et al. 2004; Nativio et al. 2011). Analogous to BWS, the *H19/IGF2* ICR is also hypermethylated in Wilms' Tumor. This type of cancer (nephroblastoma) typically occurs during childhood and is derived from kidney precursor cells in embryogenesis. A 20- to 80-fold downregulation of *H19* together with elevated *IGF2* levels was reported in this cancer, resulting from the loss of heterozygosity and the methylation of CTCF target sites (Cui et al. 2001; Moulton et al. 1994; Steenman et al. 1994).

1.3 Mechanisms of DNA maintenance methylation

In mammals, the methylation of cytosines is introduced by enzymes of the DNA (cytosine-5-)-methyltransferase (DNMTs) family, and occurs mostly in palindromic CpG sequence context. Currently, it is still highly debated by which means specific DNA methylation patterns are established, and how they are maintained over multiple cell generations. The classical model of the DNA methylation cycle is depicted in Figure 8, and represents the general mechanisms regulating the gain and loss of methylation. Cell type-specific DNA methylation patterns are established in the course of embryogenesis by the two *de novo* DNMT3 enzymes (DNMT3A and DNMT3B). These patterns can be inherited over multiple cell generations in somatic cells. After DNA replication, the parental strand is methylated, whereas the newly synthesized strand is unmethylated. The

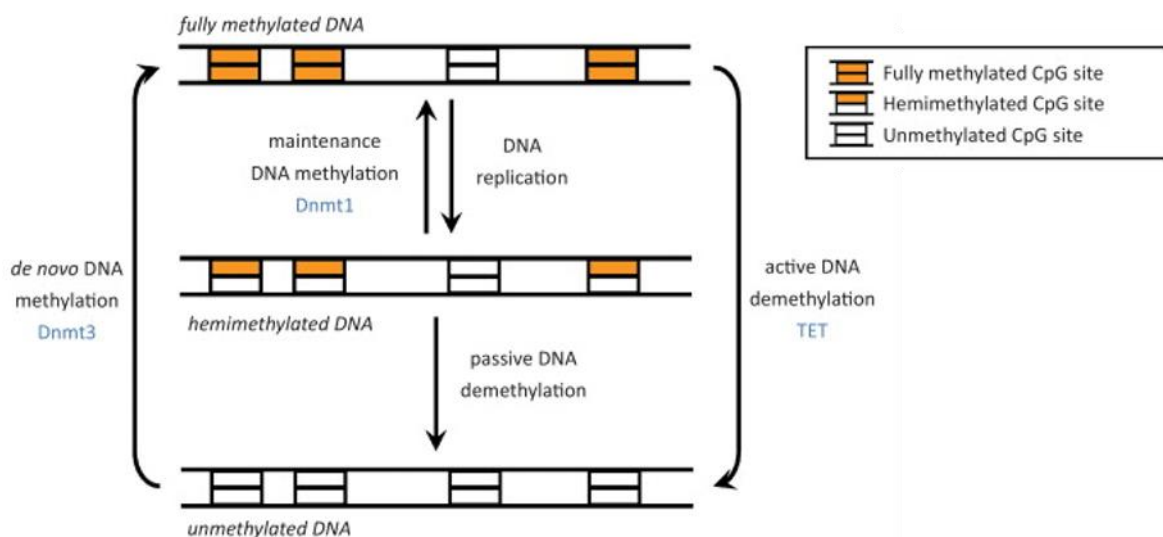


Figure 8: Classical DNA methylation maintenance model. The DNA methylation pattern can be established by the *de novo* DNA methyltransferases DNMT3A/B, leading to fully methylated DNA. After each round of DNA replication, half of the methylation is passively lost. The hemimethylated DNA double strand can then be remethylated by DNMT1 or be further diluted by DNA replication. DNA methylation can also be actively removed by enzymes of the TET family. Figure adopted from Jeltsch and Jurkowska (2014).

The so-called 'maintenance methyltransferase' DNMT1 remethylates the hemimethylated CpG sites, and thus maintains the original methylation pattern. At regions where maintenance methylation does not occur, half of the DNA methylation is passively lost as it is diluted after each round of DNA replication (Jeltsch and Jurkowska 2014; Jones and Liang 2009). Alternatively, active demethylation can be induced by enzymes of the

ten-eleven translocation (TET) family, which catalyze the oxidation of 5mC stepwise. The resulting derivatives can then be removed by thymine DNA glycosylase (TDG)-mediated base-excision repair (BER) (Schuermann et al. 2016).

However, the classical maintenance methylation model is challenged by several studies, and can be regarded as an oversimplification. It has been observed, that after DNA replication, methylation patterns are not perfectly preserved and site-specific. At a certain locus, the distribution of methylated CpG sites can vary between different cells of the same tissue, but still display a similar methylation density, indicating that DNA methylation maintenance takes place in a partially stochastic and dynamic process, which depends on the relative local enrichment and activity of DNMTs and demethylases, as well as additional properties of the respective genomic region. Moreover, the DNMT3s and DNMT1 exhibit overlapping functions. Despite being often labelled as merely a maintenance methyltransferase, *in vitro* and *in vivo* studies revealed that DNMT1 also exhibits *de novo* DNA methylation activity. On the other hand, the DNMT3 enzymes can cooperate with DNMT1 and support the remethylation of CpG-dense regions or repeats (Gujar et al. 2019; Jeltsch and Jurkowska 2014).

1.4 Mammalian DNA methyltransferases

As described in the previous chapter, three catalytically active DNMTs have been described in human cells so far, which are able to introduce DNA methylation at the C5 position of cytosine. In addition to DNMT3A and DNMT3B, a paralog within the DNMT3 family has been described, the DNA Methyltransferase 3-Like (DNMT3L) which however does not possess catalytic activity (Gowher and Jeltsch 2018). All DNMT3s and DNMT1 contain a N-terminal part which has an important regulatory function for the enzymes, and enables the association with various interaction partners (Figure 9). The C-terminal parts of the DNMTs comprise the catalytic domains, which are highly conserved with several other DNA MTases found in all kingdoms of eukaryotes and prokaryotes (Gowher and Jeltsch 2018). Based on comparative sequence analysis, ten conserved motif blocks were identified within the catalytic domain, out of which six (I, IV, VI, VIII, IX, X) are highly conserved in eukaryotic 5mC methyltransferases. The motifs occur in sequential order and vary in distance (Kumar et al. 1994). Binding of

the methyl donor *S*-adenosyl-L-methionine (AdoMet, or SAM) is promoted by the motifs I, II, III, and X, whereas motifs IV - VIII shape the catalytic pocket. Besides, motifs VIII and IX are part of the target-recognition domain (TRD) and are involved in DNA sequence recognition (Albert et al. 2018).

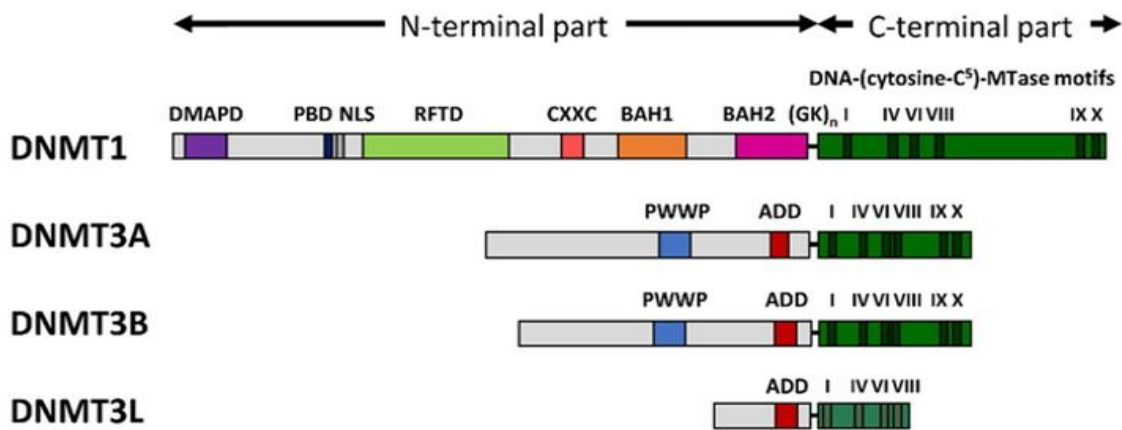


Figure 9: Domain compositions of the four mammalian DNA methyltransferases (DNMTs). All enzymes contain a N-terminal regulatory part and a catalytic domain in the C-terminal part. DMAPD, DNA methyltransferase associated protein 1 interaction domain; PCB, PCNA binding domain; NLS, nuclear localization signal; RFTD, replication-foci-targeting sequence; CXXC, cysteine-rich domain; BAH1/2, bromo-adjacent-homology domain 1/2; GK_n, glycine lysine repeats; PWWP, PWWP domain; ADD, ATRX-DNMT3-DNMT3L domain; I-X, conserved motifs in the catalytic domain. Scheme taken from Gowher and Jeltsch (2018).

DNMTs contain a characteristic core structure, which is comprised of a mixed seven-stranded β -sheet, also termed as the ‘AdoMet-dependent MTase fold’. The enzyme can access the cytosine by a specific base flipping mechanism. In this process, the hydrogen bonds to the complementary guanine are disrupted, and the base is rotated on its sugar-phosphate backbone, enabling it to enter the catalytic pocket of the MTase (Cheng and Roberts 2001). The subsequent methylation of cytosine is a multi-step process and begins with the nucleophilic attack of a SH-group of cysteine (PCQ/N; motif IV) on carbon 6 (Figure 10). This results in a covalent bond between the enzyme and the base, which is supported by the temporary protonation of the N3 atom by a glutamate (ENV; motif VI). Next, the active C5 atom attacks the methyl group of AdoMet, and after deprotonation of C5 in an uncharacterized mechanism, the previously formed covalent bond by cysteine is resolved. Thus, the aromatic ring is reformed, and the demethylated cofactor *S*-Adenosyl-L-homocysteine (AdoHcy, or SAH) remains as a product. Due to its lack of a methyl group, AdoHcy acts as an inhibitor of MTase activity (Jeltsch 2002).

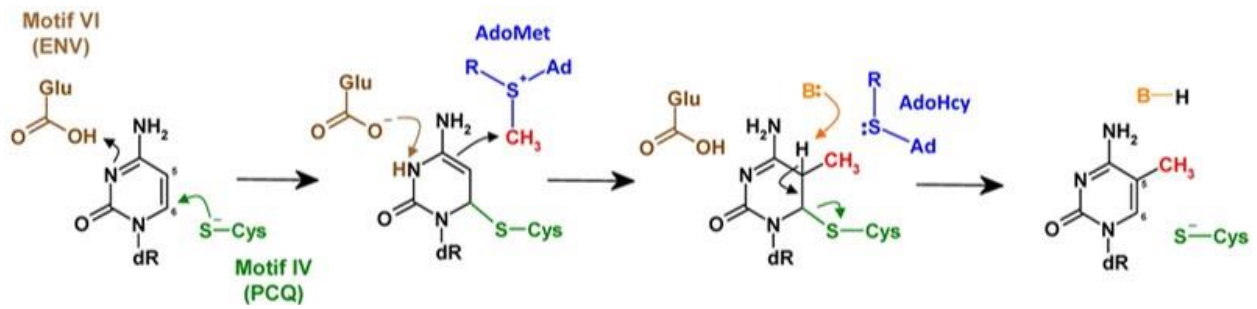


Figure 10: Catalytic mechanism for the transfer of a methyl group (red) to the C5-position of cytosine by DNA methyltransferases (DNMTs). The co-factor S-Adenosyl-L-methionine (AdoMet) serves as a methyl donor for the reaction. Two catalytic motifs (IV and VI) of the DNMTs are involved in the catalytic process. The unknown base (B) for proton abstraction is displayed in orange. Figure taken from Gowher and Jeltsch (2018).

1.4.1 DNMT1

The murine DNA (cytosine-5)-methyltransferase 1 (DNMT1) has been identified as a CpG-specific methyltransferase for hemimethylated DNA by Browne et al. in 1977, and has thus been the first MTase discovered in vertebrates (Browne et al. 1977). Studies in *Dnmt1*-deficient mouse ES cells showed that the loss of activity resulted in drastically reduced global DNA methylation levels, derepression of intracisternal A particle (IAP) retroviruses, and partially impaired XCI (Li et al. 1992; Sado et al. 2000; Walsh et al. 1998). In general, DNMT1 is required for the maintenance of methylation patterns after DNA replication, and *in vitro* studies demonstrated that DNMT1 has a 10 - 50-fold preference for hemimethylated DNA compared to unmethylated DNA. Moreover, DNMT1 activity is further stimulated by DNMT3A-mediated *de novo* methylation, and also by itself exhibits *de novo* methylation activity (Bashtrykov et al. 2012; Fatemi et al. 2002).

DNMT1 has a large N-terminal part of around one thousand amino acids, which contains multiple functional domains (Figure 9). At the very N-terminal part, the DNA methyltransferase associated protein 1 interaction domain (DMAPD) is located, which however is not present in the DNMT1o isoform (Veland and Chen 2017). The associated DMAP1 protein is a stimulator of DNMT1 activity, and interaction foremost occurs at sites of double strand break repair (Lee et al. 2010). Moreover, DNMT1 contains a cysteine-rich CXXC domain which is able to specifically bind to unmethylated CpG dinucleotides (Pradhan et al. 2008). Two bromo-adjacent-homology (BAH) domains

(BAH1 and BAH2) are positioned close to the catalytic domain of DNMT1. A recent discovery showed that BAH1 is a specific reader of H4K20me3. Hence, it can recruit DNMT1 to LINE-1 retrotransposons, and reinforce silencing by the introduction of DNA methylation (Ren et al. 2021).

As mentioned before, DNMT1 has an important role in maintaining methylation patterns during S-phase. The proliferating cell nuclear antigen (PCNA) binding domain (PBD) of DNMT1 facilitates the interaction with PCNA, which is part of the replisome. This interaction in turn leads to a higher methylation efficiency at the daughter strand (Iida et al. 2002). Another crucial interaction partner of DNMT1 for the recruitment to replication foci is Ubiquitin Like With PHD And Ring Finger Domains 1 (UHRF1). UHRF1 is a multi-domain protein, and just as DNMT1, expressed in proliferating cells. The interaction between both proteins is mediated by the ubiquitin-like (UBL) domain of UHRF1, and the auto-inhibitory RFTS domain of DNMT1. At the replication fork, UHRF1 can bind to hemimethylated DNA through its SET and RING associated (SRA) domain (Figure 11). Moreover, it recognizes the H3K9me2/3 modification with its tandem tudor domain (TTD) and H3R2me0 with its plant homeodomain (PHD) (Greenberg and Bourc'his 2019; Patnaik et al. 2018; Veland and Chen 2017). The Really Interesting New Gene (RING) domain, which has E3 ubiquitin ligase activity, ubiquitinates H3

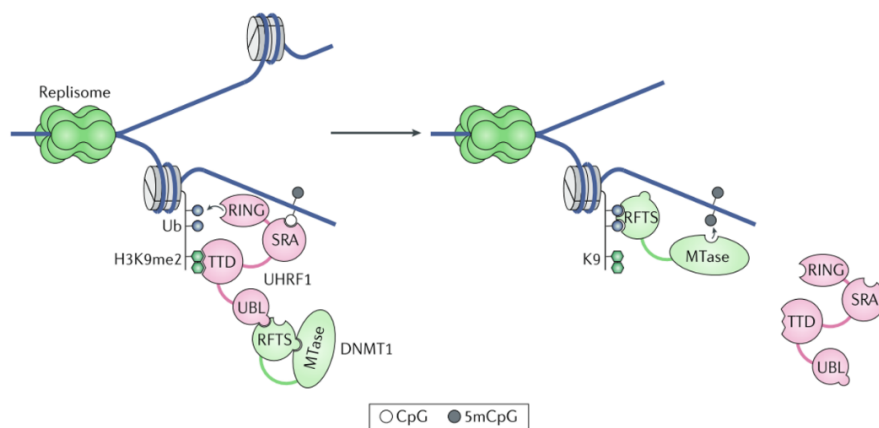


Figure 11: DNA maintenance methylation mechanism at the replication fork. The SET- and RING-associated (SRA) domain of UHRF1 binds to hemimethylated DNA (black and white lollipop), and the tandem tudor domain (TTD) to H3K9me2. The Really Interesting New Gene (RING) domain ubiquitinates the H3 tail. By the interaction of the replication-foci-targeting sequence (RFTS) of DNMT1 with the ubiquitin-like (UBL) domain of UHRF1, DNMT1 is recruited to the replication fork. The RFTS domain then binds the ubiquitinated H3 tail, which releases the inhibitory state of the catalytic domain (MTase) and enables the remethylation of the freshly synthesized DNA strand. Figure adopted from Greenberg and Bourc'his (2019).

at Lys 18 and Lys 23. The modified H3 tail is then bound by the RFTS domain of DNMT1, which concomitantly releases the catalytic domain. As a consequence, DNMT1 is activated and methylates the nearby daughter strand (Li et al. 2018).

1.4.2 DNMT3 family

Based on the classical DNA maintenance methylation model (Figure 8), the enzymes of the DNA (cytosine-5)-methyltransferase 3 (DNMT3) family are mainly responsible for the establishment of *de novo* DNA methylation patterns in the genome. Among the family members, only the paralogs DNMT3A and DNMT3B exhibit catalytic activity, whereas DNMT3L is catalytically inactive and functions as a regulator (Jeltsch and Jurkowska 2014). The amino acid sequences of mouse and human DNMT3s are highly conserved. While the human DNMT3B full-length protein has 94 % sequence identity with its murine counterpart, the identity for DNMT3A is even at 98 %. Moreover, the homology between the catalytic domains of human DNMT3A and DNMT3B is also very high (81 %) (Xie et al. 1999). In comparison, DNMT3L is truncated at the C-terminus of the catalytic domain, thus lacking the TRD with the motifs IX and X. It also contains mutations within the conserved motifs responsible for the methylation activity of 5mC methyltransferases (Albert et al. 2018; Gowher et al. 2005). Recently, an additional variant, namely DNMT3C, was discovered in rodents, which represents an evolutionary recent duplication of the *Dnmt3b* gene (Barau et al. 2016).

The N-terminal regulatory part of DNMT3 enzymes

All DNMT3s contain an ATRX-DNMT3-DNMT3L (ADD) domain at the N-terminal part of the protein in common (Figure 9), which itself separates into a GATA-like zinc finger, a PHD finger, and a C-terminal α -helix. The ADD domain inhibits the activity of the catalytic domain through an allosteric mechanism. It has a strong binding preference for H3 tails unmethylated at Lys4 (H3K4me0). When ADD binds to H3K4me0, the auto-inhibitory conformation of the enzyme is abrogated, which however does not apply for H3K4me3 (Jeltsch and Jurkowska 2016; Veland and Chen 2017). In addition, the auto-inhibitory mode of DNMT3A can be further strengthened by association of MeCP2 to

the ADD domain (Rajavelu et al. 2018). Mutagenesis of the DNMT3A-ADD binding pocket lead to dysregulated binding of the enzyme, and resulted in aberrant deposition of DNA methylation at H3K4me2/3-marked, active promoters in mouse ESCs (Noh et al. 2015).

DNMT3A and DNMT3B also contain a PWWP domain (Figure 9), which got its name from a typical short proline/tryptophane motif. It is folded into a β -barrel composed of five anti-parallel β -strands, followed by a helix bundle of five α -helices. In general, PWWP domains can bind methylated lysines through a conserved aromatic cage, formed by three aromatic residues (Qin and Min 2014). It was shown that the PWWP domain binds histone tails trimethylated at Lys36 (H3K36me3), a modification usually located at the gene bodies of active genes (Dhayalan et al. 2010). However, a recent study revealed that the PWWP domain of DNMT3A not only has the same binding affinity for H3K36me2, it also binds to DNA through its basic surface patch with a slight preference for AT-rich sequences (Dukatz et al. 2019). Moreover, the mutation of a tryptophane within the aromatic cage resulted in impaired binding to H3K36me2/3, and aberrant methylation of polycomb-marked regions, which in turn phenotypically manifested as microcephalic dwarfism in mice (Heyn et al. 2019).

For the very N-terminal part of DNMT3A and DNMT3B, the exact function has not been fully resolved. Yet, it could be shown that this part of the enzyme promotes the anchoring to DNA or nucleosomes. Mutations of certain basic residues, as well as isoforms, and mutants lacking the N-terminal part of the protein displayed reduced binding affinity and methylation activity (Jeong et al. 2009; Suetake et al. 2011).

Multimerization and subnuclear localization of DNMT3 enzymes

Although DNMT3L does not exhibit catalytic activity on its own, it has a critical role in stimulating DNMT3A and DNMT3B activity for efficient silencing of retrotransposon, or establishing genomic imprints during gametogenesis that has been well documented *in vivo* (Bourc'his et al. 2001; Bourc'his and Bestor 2004; Hata et al. 2002). This observation was supported by *in vitro* studies, which reported a 15 to 20-fold increased methylation activity of the catalytic domain of DNMT3A (3AC) and DNMT3B (3BC). Moreover, a 20-fold accelerated DNA binding of 3AC and 3BC was observed in the

presence of the C-terminal part of DNMT3L (3LC) (Gowher et al. 2005; Kareta et al. 2006).

Crystallography studies of DNMT3 enzymes revealed that two monomers of 3LC can form a tetrameric, butterfly-shaped complex with two monomers of either 3AC or 3BC (Figure 12). The heterotetramer assembles in the order of 3LC-3AC-3AC-3LC (or 3LC-3BC-3BC-3LC), and the 3AC dimer contacts the negatively charged phosphate backbone through its basic surface patch. When bound to the DNA substrate, the active sites of the 3AC dimer are separated by 14 bp, and the flipped-out cytosines are thereby located on opposing strands. Between both 3AC monomers, a polar Arg/Asp ('RD') interface is formed, and strengthened by further polar interactions. 3LC binds to 3AC through a hydrophobic interface mediated by, *inter alia*, two phenylalanine ('FF') residues. The improved binding to DNA and the higher methylation activity in presence of 3LC are facilitated by the stabilization of the flexible active-site loop of 3AC. Consequently, the loop gets into close contact with the minor groove of the DNA, and the Val716 residue enters the cavity of the flipped-out cytosine. At the same time, the association to DNA is also induced by the TRD loop, which only becomes ordered in DNA-bound state, and is stabilized by hydrogen bonding with the Arg882 residue (Jia et al. 2007; Lin et al. 2020; Zhang et al. 2018).

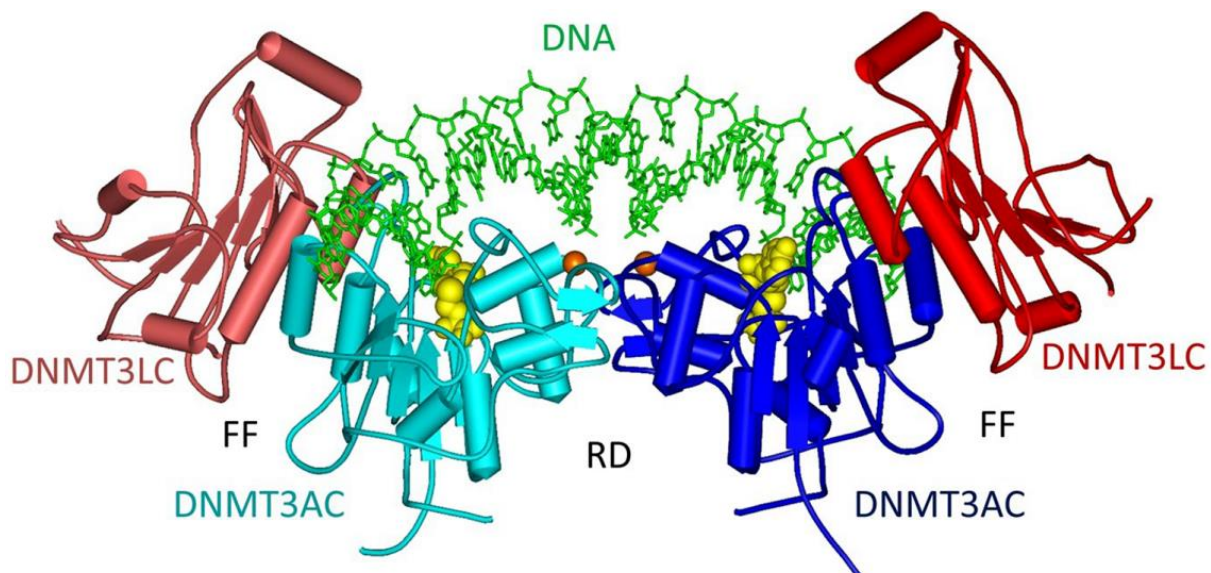


Figure 12: Crystal structure of a DNMT3A/DNMT3L heterotetramer bound to DNA. Two catalytic domains of DNMT3A (DNMT3AC) form a homodimer at the Arg/Asp (RD) interface, and each DNMT3AC forms a heterodimer with the C-terminal part of DNMT3L (DNMT3LC) at the hydrophobic Phe/Phe (FF) interface. The cofactor AdoMet is displayed in yellow. R882 is highlighted as two orange balls. The figure is taken from Emperle et al. (2018).

Despite the strong homology between 3AC and 3BC, there is a patch of sequence variability between both enzymes at the TRD loop. Due to these structural differences, 3AC and 3BC exhibit a diverse flanking sequence preference at the +1 flank position of the methylated CpG dinucleotide. While 3AC preferentially methylates the CpG(C/T) motif, 3BC favors the CpG(G/A) motif instead. Thus, the 3BC motif contains parts of the satellite II (Sat II) repeat sequence (5'-ATTTCGATG-3'), which is more readily methylated by DNMT3B in mice. (Emperle et al. 2019; Gao et al. 2020).

After the binding of a 3AC/3LC heterotetramer to DNA, additional tetramers can append to the protein-DNA complex cooperatively, and create a nucleoprotein filament (Figure 13A) which exhibits high methylation efficiency (Jurkowska et al. 2008). The oligomerization can only take place in one dimension, as 3LC is lacking the residues required for an interaction at the RD interface. However, 3AC is capable to assemble into homotetramers at its RD and FF interfaces (Norvil et al. 2018). Moreover, 3AC can form large multi-protein complexes, and thereby polymerize in two dimensions. First, it can horizontally polymerize on the same DNA strand like the 3AC/3LC heterotetramer, and second, it can polymerize vertically and bind to an additional DNA strand in parallel orientation (Figure 13B) (Jurkowska et al. 2011).

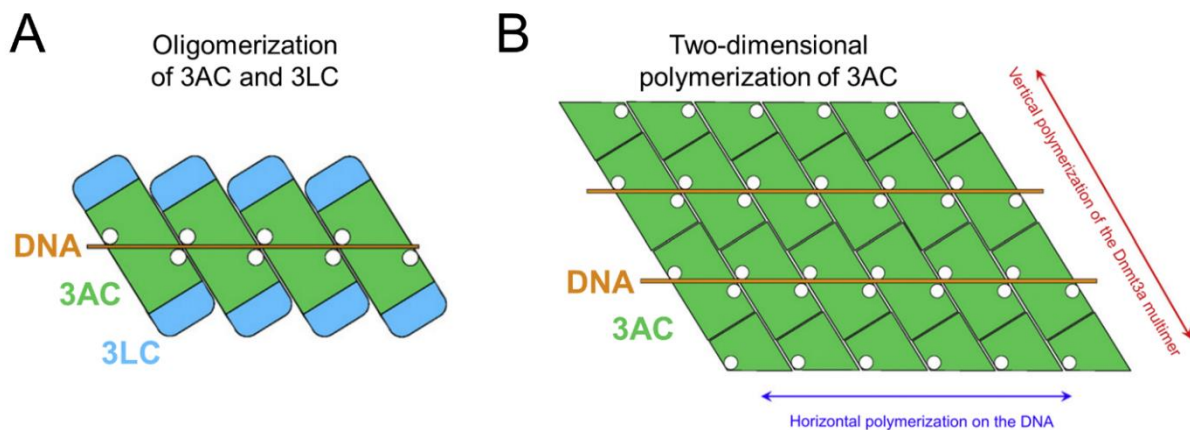


Figure 13: Multimerization of DNMT3 enzymes on DNA. A) The catalytic domains of DNMT3A (3AC) and DNMT3L (3LC) assemble into a heterotetramer, which can multimerize with additional heterotetramers on the DNA and form a one-dimensional filament. B) In the absence of 3LC, 3AC can polymerize in two dimensions through its RD and FF interfaces, and further associate to a parallel DNA strand. Both figures A) and B) were adopted from Jeltsch and Jurkowska (2013).

The mode of polymerization has direct implications on the subnuclear localization of the protein. Presumably, DNMT3A multimers are preferentially localized at heterochromatic regions, and binding of DNMT3L relocates DNMT3A to euchromatic loci, e.g.

during the process of genomic imprinting (Jeltsch and Jurkowska 2013). Moreover, the enrichment at specific heterochromatic regions (e.g. pericentromeric repeats) was also described for DNMT3B in mouse ES cells. Mutations within the catalytic site of DNMT3B lead to the hypomethylation of pericentromeric satellite repeats, which is the same phenotype as observed in the Immunodeficiency, Centromeric instability and Facial anomalies (ICF) syndrome in humans, and associated to centromeric heterochromatin instability (Okano et al. 1999).

1.5 Active demethylation by TET family members

As already described in the classical DNA maintenance methylation model, DNA methylation can be either passively lost by a lack of DNMT1-mediated maintenance methylation after replication, or by active demethylation induced by specific enzymes. The existence of a 5mC derivative named 5-Hydroxymethylcytosine (5hmC) was first detected in animals in 1972 by Penn et al. (1972). It took until 2009 to discover the three members of the Ten-Eleven Translocation (TET) enzyme family (TET1, TET2, and TET3) in mammals responsible for the oxidation of 5mC to 5hmC (Tahiliani et al. 2009).

The three TET enzymes are alpha-ketoglutarate (α KG)-dependent dioxygenases, and have a C-terminal core catalytic domain in common (Figure 14). The catalytic domain is composed of a cysteine (Cys)-rich domain, and a double-stranded β -helix (DSBH) domain where the binding of both cofactors, α KG and Fe(II), occurs. Moreover, the DSBH domain includes a low complexity insert, which is suspected to have a regulatory function, and contains residues susceptible for PTMs. At the N-terminal part of TET1 and TET3, a CXXC domain is present. The role of this domain is in the binding to DNA containing unmethylated CpGs, which consequently has implications for the genomic localization of the enzymes at unmethylated CGIs and enhancers. In contrast, TET2 lost its CXXC domain during evolution, and the ancestral domain is coded as a separate gene called Inhibition Of The Dvl And Axin Complex (*IDAX*). Presumably, the lack of the CXXC domain results in a partially differential localization of TET2 to gene bodies or enhancers (An et al. 2017; Ravichandran et al. 2018).

The TET enzymes are differentially expressed during mammalian development. TET3 contributes to DNA methylation reprogramming after fertilization by inducing global demethylation of the paternal genome. In contrast, TET1 and TET2 both participate in the re-establishment of germ cell identity of primordial germ cells during gametogenesis (Greenberg and Bourc'his 2019). Analysis of cell type-specific expression of TET enzymes in various human adult tissues revealed that TET3 was absent in most of these tissues. However, TET2 was expressed in most of the analyzed tissues, and TET1 displayed a variable expression pattern (Lorsbach et al. 2003).

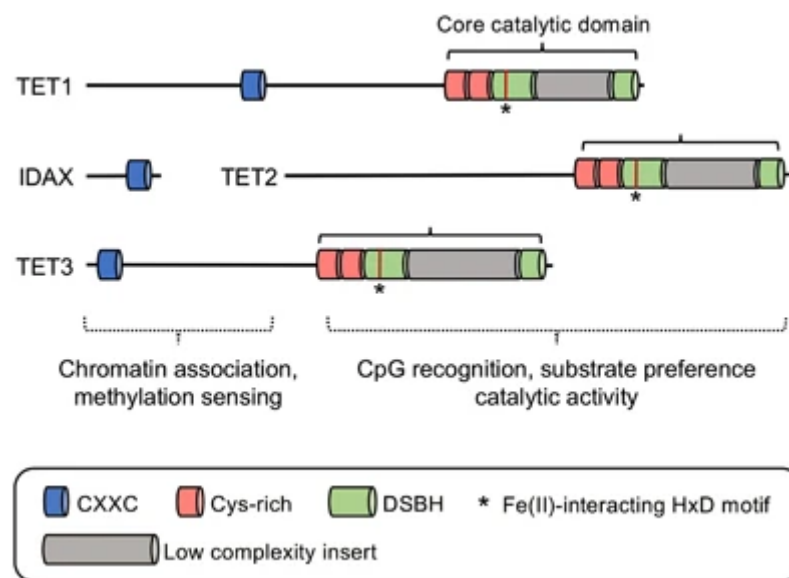


Figure 14: Domain structures of the mammalian TET1-3 enzymes and the TET2-associated IDAX protein. The TET enzymes contain a N-terminal core catalytic domain. CXXC, cysteine (Cys)-rich domain; DSBH, double-stranded β -helix. The figure is taken from An et al. (2017).

The TET enzymes are differentially expressed during mammalian development. TET3 contributes to DNA methylation reprogramming after fertilization by inducing global demethylation of the paternal genome. In contrast, TET1 and TET2 both participate in the re-establishment of germ cell identity of primordial germ cells during gametogenesis (Greenberg and Bourc'his 2019). Analysis of cell type-specific expression of TET enzymes in various human adult tissues revealed that TET3 was absent in most of these tissues. However, TET2 was expressed in most of the analyzed tissues, and TET1 displayed a variable expression pattern (Lorsbach et al. 2003).

In 2011, 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) were identified as two additional derivatives of cytosine methylation, emerging from TET-dependent demethylation (Ito et al. 2011). 5mC is first converted to 5hmC, which is followed by further oxidation to 5fC and 5caC (Figure 15). The oxidation of a modified cytosine is achieved in a stepwise manner. It is suspected that the enzyme is stalled at the intermediates and may even release its substrate temporarily. Analogously to the DNMTs, TETs use a base-flipping mechanism, in which the modified base is inserted into the catalytic pocket of the enzyme. At the catalytic site, Fe(II) is transiently oxidized by O₂, facilitating the oxidation of modified cytosine by oxidative decarboxylation of αKG (An et al. 2017).

After the conversion to 5fC or 5caC, the thymine DNA glycosylase (TDG) enzyme specifically recognizes and excises the modified base. This process leaves behind an abasic site, and induces the base-excision repair (BER) pathway, which reintroduces an unmodified cytosine (Figure 15) (Ravichandran et al. 2018). The involvement of TDG in the removal of oxidized 5mC derivatives was demonstrated by knockdown experiments, showing a 6-fold enrichment of 5fC in the genome of mouse ES cells after TDG depletion (Raiber et al. 2012).

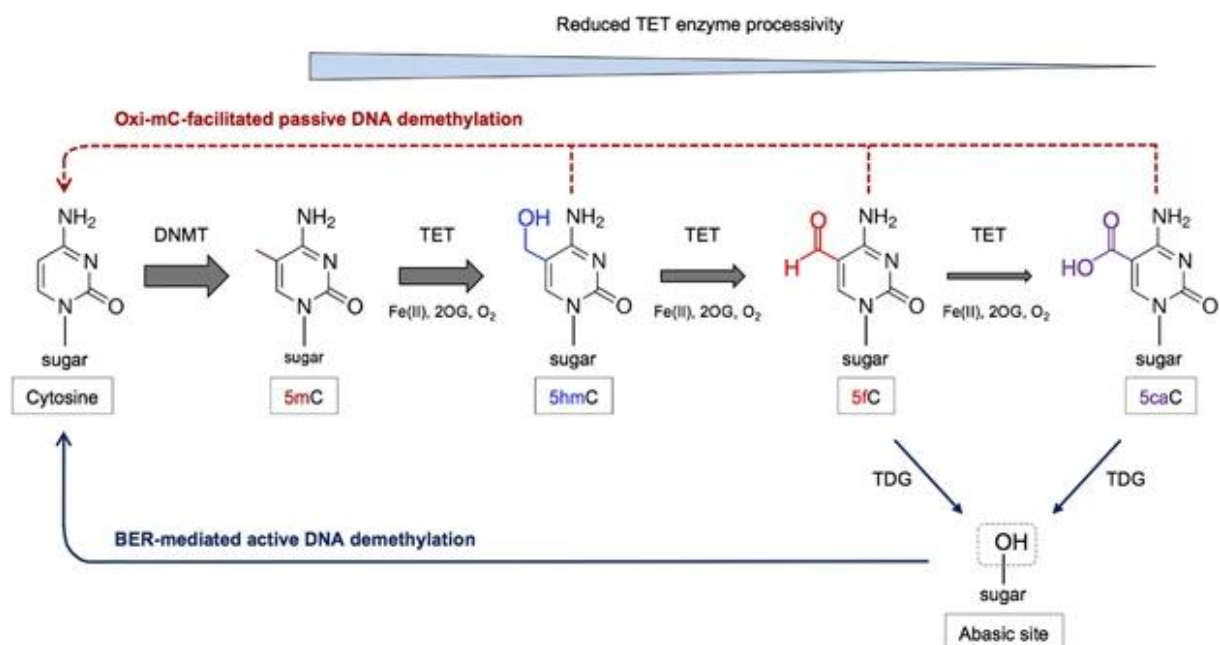


Figure 15: Active demethylation of 5-methylcytosine (5mC) in a stepwise process. TET enzymes first oxidize 5mC to 5-hydroxymethylcytosine (5hmC), then to 5-formylcytosine and 5-carboxylcytosine (5caC). For the reactions, TET requires Fe(II), 2-oxoglutarate (2OG) and molecular oxygen (O₂) as co-factors. The processivity decreases with each oxidation step. The 5fC- or 5caC bases are then excised by the Thymine-DNA glycosylase (TDG), and the resulting abasic site is exchanged to cytosine by the Base excision repair (BER) pathway. The scheme was taken from An et al. (2017).

The relative amount of 5hmC compared to guanine was measured to be in the range of 0.03 - 0.7 % in different mammalian tissues, and hence it is less abundant than 5mC (Globisch et al. 2010). 5hmC is enriched at enhancers or gene bodies of tissue-specific genes, and generally associated with gene activation. It is also found at TSSs of CpG-rich or bivalent promoters, and coincides with the binding of TET1 (Cui et al. 2020; Pastor et al. 2011; Williams et al. 2011). Several studies showed that the relative levels of 5fC and 5caC were considerably lower than 5hmC in various cell types. The numbers ranged from 5 - 15 ppm for 5fC, or 3 - 7 ppm for 5caC, which is up to three orders of magnitude less than 5hmC (Iurlaro et al. 2016; Mulholland et al. 2020; Zhang et al. 2016). Both modifications, 5fC and 5caC, were found at active promoters and enhancers in mouse ES cells and co-localized with TET1 and TET2 (Neri et al. 2015). The presence of 5hmC in differentiated tissues highlights the ambivalent characteristics of this modification. On the one hand, it is a transient intermediate of the DNA demethylation cascade. On the other hand, it can stably demarcate distinct genomic elements, and serve as an epigenetic mark which is specifically recognized by 5hmC readers, such as MeCP2 or UHRF2 (Hahn et al. 2014).

1.6 Post-translational modifications at histone tails and their connection to DNA methylation

Histone post-translational are one major group of epigenetic signals that regulate the compaction of chromatin, and thus control the DNA accessibility for proteins involved in activating gene expression. The majority of modifications mostly occur at the unstructured N-terminal tails of the histone core proteins (Bannister and Kouzarides 2011). More than 500 different PTMs of at have been discovered so far, and distribute over a variety of residues. They include methylation (Lys, Arg), acetylation (Lys, Ser, Thr), phosphorylation (Ser, Thr, Tyr, Lys), ubiquitination (Lys), citrullination (Lys) or sumoylation (Lys) (Figure 16A) (Huang et al. 2015).

The probably best-studied modifications are methylation and acetylation. Lysines can exist in four different methylation states, which are either 'unmethylated' (Kme0), mono-/di- or trimethylated (Kme1/2/3) (Figure 16B). In unmethylated state, the ϵ -nitrogen of lysine carries three hydrogen atoms and a positive charge. In methylated state,

the hydrogen atoms are substituted by methyl groups (-CH₃). With each addition of a methyl group, the size and hydrophobicity of the lysine increases, whereas its potential to form H-bonds decreases. The charge of +1, however, remains constant (Luo 2018; Taverna et al. 2007). Acetylation of the ε-nitrogen of lysine results in the neutralization of its positive charge, and hence can interfere with the formation salt bridges, for example in DNA interaction (Christensen et al. 2019).

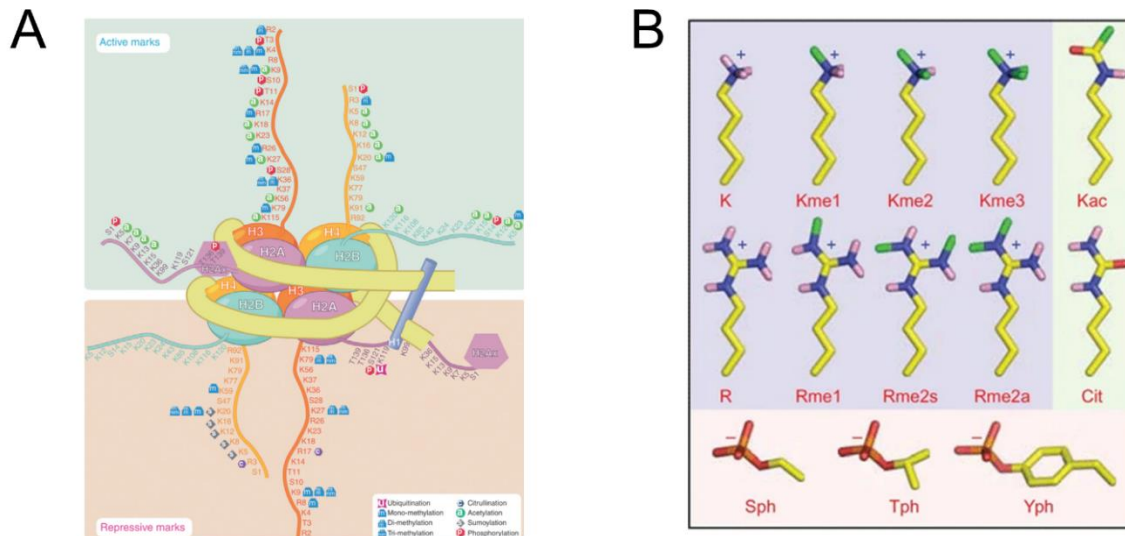


Figure 16: Post-translational modification (PTM) of histones defining the local chromatin states. A) Lysines (K), arginines (R), serines (S) or threonines (T) on histones tails are target for various PTMs. The modifications are either associated with active or repressive chromatin states. Figure taken from Sawan and Herceg (2010). B) Different modification states of lysine (K), arginine (R), serine (S), threonine (T), and tyrosine (Y) residues. me1/2/3, mono-/di-/trimethylation; ac, acetylation; me2s, symmetric dimethylation; me2a, asymmetric dimethylation; Cit, citrullination; ph, phosphorylation. Figure adopted from Ta-

The histone PTMs are introduced by so-called ‘writers’, which are enzymes capable to transfer a chemical group to a specific residue. Two examples of this are histone lysine methyltransferases (HKMTs) or histone acetyl transferases (HATs). Although the PTMs are covalently attached, they can be actively removed by ‘erasers’, for instance histone lysine demethylases (KDMs) or histone deacetylases (HDACs). The readout of the histone PTMs is performed by ‘readers’, which can specifically recognize a single mark or even a combination of multiple marks, and then induce downstream processes (Podobinska et al. 2017; Ueda and Seki 2020).

In this chapter, a selection of five histone PTMs (H3K4me3, H3K9me3, H3K27me3, H3K27ac, and H3K36me3) is described in more detail, as their dynamics were investigated in the context of *de novo* DNA methylation in this PhD work. Thereby, the focus is set on the occurrence of these marks, their regulator function, and the enzymes

involved in their control. Finally, their interplay with DNA methylation is elucidated, in order to gain an insight into the complex regulatory mechanisms of these epigenetic marks.

1.6.1 Histone 3 lysine 4 trimethylation

The trimethylation of histone H3 at lysine 4 (H3K4me3) was first described as a mark associated with gene activation in the transcriptionally active macronuclei of *Tetrahymena*, and its presence could also be experimentally proven in the human HeLa cell line (Strahl et al. 1999). At the promoters of active genes, H3K4me3 accumulates in a defined pattern around the TSS. The DNA region at the positions -200 to +50 relative to the TSS is generally depleted in nucleosomes, and it is highly accessible at these sites. H3K4me3 is strongly enriched at the nucleosome next to this gap, where the H3K4me3 peak downstream of the TSS is more prominent. At these sites, H3K4me3 is flanked by H3K4me2 (-500 and +700) and H3K4me1 (-900 and +1000). The height and width of H3K4me1/2/3 peaks are closely correlated with the levels of gene transcription (Barski et al. 2007; Heintzman et al. 2007; Howe et al. 2017; Thurman et al. 2012). H3K4me3 can also co-occur with the silencing mark H3K27me3 and together form bivalent domains at specific developmental genes, which will be elucidated in more detail in the chapter 'Histone 3 lysine 27 trimethylation' (Bernstein et al. 2006).

The TATA-Box Binding Protein Associated Factor 3 (TAF3) is a reader of H3K4me3. It recruits the general initiation factor TFIID multi-protein complex to the core promoter, and thus induces the formation of the transcription preinitiation complex (PIC) (Laubert et al. 2013). This mechanism provides a direct link between H3K4me3 and gene activation, and is supported by the observation that the occupancy of H3K4me3 is correlated with RNA polymerase II (RNAPII) at active promoters (Figure 17A) (Heintzman et al. 2007).

To date, at least six different histone methyltransferase (HMTases) have been identified in humans which are capable to methylate H3K4 with their AdoMet-dependent SET domain. These enzymes are mixed-lineage leukemia 1 – 4 (MLL1-4, also KMT2A-D), as well as the SET domain-containing protein 1A (SETD1A, also KMT2E) and 1B (SETD1B, also KMT2F). They are all parts of multi-protein complexes which belong to

the family of so-called COMPASS-like complexes. These complexes in turn are subdivided into a WRAD complex and contain additional variable subunits (Figure 17B). SETD1A/B-COMPASS-like complexes are solely responsible for the trimethylation of H3K4, whereas MLL1/2-COMPASS-like complexes are restricted to mono- or dimethylation, and MLL3/4-COMPASS-like complexes to monomethylation. The methyl groups on Lys4 can be dynamically removed by specific demethylase enzymes, of which seven have been described in humans. Of these, the four KDM5A-D family members and nuclear protein 66 (NO66) are capable to demethylate H3K4me3 and H3K4me2 with their Jumonji C (JmjC) domain, using Fe(II) and α KG as co-factors. (Collins et al. 2019; Hyun et al. 2017; Shilatifard 2012).

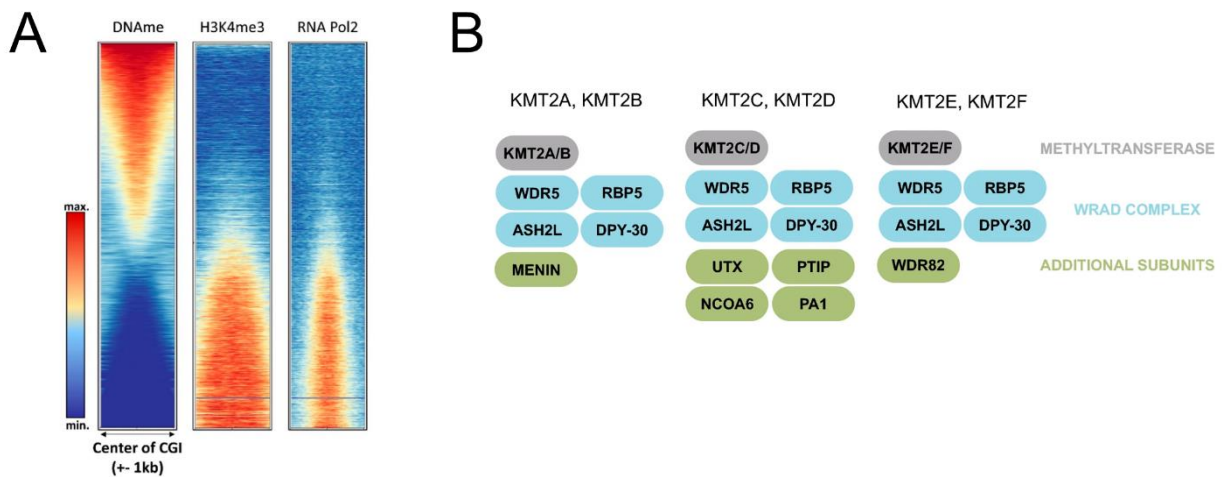


Figure 17: A) Heatmaps showing the signals of DNA methylation (DNAm), H3K4me3, and RNA Polymerase II (RNA Pol2) at CpG islands (CGI) in HEK293 cells, sorted by DNAm signal intensity. DNAm enrichment is strongly anticorrelated with H3K4me3 and RNA Pol2 occupancy. Adopted from Jeltsch et al. (2019). B) Composition of COMPASS-like complexes. All complexes contain a WRAD complex, and differ by the histone methyltransferase subunit, and by additional subunits. Figure taken from (Collins et al. 2019).

In general, H3K4me3 and DNA methylation are strongly anti-correlated (Figure 17A), and loss of H3K4me3 at certain genomic loci was associated with the concurrent gain of DNA methylation at the same sites in mouse ES cells (Meissner et al. 2008). A similar observation was made in the human HepG2 cell line, where H3K4me3 and DNA methylation were mutually exclusive at gene promoters (Zhang et al. 2009). As previously described in the chapter 'DNMT3 family', the DNA methyltransferases of the DNMT3 family contain an ADD domain, which autoinhibits the activity of the catalytic domain. While binding of ADD domain to H3K4me0 induces a conformational change

and activates the enzyme, trimethylated H3K4 cannot be bound. This mechanism protects H3K4me₃-marked active promoters against aberrant DNA methylation by DNMT3 enzymes (Jeltsch and Jurkowska 2016).

Vice versa, the presence of DNA methylation can also affect the deposition of H3K4 methylation. A local intermediate DNA methylation density allows the enrichment of the enhancer mark H3K4me₁, while regions with low DNA methylation are more susceptible for trimethylation of H3K4. This discriminative effect between methylation states can be explained by the lack of a CXXC domain in the monomethyltransferases MLL3 and MLL4, responsible for the recognition of unmethylated CpG sites. In contrast, a CXXC domain is present in MLL1 and MLL2, whereas SETD1A and SETD1B interact with the CXXC domain-containing protein CFP1. Hence, the CXXC domain can guide the COMPASS complexes to loci with low DNA methylation and facilitate the introduction of higher H3K4 methylation states (Hashimoto et al. 2010; Sharifi-Zarchi et al. 2017).

1.6.2 Histone 3 lysine 9 trimethylation

The trimethylation of histone 3 at Lys9 (H3K9me₃) is a well-studied histone PTM closely connected with the formation of constitutive heterochromatin. This modification is strongly enriched in late-replicating pericentric heterochromatin and in repetitive sequences. Endogenous retroviruses (ERVs) and LINEs are marked with H3K9me₃, which retains them in a repressed state, and thus preserves genomic stability. H3K9me₃ is often found at silent gene promoters, and principally associated with gene repression (Barski et al. 2007; Ernst et al. 2011; Rice et al. 2003; Wu et al. 2005). Moreover, it was shown that H3K9me₃-marked chromatin controls telomere length, and is required for accurate mitotic segregation (García-Cao et al. 2004; Martins et al. 2020).

Various HMTases have been described in mammalian cells, which include Suppressor of variegation 3-9 homolog 1 (SUV39H1) and 2 (SUV39H2), Euchromatic histone-lysine N-methyltransferase 2 (EHMT2, also G9a), G9a-like protein (GLP), SET domain bifurcated 1 (SETDB1) and 2 (SETDB2), and members of the PR/SET Domain (PRDM) family. Besides the PRDMs, all of these proteins are SET domain-containing enzymes,

and use AdoMet as cofactor. *In vivo*, G9a and GLP together form heterodimers, and are able to mono- and dimethylate H3K9 at euchromatic regions. At promoters, G9a/GLP can oligomerize, and hence spread the H3K9me_{1/2} mark over multiple nucleosomes. SETDB1 has the capability to introduce all three methylation states at H3K9. Likewise, SUV39H1/2 can mono-/di-/trimethylate H3K9 *in vitro*. However, at pericentromeric regions, SUV39H1/2 uses H3K9me₁, as a substrate to di-/trimethylate H3K9, which has to be established by SETDB1 beforehand (Hyun et al. 2017; Monaghan et al. 2019; Zhu et al. 2020). Another study showed that PRDM3 and PRDM16 also contribute to the monomethylation of H3K9 (Pinheiro et al. 2012).

SETDB1 can form a complex with a KRAB domain-containing zinc-finger proteins (KRAB-ZFP) and KRAB domain-associated protein 1 (KAP-1), and is recruited to specific genomic loci where it introduces H3K9 trimethylation (Figure 18A). Alternatively, MBD1 can direct SETDB1 to regions which contain DNA methylation (Figure 18B) (Zhu et al. 2020). In a different mechanism, SUV39H1/2 deposits H3K9me₃ at loci, and then reads its newly established mark via its chromodomain. Subsequently, it continues to methylate the adjacent nucleosome, leading to the spreading of the mark (Hyun et al. 2017).

The formation of constitutive heterochromatin is facilitated by the binding of HP1 to H3K9me₃. HP1 is composed of a chromodomain, responsible for H3K9me₃ readout, as well as a chromoshadow domain, inducing the homodimer formation and bridging between two H3K9me₃ nucleosomes. The spreading of the H3K9me₃ mark by SUV39H1/2 is followed by the continuous association of HP1, and eventually results in chromatin condensation (Figure 18C) (Al-Sady et al. 2013; Machida et al. 2018). There is increasing evidence that the accumulation of HP1 α above a certain threshold locally induces a phase separation, and the formation of liquid-like droplets at HP1 α heterochromatin (Larson and Narlikar 2018).

Several studies have shown that SETDB1 and SUV39H1/2 are capable to directly interact with HP1, but also with different DNMTs, thus providing a direct link between H3K9 methylation and DNA methylation (Fuks et al. 2003b; Li et al. 2006). However, DNA methylation set by DNMT3B as a consequence of SUV39H1/2-dependent recruitment was only observed at pericentric repeats, but not at centromeric repeats. *Vice*

versa, the knockout of DNMTs did not affect H3K9me3 levels at pericentric heterochromatin (Lehnertz et al. 2003). These observations indicate that the complex regulatory network involving various H3K9 HMTases and DNMTs may have to be revised or refined in parts.

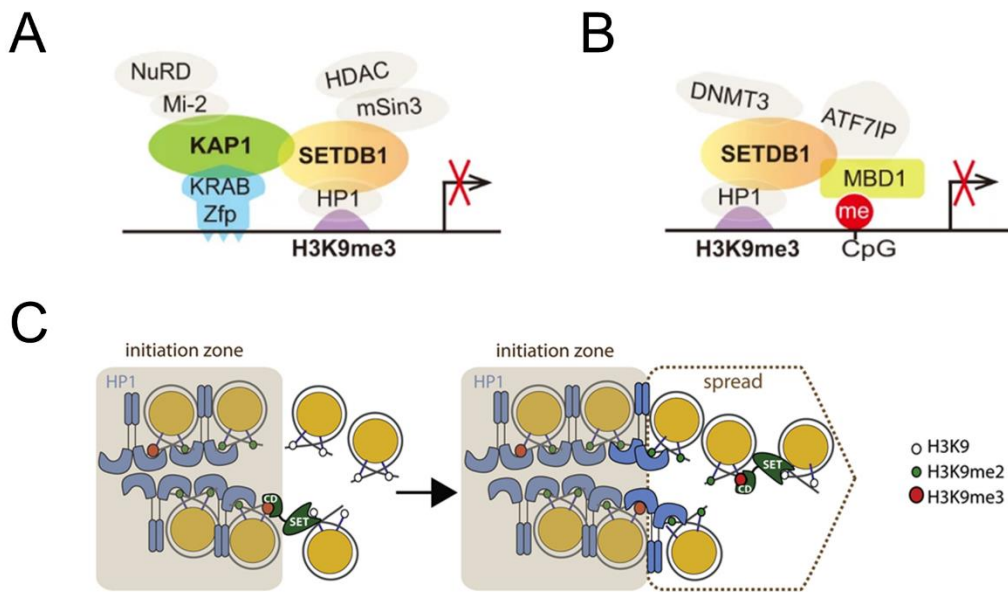


Figure 18: Epigenetic mechanisms explaining the establishment of H3K9 trimethylation at different genomic loci. A) Complex formation of SETDB1 with KAP1, and KRAB Zinc finger protein (KRAB-Zfp). KRAB-Zfp-mediated recruitment is DNA sequence-specific, and enables the introduction of H3K9 trimethylation by SETDB1. The subsequent association of HP1, HDAC/mSin3 corepressor complex or Mi-2/NuRD leads to the establishment of a repressed epigenetic state. B) MBD1 binds to methylated (me) DNA, and recruits SETDB1 for the establishment of H3K9me3, further stimulated by ATF7IP. The co-recruitment of DNMT3 enzymes or HP1 enforces the induction of a repressed chromatin state. The figures in A) and B) were taken from Zhu et al. (2020). C) Spreading of H3K9me3 by SUV39H1/2. At the edge of the initiation zone, SUV39H1/2 reads a H3K9me2/3 via its chromodomain (CD) and methylates the adjacent nucleosome. The fresh mark then serves as an anchor point for successive SUV39H1/2 binding and spreading of H3K9me3. HP1 recognizes H3K9me2/3 and induces heterochromatin. Figures adopted from Al-Sady et al. (2013).

On top of this, UHRF1 has been identified to be an additional player directly connecting H3K9me3 and DNA methylation. As described in the chapter 'DNMT1', UHRF1 is a critical interaction partner of DNMT1 for the maintenance of DNA methylation during replication. UHRF1 reads the H3K9me2/3 mark with its TTD, recruits DNMT1 to the replication fork and stimulates its catalytic activity (Figure 11) (Greenberg and Bourc'his 2019). It was shown that the knockout of UHRF1 in mouse embryos was entailed by severe hypomethylation of retrotransposons and imprinting control regions (Sharif et al. 2007). Overall, the association of UHRF1 and DNMT1 at H3K9me3-marked regions has significant implications for the (re-)establishment of global methylation patterns.

1.6.3 Histone 3 lysine 27 acetylation

The acetylation of histone 3 at Lys27 (H3K27ac) is a PTM strongly associated with gene activation, and can be foremost found at active promoters or enhancers. More precisely, H3K27ac shows the strongest correlation with gene expression at high CpG content promoters, and coincides with DNase hypersensitive sites. The highest peaks of H3K27ac are localized around 100 bp downstream, and 300 bp upstream of the TSSs (Ernst et al. 2011; Karličić et al. 2010). Besides H3K27ac, active promoters are usually decorated with additional acetylated lysine residues at histone tails, which include H2AK9ac, H2BK5ac, H3K9ac, H3K18ac, H3K27ac, H3K36ac and H4K91ac (Wang et al. 2008). Thereby, the acetylation neutralizes the positive charge of the histone tails, leading to a decreased interaction with the negatively charged phosphate backbone of the DNA, and thus to an increased chromatin accessibility (Verdone et al. 2006). *In vivo* experiments showed that H3K27ac stimulates RNAPII activity by about 50 %, which is attained by an accelerated process of transcription initiation, and a faster transition to elongation (Stasevich et al. 2014).

While H3K27ac at promoters was shown to be cell-type-invariant, its enrichment at different enhancers is very cell-type-specific among various human cell lines (Heintzman et al. 2009). The presence of H3K27ac is the defining criterion for the activity of an enhancer, whose states can be classified as either 'repressed', 'primed' or 'active' (Figure 19) (Atlasi and Stunnenberg 2017). In the process of enhancer activation, H3K27ac is established by the highly homologous co-activators P300 and CREB binding protein (CBP), which are both histone acetyltransferases (HATs) using acetyl-CoA as a cofactor. Active RNAPII was shown to be present at H3K27ac-marked enhancers, and required for the transcription of enhancer RNA (eRNA). The loss of P300/CBP-dependent H3K27ac at enhancers resulted in a reduced eRNA production (Raisner et al. 2018; Santa et al. 2010).

Enhancer priming and activation are also accompanied by the oxidation of 5mC to 5hmC, which contributes to chromatin opening, and is facilitated by TET enzymes. 5hmC levels are intermediate in primed enhancers containing only H3K4me1, and the highest in H3K4me1/H3K27ac-marked active enhancers (Mahé et al. 2017; Tsagaratou et al. 2014). TET1 was shown to be enriched at enhancers marked by

H3K4me1 (Yildirim et al. 2011). Knockout of *Tet2* in mouse ES cells lead to DNA hypermethylation of enhancers and depletion of H3K27ac. Conversely, the triple-knockout of *Dnmt1*, *Dnmt3a* and *Dnmt3b* resulted in DNA hypomethylation and gain of H3K27ac at enhancers, an effect that could be recovered after reconstitution of the DNMTs. These observations support a model in which DNA methylation acts upstream of H3K27ac in the hierarchy of enhancer activation (Hon et al. 2014; King et al. 2016).

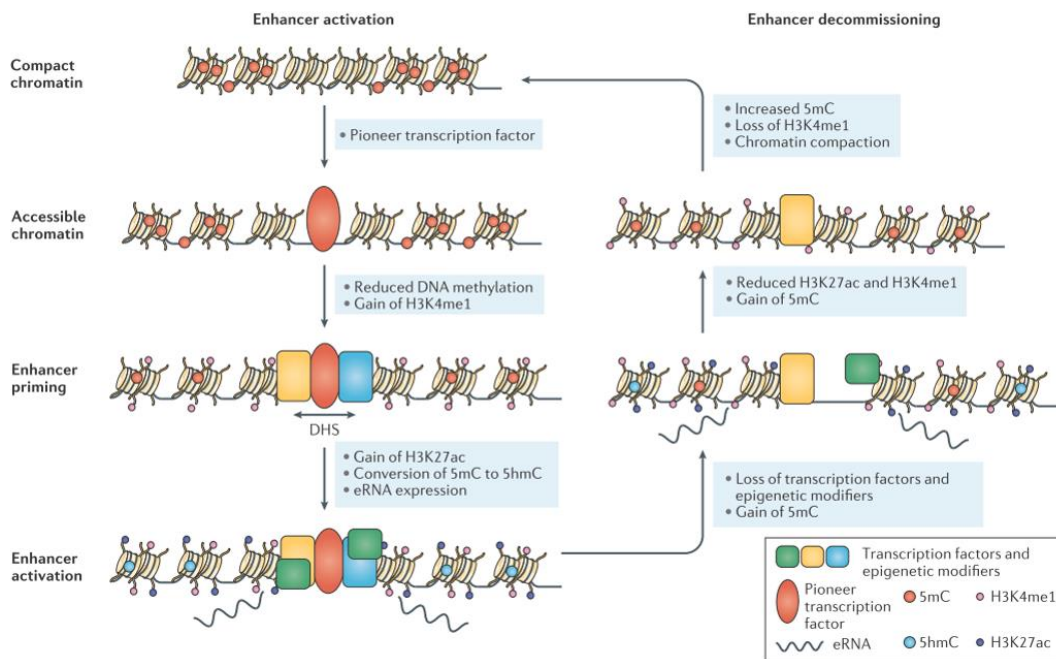


Figure 19: Schematic depiction of enhancer activation and inactivation, illustrating the central role of H3K27ac in defining enhancer activity states. Inactive enhancers are depleted in H3K27ac, and pioneer transcription factors induce the decompaction of chromatin. Enhancer priming is induced by DNA methylation (5mC) loss and H3K4me1 deposition. Acetylation of H3K27 leads to enhancer activation and enables enhancer RNA (eRNA) expression. Enhancer decommissioning is initiated by deacetylation of H3K27, H3K4me1 loss, and gain of 5mC. Taken from Atlasi and Stunnenberg (2017).

The acetylation of H3K27 is a highly dynamic process. A study applying quantitative proteomics revealed that the half-lives of this modification can actually fall below 30 minutes. This observation necessitates the involvement of specific histone deacetylases (HDACs) actively removing the acetylation mark (Weinert et al. 2018). At least eighteen HDACs (HDAC1-11 and SIRT1-7) of different classes have been identified so far. HDAC1 and HDAC2 are components of Nucleosome Remodeling and Deacetylase (NuRD) multi-protein complexes. NuRD complexes are constituted of seven core components with varying composition, and in addition, they contain the DNA methylation-sensitive binding proteins MBD2 or MBD3 (Barneda-Zahonero and Parra 2012; Basta and Rauchman 2015). It was demonstrated that MBD3 was capable to bind both,

5hmC and unmethylated cytosine, and to be localized at active regulatory elements, such as enhancers or promoters (Shimbo et al. 2013; Yildirim et al. 2011). In contrast, MBD2 specifically binds to 5mC with high affinity (Mellén et al. 2012). Hence, the NuRD complex can discriminate between different modification states of cytosine by associating with either MBD2 or MBD3, and induce decommissioning of enhancers through deacetylation of H3K27 (Basta and Rauchman 2015). The process of enhancer silencing is accompanied by the gain of 5mC by DNMTs, and demethylation of H3K4me1 by LSD1 (Figure 19) (Alajem et al. 2021; Whyte et al. 2012).

1.6.4 Histone 3 lysine 27 trimethylation

Trimethylation of H3 at Lys27 (H3K27me3) is a repressive histone PTM and constitutes a specific chromatin state, termed 'Polycomb group (PcG) chromatin' (Filion et al. 2010). PcG chromatin can be characterized as facultative heterochromatin, which can be temporally or locally decondensed, and thus represents a dynamic intermediate state between euchromatin and constitutive heterochromatin. Due to these features, PcG chromatin precisely controls the expression of key developmental genes, as for instance the homeobox (*HOX*) genes, and can form large domains that span up to more than 100 kb. In mouse embryos, the levels of H3K27 trimethylation strongly increase after embryo implantation (Atlasi and Stunnenberg 2017; Trojer and Reinberg 2007). H3K27me3 also has a repressive function in X-chromosome inactivation (XCI) (Brockdorff 2017). In addition, poised enhancers, which can be readily activated, or inactive enhancers are frequently decorated with the H3K27me3 mark (Nguyen et al. 2015). The presence of H3K27me2/3 at promoters is anti-correlated with gene expression, whereas H3K27me1 is localized around the TSS of actively transcribed genes (Barski et al. 2007).

A subset of developmental genes, mostly coding for tissue-specific transcription factors, have a specific chromatin signature at their promoters, in which H3K27me3 and H3K4me3 coincide. These so-called 'bivalent' domains hence contain a repressive and an activating mark simultaneously and represent a poised promoter state (Nguyen et al. 2015). Compared to monovalent PcG chromatin or constitutive heterochromatin, bivalent domains exhibit the lowest repressive potential (Blomen and Boonstra 2011).

At bivalent promoters, the H3K27me3 mark prevents pre-mature expression of lineage-specific genes, while the presence H3K4me3 prepares the promoter for rapid switching towards active gene expression (Azuara et al. 2006; Mikkelsen et al. 2007). Hence, the removal of either of the modifications results in lineage-commitment of the underlying gene. During differentiation, the epigenetic state can be reinforced by the acquisition of additional modifications, such as H3K9me3 or DNA methylation (Figure 20) (Margueron and Reinberg 2011). For instance, the bivalent state can be resolved by two different H3K27 demethylases, namely Ubiquitously-transcribed TPR protein on the X chromosome (UTX), or Jumonji domain-containing protein 3 (JMJD3).

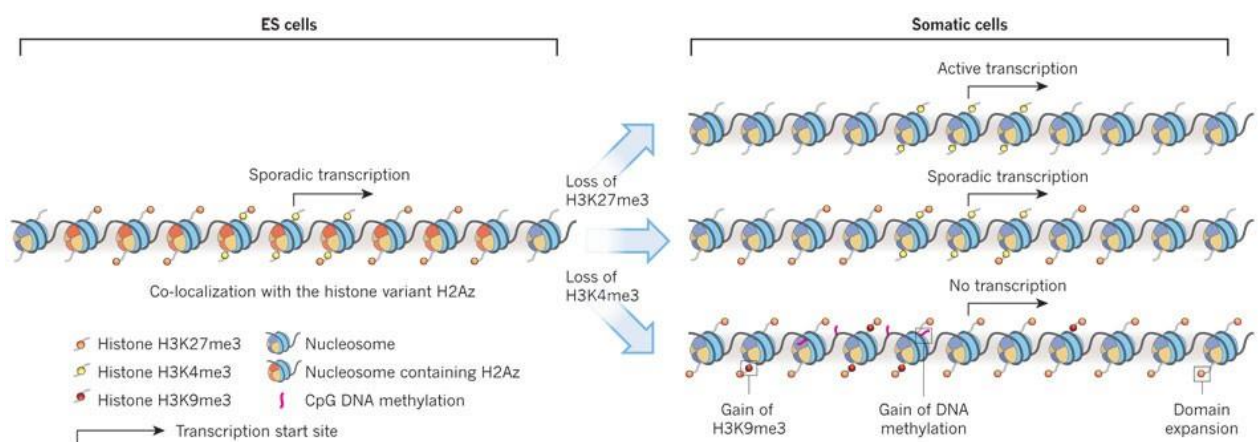


Figure 20: Resolution of the H3K4me3/H3K27me3 bivalent chromatin state of PcG target genes during development. In embryonic stem (ES) cells, bivalent domains are also enriched in H2Az, and the underlying genes are sporadically expressed. In the course of cellular differentiation, H2Az is lost, and the lineage-commitment of the genes is induced by the depletion of either the activating H3K4me3, or the inactivating H3K27me3 mark. The loss of either mark leads to the stable repression or expression of the gene in somatic cells, respectively. The repressive chromatin state is further reinforced by the acquisition of additional silencing marks such as DNA methylation or H3K9me3. The figure is taken from Margueron and Reinberg (2011) .

The deposition of all H3K27 methylation states is exclusively realized by the interplay between the two Polycomb Repressive Complexes 1 and 2 (PRC1/2). PRC2 is composed of four core subunits, which are Suppressor of zeste 12 protein homolog (SUZ12), Enhancer of zeste homolog 1 or 2 (EZH1/2), Embryonic ectoderm development (EED), and Retinoblastoma-binding protein P46/48 (RBAP46/48) (Hyun et al. 2017). SUZ12 exhibits multiple functions, as it stabilizes the complex, and enables a locus-specific recruitment by binding to ncRNA. Both, EZH1 and EZH2 are SET domain-containing HMTases, responsible for H3K27 methylation. RBAP46/48 can interact with H4, whereas EED has a critical role in the binding to H3K27me3, and stimulates the activity of PRC2. Hence, EED-mediated PRC2 recruitment to pre-existing

H3K27me3 allows the trimethylation of H3K27 at the neighboring nucleosome, and serves as a positive feedback mechanism (Figure 21).

Moreover, H3K27me3 can be read by chromobox (CBX) proteins of PRC1, which then mono-ubiquitinates histone H2A at Lys119 (H2AK119ub1) via the cooperation of the RING1A/B ubiquitin ligase, and the PCGF2/4 subunits (Nayak et al. 2011; Zhang et al. 2015). An additional PRC2 variant (PRC2.2) has been described which contains the accessory subunits Jumonji/ARID domain-containing protein 2 (JARID2) and Adipocyte enhancer-binding protein 2 (AEBP2). Recently, it was shown that JARID2 and AEBP2 physically interact with H2AK119ub1 introduced by PRC1, and thereby stimulate PRC2.2 activity (Kasinath et al. 2021). In conclusion, PRC1 and PRC2 function in a synergistic mechanism, which allows the spreading of the H3K27me3 mark over large genomic domains (Zhang et al. 2015).

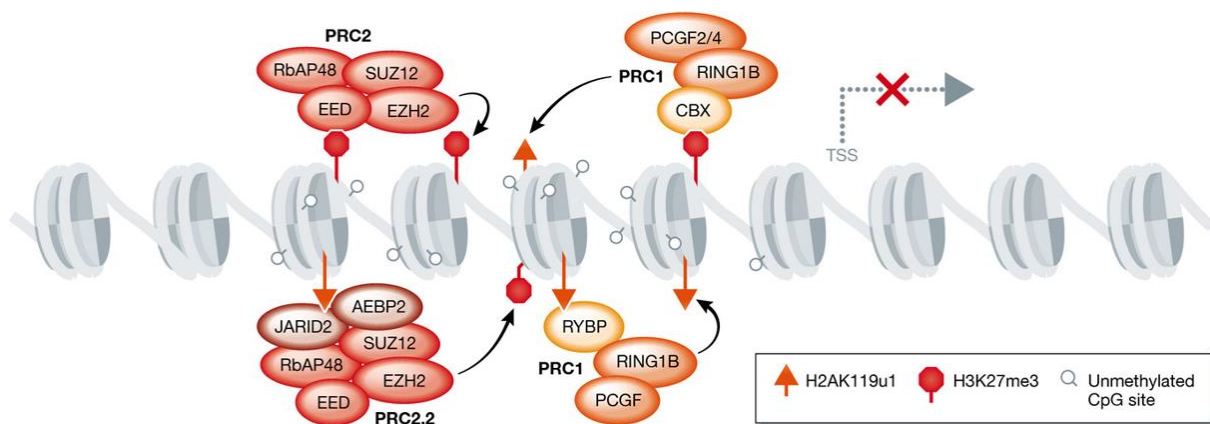


Figure 21: Schematic depiction of the interdependence of polycomb repressive complexes (PRC) for the establishment of the PcG chromatin at transcriptional start sites (TSS). The PRC2 core-subunit EED recruits the complex to a H3K27me3-marked nucleosome, enabling it to trimethylate H3K27 at the adjacent nucleosome via its EZH2 histone methyltransferase. The H3K27me3-reader CBX of PRC1 binds to the modification, and allows the RING1A/B ubiquitin ligase, and the PCGF2/4 subunits to mono-ubiquitinate H2A at Lys119 (H2AK119ub1). Spreading of H2AK119ub1 is facilitated by the association of a different PRC1 variant to the mark, and subsequent ubiquitination of the neighboring nucleosome. In addition, H2AK119ub1 is recognized by the AEBP2/JARID2 subunits of PRC2.2, thereby stimulating PRC2.2 activity and the spreading of H3K27me3. Taken from Zhang et al. (2015)

The interplay between H3K27me3 and DNA methylation is multifaceted and context-dependent. Various studies analyzing the global distribution of 5mC and H3K27me3 demonstrated that both modifications are anti-correlated, and thus may act antagonistically (Fu et al. 2020; Hagarman et al. 2013; Hon et al. 2012). One study conducting sequential ChIP-bisulfite sequencing discovered a mutually exclusive distribution of 5mC and H3K27me3 at CGIs (Brinkman et al. 2012). This was supported by the observation in hESCs, that high-CpG promoters of tissue-specific genes were enriched

in H3K27me₃, whereas medium- and low-CpG promoters were predominantly marked by DNA methylation (Xie et al. 2013). Moreover, the loss of DNA methylation at CpG-rich sequences was followed by PRC2 recruitment and the deposition of H3K27me₃ (Lynch et al. 2012). This might be explained by the association of the PRC2 subunits MTF2 and PHF1 to unmethylated, and CpG-rich sequences (Li et al. 2017). In addition, TET enzymes can directly interact with PRC2 and maintain the hypomethylated state at bivalent CGIs, resulting in the protection of H3K27me₃ against aberrant DNA methylation. At the same time, the DNMT3A1 isoform can bind the flanking shores of bivalent promoters, and thus demarcate the bivalent domain by establishing DNA methylation (Kong et al. 2016; Manzo et al. 2017; Neri et al. 2013). Further studies showed that a large fraction of DNA methylation canyons containing low levels of 5mC exhibited strong enrichment of trimethylated H3K27 (Jeong et al. 2014; Zhang et al. 2020).

Despite the strong antagonism of H3K27me₃ and DNA methylation at various genomic elements, there is still plenty of evidence that both modifications can also co-exist at the same locus and may also act synergistically. For instance, in hematopoietic cells, all regions containing H3K27me₃ displayed an average DNA methylation of around 43 % (Jeong and Goodell 2014). In HCT116 cells, high 5mC levels were also detected in non-CGI promoters, genic and intergenic regions which were simultaneously marked by H3K27me₃. However, the same study also showed that in mouse ES cells, both marks mainly co-occurred in regions with very low CpG density (Brinkman et al. 2012). Intriguingly, *Dnmt1*, *Dnmt3a* and *Dnmt3b* triple-knockout experiments in mouse ES cells with subsequent rescue of *Dnmt3a* and *Dnmt3b* demonstrated the critical role for DNA methylation for the maintenance of H3K27 trimethylation at several enhancers and promoters (King et al. 2016). Moreover, EZH2 was shown to directly interact with all DNMT enzymes, and thus recruit them to PRC2 target regions (Viré et al. 2006). Hypermethylation of bivalent promoters was observed in mice with microcephalic dwarfism, which was caused by a mutation in the PWWP domain of DNMT3A. The consequent abrogation of H3K36me_{2/3}-binding illustrated that this modification detains DNMT3A from erroneous methylation of nearby bivalent promoters (Heyn et al. 2019).

A synergistic effect of H3K27me₃ on DNA methylation could also be identified in cancer. Promoters enriched with H3K27 methylation were highly susceptible to acquire aberrant DNA methylation in human colon, breast and prostate cancer cells. Genes

with both modifications at their promoter were stronger repressed than those with individual marks only (Schlesinger et al. 2007; Takeshima et al. 2015; Widschwendter et al. 2007).

1.6.5 Histone 3 lysine 36 trimethylation

Numerous studies have illustrated the multi-faceted functions of trimethylated histone 3 Lys36 (H3K36me3) in the epigenome. In general, the enrichment of H3K36me3 within gene bodies was found to be correlated with transcript levels. ChIP experiments in human cells and yeast showed that H3K36me3 is not evenly distributed over the gene body of active genes. Instead, it is depleted at the transcriptional start sites (TSS) and elevated towards the 3' end of the genes (Figure 22A). The biased positioning of nucleosomes at exons compared to introns is accompanied by elevated H3K36me3 levels at these sites. RNAPII however shows an inverse correlation, as it accumulates at the 3' end of introns, and thus at the intron/exon boundaries (Wilhelm et al. 2011). The enrichment of H3K36me3 at exons has a role for the alternative splicing of pre-mRNA (Luco et al. 2010). Although H3K36me3 is mostly associated with active genes, its presence in constitutive and facultative heterochromatin could be detected in mouse ES cells (Chantalat et al. 2011). Moreover, the H3K36me3 modification is set at sites of DNA double-strand break (Pfister et al. 2014).

In humans, eight H3K36 HMTases have been discovered so far, of which nuclear receptor-binding SET domain (NSD) 1-3 (NSD1-3), and SET Domain Containing 2 (SETD2) are regarded as the key actors. NSD1-3 are able to mono- or dimethylate H3K36, while SETD2 is the only enzyme known to introduce all methylation states up to H3K36me3, at least *in vitro*. Albeit *in vivo* studies suggest that SETD2 requires H3K36me1/2 deposited by NSD1 and NSD2 in order to trimethylate H3K36 (Hyun et al. 2017). It was shown that SETD2 binds to the C-terminal domain of the elongating RNAPII and the splicing factor hnRNP. These interactions were critical for the deposition of H3K36me3 within gene bodies (Bhattacharya et al. 2021). The demethylase KDM5B is then recruited to the gene body, in order to remove intragenic H3K4me2/3, and thus prevent cryptic transcription (Figure 22B) (Li et al. 2019).

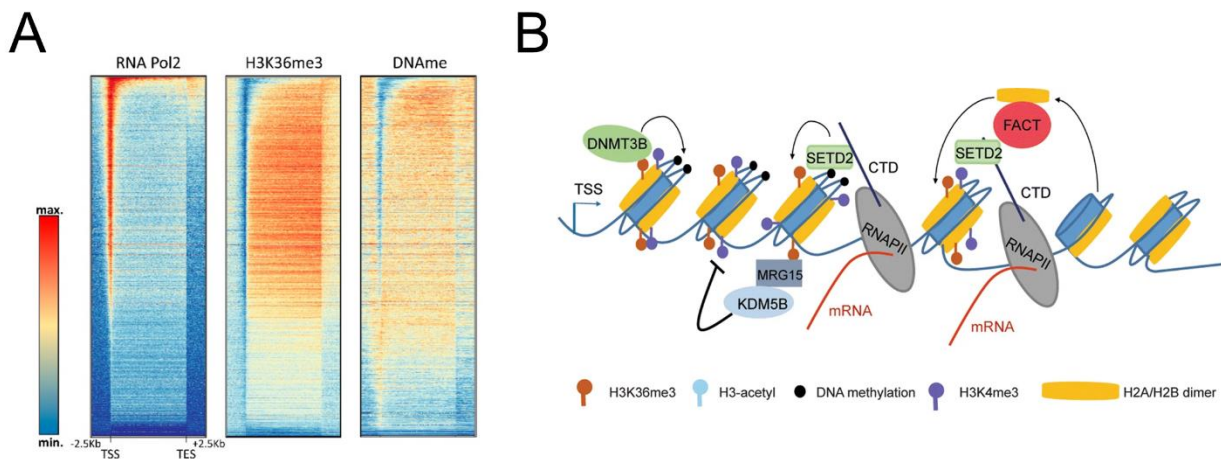


Figure 22: The deposition of H3K36 trimethylation at the gene bodies of actively transcribed genes. A) Heatmaps showing the enrichment of RNA polymerase II (RNA Pol2), H3K36me3, and DNA methylation (DNAm) within gene bodies, sorted by RNA Pol2 signals. The presence of RNA Pol2 at the transcriptional start sites (TSS) of active promoters coincides with the enrichment of H3K36me3 and DNAm in the gene bodies. TES, transcriptional end site. Modified from Jeltsch et al. (2018). B) Deposition of H3K36me3 in gene bodies through recruitment of SETD2 by the elongating RNAPII. MRG15 binds to H3K36me3, and recruits the H3K4 demethylase KDM5B to remove intragenic H3K4me3. Accumulation of spurious H3K4me3 is further prevented by FACT-mediated chromatin remodeling. DNMT3B binds to the H3K36me3 mark and introduces DNA methylation at the gene body. Figure adapted from Li et al. (2019).

Notably, DNA methylation is connected to similar genomic elements and epigenetic mechanisms as H3K36me3. For instance, the presence of DNA methylation within gene bodies is associated with active transcription of the underlying genes (Figure 22A). Moreover, DNA methylation is also involved in alternative splicing, and counteracts cryptic transcription (Jeltsch et al. 2018). As described for H3K36me3, DNA methylation was likewise found at pericentromeric heterochromatin (Bachman et al. 2001; Chantalat et al. 2011). The strong co-occurrence of both epigenetic modifications was also observed in mouse hematopoietic stem cells, where H3K36me3-marked regions exhibited a strong average CpG methylation of 94 % (Jeong and Goodell 2014).

As already mentioned in the chapter 'DNMT3 family', the DNA methyltransferases DNMT3A and DNMT3B contain a PWWP domain, which can bind to di- or trimethylated H3K36. However, different observations were made regarding the preference of DNMT3A^{PWWP} for both methylation states (Dukat et al. 2019; Sankaran et al. 2016; Weinberg et al. 2019). *In vivo* experiments with full-length constructs or just the PWWP domains revealed that DNMT3B had a strong preference for the binding to active gene bodies, and thereby closely resembled the H3K36me3 pattern. In comparison, DNMT3A binding showed a less defined distribution within gene bodies, and was strongly enriched at intergenic regions. These regions contained high levels of

H3K36me2 and ablation of NSD1/2, which introduce this mark, resulted in the redistribution of DNMT3A towards H3K36me3-enriched regions (Baubec et al. 2015; Weinberg et al. 2019).

1.7 Targeted epigenome editing

Over the past decades, a plethora of severe diseases have been described, which can be traced back to the dysregulation of epigenetic mechanisms, and result in aberrations in the epigenetic landscape. For instance, erroneous DNA methylation of promoter CGIs has been found in various types of cancers (e.g. sporadic colorectal cancer), as well as in Alzheimer's and Parkinson's disease. Moreover, allele-specific hypo- or hypermethylation at ICRs were shown to be the cause of imprinting disorders (e.g. BWS, Prader-Willi syndrome), and thereby leading to abnormal growth and mental retardation (Egger et al. 2004; Portela and Esteller 2010). Similarly, the malfunction or dysregulation of HATs (e.g. CBP in Rubinstein-Taybi-Syndrome), HDACs (e.g. HDAC2 in gastric cancer), HKMTs (e.g. MLL2 in Kabuki 1 Syndrome) or KDMs (e.g. UTX in multiple myeloma) have been associated with deviations in the global pattern of histone PTMs (Glozak and Seto 2007; Park et al. 2014; Rangasamy et al. 2013; Schulz et al. 2019). These deviations from healthy epigenetic patterns are termed 'epimutations'. They often result in the overexpression of tumorigenic genes or downregulation of tumor suppressor genes. Epimutations can be somatically acquired and maintained through cell division. When present in the germline, they can even be transmitted to the next generation (Egger et al. 2004; Mossman and Scott 2006).

Hence, the reprogramming of epigenetic states and correction of epimutations is of particular interest for potential clinical applications, but also for basic research in order to gain a better understanding of the complex epigenome network. So far, changes in the epigenome have predominantly been induced by pharmaceutical inhibitors, which are providing limited spatial and temporal control. However, the recent development of so-called 'EpiEditors' (also known as 'EpiEffectors') enables the editing of chromatin states in a more sophisticated and controlled way. The most frequent design of an EpiEditor implies the direct fusion of a chromatin-modifying enzyme to a DNA binding domain (DBD). The DBD is required for the recognition of a desired DNA sequence in the genome. To date, three programmable DBDs have been applied as fusion partners

for targeted epigenome editing, which are Zinc finger (ZnF) proteins, transcription activator-like effectors (TALE) arrays, and the Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR) sequences with CRISPR-associated protein 9 (dCas9) (CRISPR/Cas9) system (Figure 23) (Kungulovski and Jeltsch 2016; Laufer and Singh 2015).

In eukaryotic cells, ZnF domains of the C2H2 type are highly abundant, and particularly present in a large variety of transcription factors. They are composed of an array of multiple individual fingers, which all have a characteristic compact $\beta\beta\alpha$ fold, and are stabilized by a zinc ion. One finger is capable of recognizing a 3-bp motif on the DNA. This is facilitated by the docking of an α -helix to the major groove, and nucleotide-specific contacts of three residues on the α -helical surface. Various positively charged residues interact with the sugar/phosphate backbone, and thereby support the positioning of the fingers in the major groove of the DNA. Moreover, the linkers between the fingers contribute to the appropriate spacing at the DNA. By modulating the linker lengths and the residues involved in nucleotide recognition, a tandem array of ZnFs can be designed, capable to bind a desired and unspaced DNA motif. Theoretically, in order to target a unique sequence in the human genome, an array of six ZnFs is already sufficient as it can bind to an 18-bp motif (Kim and Kini 2017; Wolfe et al. 2000).

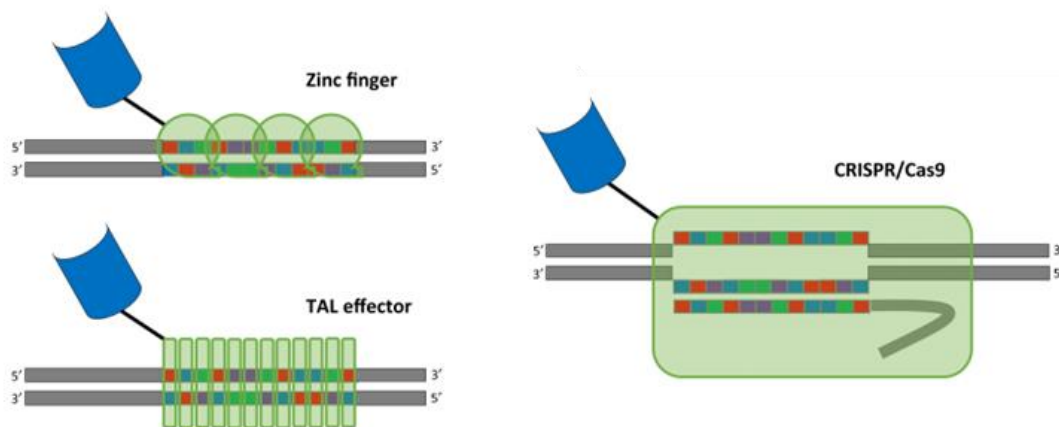


Figure 23: Schematic depiction of three programmable DNA binding domains (DBDs) for targeted epigenome editing approaches. The DBDs (green) are fused to a chromatin modifying effector domain (blue), to induce changes in the epigenome. Each finger of the zinc finger protein binds to three bases on the DNA (grey). In TAL effectors, one repeat reads a single base pair. The repeats can be arranged in an array to recognize longer DNA sequences. Target-specific binding of the CRISPR/Cas9 system is facilitated by Watson/Crick base-pairing of a guide RNA with a complementary DNA strand. Figure taken from Kungulovski and Jeltsch (2016).

Gersbach et al. (2014) developed a toolbox for targeting almost each possible DNA triplet by utilizing phage displays. In this process, they obtained a set of synthetic ZnFs, which in parts were able to bind their intended motifs with an up to 100-fold improved specificity in comparison to the naturally occurring variants (Gersbach et al. 2014). The relatively small size of the ZnF proteins, and their capability to also bind to lowly-accessible chromatin makes them a useful targeting device for epigenome editing approaches. Though, they are still inferior in overall specificity compared to other available targeting systems (Cano-Rodriguez and Rots 2016).

TALE proteins were originally discovered in the plant pathogen *Xanthomonas*, where they play a critical role for the control of host gene expression after infection. The DNA sequence readout of TALEs is enabled by a tandem array of 10 to 30 repeats, which are mostly around 34 aa in length, and form a left-handed two-helix bundle. The repeats differ in two neighboring residues (position 12/13), also termed the 'repeat variable diresidue' (RVD). Unlike the ZnF proteins, each RVD recognizes only one single nucleotide. This property allows a straightforward design of an array of repeats that can target a DNA sequence of interest. Importantly, TALEs exhibit a higher specificity compared to ZnFs. However, their relatively large size poses a restrictive factor for diverse therapeutic or biotechnological applications. In addition to that, the ZnF and TALE systems are limited in flexibility, since the targeting of a different genomic locus necessitates the complete redesigning of the corresponding domain (Bogdanove and Voytas 2011; Gupta and Musunuru 2014; Mak et al. 2013).

The intricacy of target-dependent *de novo* protein design was overcome by the CRISPR/Cas9 system. In 2012, the CRISPR/Cas9 derived from *Streptococcus pyogenes* was discovered as RNA-guided DNA endonuclease for programmable genome editing. In bacteria and archaea, CRISPR/Cas9 is involved in the adaptive immune defense against pathogenic DNA. Parts of this foreign DNA are incorporated into the host CRISPR locus. When expressed, the transcripts are further processed into small CRISPR RNA (crRNA) of around 42 nt in length. In type II systems, the crRNA hybridizes with the trans-activating crRNA (tracrRNA) required for endonuclease activity, followed by incorporation of the crRNA:tracrRNA hybrid into the Cas9 protein. For genome engineering, this hybrid structure can be substituted by a single chimeric RNA, also termed single guide RNA (sgRNA) of around 100 nt in length (Jinek et al.

2012). The association of CRISPR/Cas9 to DNA results in the melting of the DNA double strand and the formation of an RNA-DNA hybrid of 20 nucleotides. This interaction depends on the presence of a protospacer adjacent motif (PAM), which is located directly downstream of the target DNA sequence and recognized by the C-terminal part of the Cas9 protein (Anders et al. 2014). While the PAM motif was found to be 5'-NGG-3' in *S. pyogenes*, it varies among different bacterial species (Karvelis et al. 2015). The seven nucleotides upstream of the PAM are called the 'seed' region, and most crucial for CRISPR/Cas9 binding to its target sequence. In this region, the Cas9 protein introduces a double strand break using its HNH nuclease and RuvC-like nuclease domains. However, introduction of the D10A and H841A mutations were shown abrogate endonuclease activity, and thus generated a deactivated Cas9 (dCas9) protein (Qi et al. 2013).

With this advancement, CRISPR/dCas9 could be utilized as a platform for targeted epigenome editing without cleaving the underlying DNA. Since dCas9 association to DNA is based on Watson/Crick base-pairing with the sgRNA, the system is highly versatile. In order to target a different genomic locus, only the sgRNA has to be recloned, instead of redesigning the entire protein as required for the ZnF- and TALE-based systems. However, diverging observations regarding the specificity of the CRISPR/dCas9 system were made in different studies (Kungulovski and Jeltsch 2016). A further advancement was the establishment of a multiplexed sgRNA expression vector, enabling to target multiple loci simultaneously (Shao et al. 2018). Like the ZnF and TALE targeting modules, dCas9 can be directly fused to a chromatin modifying domain, which however limits the total number of effector domains at a certain target locus. For this reason, Tanenbaum et al. (2014) developed the so-called SUPERNOVA tag (SunTag) system, allowing to non-covalently recruit up to 24 antibody (Ab)-fused effector domains to a dCas9-fused GCN4 peptide repeat. Compared to the direct-fusion design for EpiEditors, the SunTag system not only amplifies the local concentration of effector domains, but also offers a higher protein flexibility, and a longer range for modifying the local chromatin (Tanenbaum et al. 2014). Recently, an alternative tunable concatenated CC peptide tag (CCC-tag) system has been introduced by Lebar et al. (2020). It functions by a similar principle as the SunTag system though utilizing specific coiled-coil interactions for the recruitment of multiple effector domains (Lebar et al. 2020).

Lastly, the Casilio system also allows the simultaneous multimerization of several effectors, though by a different mechanism. In this system, the sgRNA is designed to contain multiple 8-mer sequences, which can be recognized by various Pumilio/FBF (PUF) RNA-binding domains. By the direct fusion of different chromatin modifiers to the PUFs, up to 25 effector domains can be recruited at the same time, and in a desired arrangement (Cheng et al. 2016).

Over the past two decades, a multitude of studies have been published, demonstrating the opportunities of epigenome editing, but also the challenges of stable reprogramming of epigenetic states. Most groups thereby focused on the delivery of DNA methylation to gene promoters using the fusion of DNMT enzymes to the aforementioned targeting modules. In breast cancer cells, the deposition of 5mC to the *SOX2* and *Maspin* promoters by the fusion of the DNMT3A catalytic domain (3AC) to a ZnF was found to be stable, but not very efficient (Rivenbark et al. 2012). The *de novo* DNA methylation of the *VEGFA* locus using the same construct, however, was shown to be unstable after depletion of the EpiEditor, and accompanied by only transient gene silencing. In the same experimental setup, DNA methylation efficiency was higher when using a single-chain construct with the C-terminal part of DNMT3L (ZnF-3AC-3LC) instead of ZnF-3AC only (Kungulovski et al. 2015; Siddique et al. 2013). The same observation was made by Stepper et al. (2017), illustrating the increased activity of 3AC-3LC compared to 3AC when fused to dCas9 (Stepper et al. 2017). In addition, dCas9-3AC-3LC-mediated DNA methylation in the *CDKN2A* and *RASSF1* promoters was stably maintained in myoepithelial cells for several weeks (Saunderson et al. 2017). Amabile et al. (2016) utilized the TALE- and dCas9 systems in combination with 3AC and 3AC-3LC in a hit-and-run approach. They observed a rapid, but transient loss of reporter gene expression after DNA methylation editing, but only low silencing potential by the mark. However, the co-expression of ZnF-KRAB resulted in a strong and stable repression of the reporter gene (Amabile et al. 2016). In contrast, another study showed that the transient recruitment of the DNMT3B catalytic domain (3BC) was sufficient for stable reporter gene silencing (Bintu et al. 2016).

A major disadvantage of EpiEditors which are based on the direct fusion of a DNMT with a targeting module are the undesired off-target effects. This problem was overcome by the combination of the dCas9-SunTag with an Ab-fused 3AC (Ab-3AC), not only showing an improved methylation efficiency at the target locus, but also strongly

reduced off-target effects compared to the direct-fusion systems (Huang et al. 2017; Pflueger et al. 2018). Morita et al. (2016) modified the dCas9-SunTag system by introducing longer linkers between the GCN4 repeats, and combining the Ab-fused TET1 catalytic domain (Ab-TET1(CD)) as well as the sgRNA into a single vector. With this design, they were able to demethylate the *H19* ICR in mouse fetuses with good efficiency (Morita et al. 2016).

Targeted epigenome editing was also successfully applied for the introduction or removal of histone PTMs. The deposition of H3K9me_{2/3} at the *VEGFA* promoter by a ZnF-GLP construct was shown to be lost after several days, and induced a transient reduction in gene expression (Kungulovski et al. 2015). A different group showed that the targeting of dCas9-HDAC3 to the *Mecp2* promoter resulted in the loss of H3K27ac, and the repression of the associated gene (Kwon et al. 2017). Combinatorial editing of H3K27me₃ and DNA methylation with dCas9-EZH2 and dCas9-3AC-3LC respectively lead to the stable repression of the *HER2* gene, and maintenance of both modifications (O'Geen et al. 2019). *Vice versa*, the re-activation of gene expression was also achieved in several studies by the targeting promoter or enhancer sequences with dCas9 fusion proteins containing either PRDM9 (H3K4 HMTase), DOT1L (H3K79 HMTase) or P300 (HAT) (Cano-Rodriguez et al. 2016; Hilton et al. 2015).

2. Principal aims of the study

DNA methylation is an epigenetic signal associated with multiple regulatory processes such as gene silencing, repression of transposons or genomic imprinting. Dysregulation of DNMTs, or mutations within the enzymes result in the establishment of aberrant DNA methylation patterns that can cause various severe diseases (Greenberg and Bourc'his 2019). Over the past decades, substantial progress has been made in characterizing the complex mechanistic crosstalk between DNA methylation and certain histone post-translational modifications (Jeltsch et al. 2019). However, varying observations were made regarding the repressive capacity of *de novo* DNA methylation introduced by EpiEditors, the long-term stability of the mark, and hence the requirement for the co-delivery of specific histone PTMs. These diverging results may be attributed to differences in the experimental settings (Amabile et al. 2016; Bintu et al. 2016; Kungulovski et al. 2015).

The aim of the main project of this PhD thesis was to continue the work of Kungulovski et al. (2015), and explore the stability and downstream effects of *de novo* DNA methylation on a genome-wide scale. For this purpose, a stable HEK293 cell line should be established, capable of expressing the ZnF-3AC fusion construct in the presence of Doxycycline. In a hit-and-run-approach, the EpiEditor could then introduce DNA methylation globally, due to the promiscuous binding of the ZnF protein to several genomic binding sites. To determine the actual genomic targets of the ZnF, ChIP-seq experiments should be conducted. Subsequently, the differential stability of DNA methylation should be explored with temporal resolution. As amplicon-based bisulfite sequencing only allows to detect genomic DNA methylation levels in a small subset of target regions, the goal was to develop an MBD2-pulldown assay for the global analysis of 5mC dynamics. In addition to that, it was of great interest to investigate the local chromatin states, and identify specific states allowing or preventing the preservation of DNA methylation over long periods of times. *Vice versa*, the effects of *de novo* DNA methylation on various histone PTMs, and on gene expression should be investigated at different experimental time points. By this means, the causative effects of DNA methylation on the epigenome network should be revealed at thousands of different genomic loci.

In a second project, the goal was to develop a CRISPR/dCas9-based targeting device for the deposition of *de novo* DNA methylation with high efficiency and maximum specificity. This is of particular interest for potential therapeutic applications, since undesired off-target effects of EpiEditors still represent a major concern about these techniques (Kungulovski and Jeltsch 2016). For this reason, it was planned to compare different EpiEditing systems for the target-specific introduction of DNA methylation side by side. First, the direct fusion of 3AC and 3AC-3LC to dCas9 should be compared to the novel dCas9-10xSunTag system, in which the DNMTs are recruited to the targeting module after fusion to an antibody against the GCN4-peptide. As previous studies already reported a high specificity and good DNA methylation efficiency of 3AC in the context of the dCas9-10xSunTag system (Pflueger et al. 2018; Tanenbaum et al. 2014), the objective was to engineer 3AC mutants in direct fusion to 3LC, exhibiting an increased on-target activity with low off-target effects. The improved properties of the new constructs should then be globally validated by performing MBD2-seq.

Finally, a third project aimed to reprogram the *H19/IGF2* imprinting control region in HEK293 cells. This approach, in principle, is of high clinical relevance, as this locus is hypermethylated on the maternal allele in patients with Beckwith-Wiedemann syndrome or in Wilms' Tumor (Moulton et al. 1994; Nativio et al. 2011). It was planned to demethylate the ICR, utilizing a dual vector system, coexpressing an antibody-fused TET1 catalytic domain together with a modified dCas9-5xSunTag construct. By co-transfecting the cells with a multi-sgRNA plasmid, the EpiEditing system was intended to bind next to multiple CpG-rich motifs within the ICR, and induce demethylation at these sites. The initial loss of DNA methylation and the stability of its reduction should then be determined by MBD2-qPCR or amplicon-based bisulfite sequencing. Moreover, it was previously shown that the insulator protein CTCF was able to bind these CpG-rich motifs in unmethylated state, and thereby block the accessibility of a downstream enhancer towards the *IGF2* promoter (Hark et al. 2000). Thus, the goal was to quantify the CTCF occupancy after targeted demethylation of the ICR by conducting ChIP-qPCR.

3. Materials and methods

3.1 Molecular cloning

The plasmids employed in this work were cloned by utilizing various cloning strategies. Usually, to insert a DNA fragment into a vector, restriction- and ligation-based cloning, or the Gibson assembly method was applied (Gibson et al. 2009). Mutants of the DNMT3A catalytic domains were generated by site-directed mutagenesis, which is a megaprimer-based approach to introduce a desired point mutations (Jeltsch and Lanio 2002). Restriction-free (RF) cloning was used for the exchange of the fluorophore within the dCas9-3AC(-3L) direct fusion plasmids (van den Ent and Löwe 2006). Finally, the single sgRNA plasmids were prepared using a modified Golden Gate cloning protocol (Shao et al. 2018). A list of all constructs applied in the studies can be found in Table 1. The cloned plasmids were either transformed into One Shot™ Stbl3™ chemically competent cells (Thermo Fisher Scientific) or XL1-Blue electrocompetent cells (Agilent Technologies). Positive clones were usually identified by colony PCR, and then used for minipreparations of plasmid DNA, utilizing the NucleoSpin Plasmid kit (Macherey-Nagel). The validity of the sequences was analyzed by Sanger sequencing (Microsynth AG), followed by large-scale isolation of plasmid DNA using the Plasmid Midi Kit (QIAGEN).

Table 1: List of plasmids generated for the experiments of this work. Additional information about the applied cloning method, and collaborators involved in cloning procedure is shown. JB = Julian Broche, LL = Laura Laistner, PB = Dr. Pavel Bashtrykov, DH = Daniel Hofacker.

Project	Plasmid	Method	Cloned by
Genome-wide methylation	pAdTrack-CMV-ZnF-3xHA--CMV-eGFP	Gibson cloning	JB
Genome-wide methylation	pSIN-TRE3G-ZnF-3AC(WT)-FLAG-IRES-GFP--PGK-Neo	Gibson cloning	JB
Genome-wide methylation	pSIN-TRE3G-3AC(WT)-FLAG-IRES-GFP--PGK-Neo	Gibson cloning	JB
Genome-wide methylation	pSIN-TRE3G-ZnF-3AC(E752A)-FLAG-IRES-GFP--PGK-Neo	Gibson cloning	JB
Genome-wide methylation	pSIN-TRE3G-ZnF-3AC(WT)-FLAG-IRES-dTomato--PGK-Hygro	Restriction enzyme cloning	JB
Genome-wide methylation	pMSCV-LTR-Ren.660--PGK-Neo-IRES-GFP	Restriction enzyme cloning	JB

Genome-wide methylation	pMSCV-LTR-TET1.2851--PGK-Neo-IRES-GFP	Restriction enzyme cloning	JB
Genome-wide methylation	pMSCV-LTR-TET2.617--PGK-Neo-IRES-GFP	Restriction enzyme cloning	JB
Genome-wide methylation	pMSCV-LTR-TET3.9887--PGK-Neo-IRES-GFP	Restriction enzyme cloning	JB
Engineering of effector domains	CMV-dCas9-3AC(WT)-3xFLAG-tagBFP	RF cloning	JB/LL
Engineering of effector domains	CMV-dCas9-3AC(WT)3L-3xFLAG-tagBFP	RF cloning	JB/LL
Engineering of effector domains	SV40-dCas9-10xSunTag-P2A-tagBFP-WPRE	Restriction enzyme cloning	JB
Engineering of effector domains	SFFV-Ab-3AC(WT)-3xFLAG-sfGFP-GB1	Restriction enzyme cloning	PB
Engineering of effector domains	SFFV-Ab-3AC(WT)-3L-3xFLAG-sfGFP-GB1	Site-directed mutagenesis	PB
Engineering of effector domains	SFFV-Ab-3AC(K766E)-3L-3xFLAG-sfGFP-GB1	Site-directed mutagenesis	PB
Engineering of effector domains	SFFV-Ab-3AC(R831E)-3L-3xFLAG-sfGFP-GB1	Site-directed mutagenesis	PB
Engineering of effector domains	SFFV-Ab-3AC(K844E)-3L-3xFLAG-sfGFP-GB1	Site-directed mutagenesis	PB
Engineering of effector domains	SFFV-Ab-3AC(R885A)-3L-3xFLAG-sfGFP-GB1	Restriction enzyme cloning	JB
Engineering of effector domains	SFFV-Ab-3AC(R887E)-3L-3xFLAG-sfGFP-GB1	Site-directed mutagenesis	PB
Engineering of effector domains	U6-sgRNA(ISG15)--CMV-dsRed	Golden Gate cloning	LL
Engineering of effector domains	U6-sgRNA(scr)--CMV-dsRed	Golden Gate cloning	PB
Demethylation of ICR	pCAG-dCas9-5xSunTag(22aa-linker)--SFFV-Ab-TET1(CD)-IRES-GFP	Restriction enzyme cloning	JB
Demethylation of ICR	pCAG-dCas9-5xSunTag(22aa-linker)--SFFV-Ab-(empty)-IRES-GFP	Restriction enzyme cloning	JB
Demethylation of ICR	multi-sgRNA1-5(ICR)--CMV-dsRed	Golden Gate cloning	DH

3.2 Cell culture experiments

3.2.1 Cell maintenance

All experiments utilizing dCas9-based epigenome editing were carried out with the Human Embryonic Kidney 293 (HEK293) cell line which has been established by Graham et al. (1977), and was purchased from DSMZ (Braunschweig, Germany). HEK293 cells are tumorigenic, and suspected to originate from the neuronal lineage or the embryonic adrenal precursor structure. Furthermore, they display a hypertriploid karyotype with varying numbers of chromosomes ($n = 62 - 78$), and due to the lack of a Y chromosome, they are regarded as female (Stepanenko and Dmitrenko 2015).

For the experiments which required the generation of stable cell lines, HEK293 cells expressing the ecotropic receptor (HEK293R) at the cell surface were used. By this means, the cells could be transduced by vectors pseudotyped with murine ecotropic envelope under biosafety level 1 conditions, as these vectors are considered non-infections for human cells (Schambach 2012) (Schambach et al. 2006). The HEK293R cell line was previously established as described in Appendix 1, and kindly provided by Dr. Max Emperle. HEK293 and HEK293R cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich, Inc.) supplemented with 10 % fetal bovine serum (FBS, Sigma-Aldrich, Inc.), 4 mM L-glutamine (Sigma-Aldrich, Inc.), 100 U/ml penicillin, and 100 µg/ml streptomycin (Sigma-Aldrich, Inc.), and will be hereinafter referred to as DMEM⁺. They were grown in an incubator at 37°C, 5 % CO₂, and a relative humidity of 95 %. When the cells reached a confluency of around 90 %, they were split by washing with Dulbecco's Phosphate Buffered Saline without MgCl₂ (DPBS⁻) (Sigma-Aldrich, Inc.), immersing with Trypsin-EDTA solution (Sigma-Aldrich, Inc.) and incubating at 37°C. After detachment, the cells were resuspended in DMEM⁺ and a fraction of the cells was re-transferred into a vessel for adherent cells.

3.2.2 Transient transfection

The transient transfection of HEK293 cells with the pAdTrack-CMV-ZnF-3xHA plasmid was accomplished with the cationic polymer polyethylenimine (PEI) MAX 40k (Polysciences). A more detailed description of the experimental procedure can be found in Appendix 1. For the dCas9-mediated epigenome editing experiments, the lipid-based

FuGENE® HD transfection reagent was utilized, as it was reported to cause only low effects on the global gene expression, but offers high transfection efficiency (Jacobsen et al. 2009). For more precise information about the respective transfection conditions, please refer to Appendix 2. The relative amount of plasmid DNA was adjusted for the ICR demethylation studies, since a dual vector was used in this experiment, coding for the dCas9-5xSunTag targeting module, and the Ab-TET1 effector domain. While the amount (500 ng) of the multi-sgRNA plasmid remained constant, 9 µg of the dual vector (Table 1) was applied for the transient transfection of HEK293 cells in a 10-cm petri dish. The sgRNA sequences used for targeting the ICR are shown in Table 2.

Table 2: Sequences of the guide RNAs for multiplexed targeting of the *H19/IGF2* imprinting control region

Single guide RNA name	Sequence (5' → 3')
sgRNA5	CCCCAGTTTGGGCGGGCTC
sgRNA6	TCTCACCGCCTGGATGGCA
sgRNA7	GGGCGAACCCCATCCAGGG
sgRNA8	CGCCCCGATGGTGCAGAAT
sgRNA9	TGCCCTGATGGCGCAGAAT

3.2.3 Generation of stable cell lines

Various stable HEK293R cell lines were established which are capable of co-expressing a (ZnF)-3AC construct together with a fluorophore over an IRES. Thereby, the expression can be controlled by a doxycycline (dox)-inducible TRE3G promoter. In addition, an antibiotic selection marker is constitutively expressed (Figure 24A). The cell lines were generated as previously described by Rathert et al. (2015). For the production of murine stem cell virus (MSCV) particles, which are able to infect HEK293R cells expressing the ecotropic receptor on their surface, Platinum-E (Plat-E) cells were used. Plat-E cells are derived from the HEK293-T cell line and were engineered for high-titer virus production (Morita et al. 2000). They were cultured under the same conditions as HEK293R cells and were transiently transfected to allow retroviral packaging (Figure 24B). Therefore, 1.4×10^6 Plat-E cells (Cell Biolabs) were seeded into a 100-mm cell

culture dish for adherent cells (Greiner Bio-One). After six hours, a calcium phosphate transfection was carried out by preparing a mixture ('solution A') consisting of 10 - 20 µg construct plasmid, 5 - 10 µg helper plasmid (pCMV-Gag-Pol, Cell Biolabs), and 250 mM CaCl₂, adjusted to a volume of 500 µl with ddH₂O. 500 µl of Hank's buffered salt solution (HBS, 'solution B') (280 mM NaCl, 50 mM HEPES, 1.5 mM Na₂HPO₄, 12 mM Dextrose, 10 mM KCl, adjusted to pH 7.0) was pipetted into a flow cytometry tube. Using a Pasteur pipet, bubbles were created in solution B, and solution A was added dropwise. After 15 min of incubation at room temperature, the DNA mixture was carefully added to the Plat-E cells and incubated overnight. Subsequently, the medium was first exchanged by fresh DMEM⁺, and later reduced to 5 ml to enable later harvesting of virus particles at high titer.

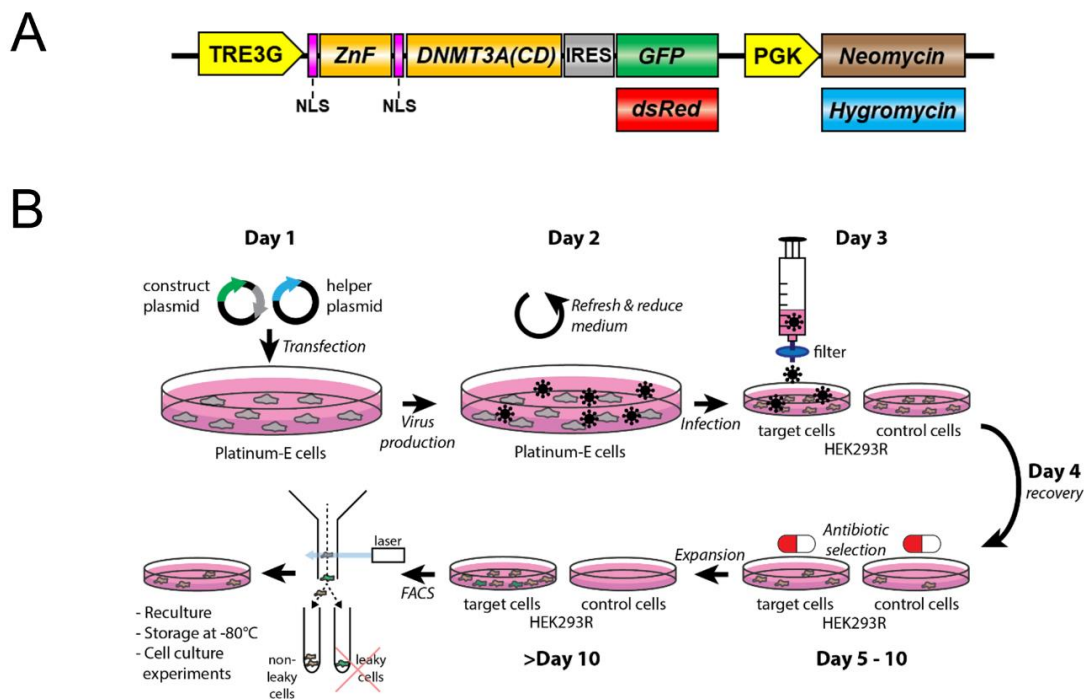


Figure 24: Generation of stable HEK293R cell lines for genome-wide methylation studies. A) Composition of the construct plasmid used for the stable integration into HEK293R genome. TRE3G = 3rd-generation Tet-responsive promoter; NLS = nuclear localization sequence, ZnF = zinc finger, DNMT3A(CD) = catalytic domain of DNMT3A, IRES = internal ribosomal entry site, GFP = green fluorescent protein, dsRed = dimeric red fluorescent protein, PGK = phosphoglycerate kinase 1 promoter. B) Schematic workflow for the generation of stable HEK293R cell lines containing the construct shown in A).

The next morning, 0.3×10^6 HEK293R target cells were seeded into the well of a 6-well cell culture plate containing 2 ml DMEM⁺. For retroviral transduction, the MSCV-containing supernatant of the transfected Plat-E cells was filtered through a 0.45 µm Filtropur S syringe filter (Sarstedt) and transferred to the target cells. The Plat-E cells

were discarded, and the medium was replaced by 2 ml of fresh DMEM⁺ 8 h post-transduction. After two days of recovery, the antibiotic selection was initiated by splitting the target cells in a 1:4 ratio, and resuspending them in 2 ml DMEM⁺ supplemented with either 300 µg/ml neomycin (Sigma-Aldrich) or 200 µg/ml hygromycin B (Sigma-Aldrich). An additional well with untransduced cells was also included and served as a control for antibiotic selection. Every 2 – 3 days, the supernatant was exchanged by fresh DMEM⁺ with antibiotics until no viable cells were visible in the control well. The selected cells were expanded in DMEM⁺ until they reached 90 % confluency in a T75 flask (Sarstedt).

Since a fraction of the cells usually showed leaky expression of the construct, a fluorescence-activated cell sorting (FACS) step was included to enrich only non-fluorescent cells. For this purpose, the cells were harvested by washing with DPBS⁻, trypsinizing and spinning down for 5 min at 300 g. Afterwards, the cell pellet was resuspended in 0.5 – 1 ml DMEM⁺, and filtered through a 70-µm pre-separation filter (Miltenyi Biotec) to remove clumps. Enrichment of single live cells displaying no fluorescence in uninduced state was achieved by applying the gating strategy shown in Figure 25 utilizing the Cell Sorter SH800S (Sony Biotechnology). Around 0.7×10^6 cells per sorted cell line were re-seeded into the well of a 6-well plate and expanded until they reached 90 % confluency in a T75 flask. For long-term storage, a fraction of the cells was resuspended in cryopreservation medium (50 % DMEM, 40 % FBS, 10 % DMSO), and frozen to -80°C using the Nalgene[®]

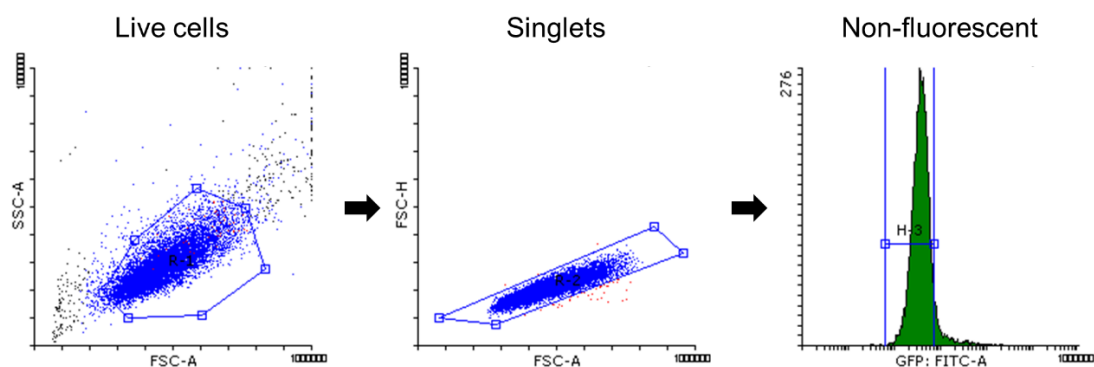


Figure 25: FACS gating strategy after the generation of doxycycline (dox)-inducible stable HEK293R cell lines for the enrichment of cells showing no leaky expression of the inducible transcript (H-3 gate). Plots were generated using the Flowing Software V2.5.1 (<http://flowingsoftware.btk.fi>).

3.2.4 Cell culture experiments for genome-wide methylation studies

Time-course experiments of EpiEditor expression were conducted utilizing the HEK293R (pSIN-TRE3G-ZnF-3AC-IRES-GFP-PGK-Neo) cell line to analyze the long-term stability of genome-wide *de novo* DNA methylation, and its interplay with the epigenome network. Control experiments using the catalytically inactive E756A mutant of 3AC, or 3AC without ZnF fusion were performed for up to three days under dox treatment. For a detailed description of the experimental workflow, please refer to Appendix 1. In brief, 6×10^6 cells were seeded into a T175 flask (Sarstedt), and treated with 1 $\mu\text{g}/\text{ml}$ doxycycline (Sigma-Aldrich) in DMEM⁺ for 3 days. Single living cells showing a green fluorescence signal were enriched using the BD FACSAria III (BD Biosciences) or Cell Sorter SF800S (Sony Biotechnology) (Figure 26). Subsequently, the cells were either snap-frozen in liquid nitrogen and stored at -80°C , or re-cultivated in a 6-well plate for later experimental time points, without the addition of dox. After 5 or 8 days of dox removal, 80 % of the cells were harvested and the rest of the cells was re-seeded. Finally, after 11 days of dox removal, all cells were harvested and culturing was discontinued.

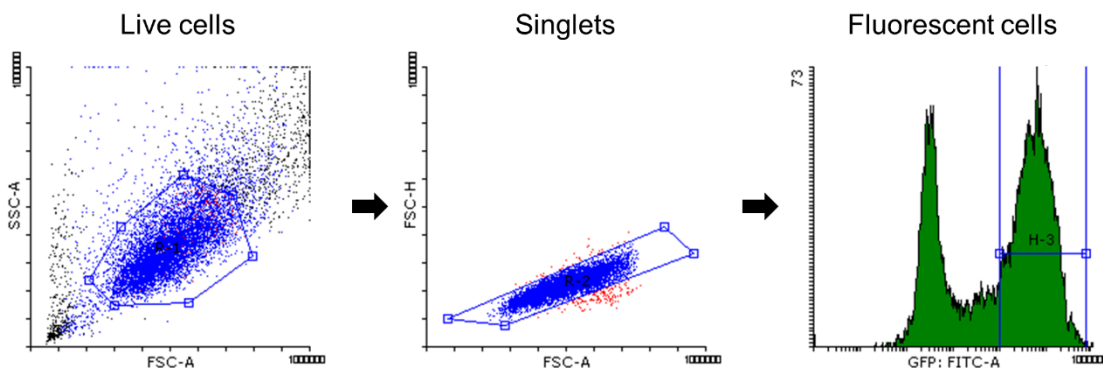


Figure 26: Flow cytometry gating strategy for the enrichment of GFP-positive HEK293R cells (gate H-3) co-expressing ZnF-3AC. The cells were treated with 1 $\mu\text{g}/\text{ml}$ doxycycline for 3 days. Plots were generated using the Flowing Software V2.5.1 (<http://flowingsoftware.btk.fi>).

3.2.5 Knockdown of TET enzymes in DNA methylation studies

In order to investigate the potential role of active DNA demethylation in the genome-wide methylation studies, knockdown of the three members of the TET family dioxygenases (TET1-3) was conducted. Therefore, short hairpin RNAs (shRNAs) for individual TET enzymes and a *Renilla luciferase* shRNA (Table 3) were cloned into

pMSCV vector and the stable HEK293R (TRE3G-ZnF-3AC-IRES-dsRed—PGK-Hygro) cell line was transduced as described in chapter 3.2.3. In this experiment, the virus supernatant was diluted 4 to 10-fold in DMEM⁺ this time to achieve a transduction efficiency around 5 - 10 %, enabling single integration of the retrovirus into the genome (Rathert et al. 2015). The transduction efficiency was validated by flow cytometry using the MACSQuant[®] VYB Flow Cytometer (Miltenyi Biotec) and the cells underwent antibiotic selection using 300 µg/ml neomycin (Sigma-Aldrich) in DMEM⁺. Afterwards, the selected cells were expanded and 19 days after transduction, when the cell numbers were sufficient, the experimental workflow to study DNA methylation dynamics was carried out as previously described in chapter 3.2.4. Around 1 x 10⁶ cells were harvested without dox treatment to analyze the efficiency of shRNA-mediated knockdown by RT-qPCR.

Table 3: List of shRNAs used for the knockdown of TET1-3. Additional information about NCBI Gene ID, and the guide sequence is shown. The location within the CCDS is included in the name of the shRNA.

Target transcript	shRNA name	NCBI Gene ID	Guide sequence
TET1	TET1.2851	80312	TTGTTTTCCAATGTCTTGCCGA
TET2	TET2.617	54790	TTTCTTATCACTCAAATCGGAG
TET3	TET3.9887	200424	TAAAACATCAACTTCTGTGGAC
Renilla luciferase	Ren.660	-	TTACTAACGGGATTTACGAGG

3.2.6 Application of enzyme inhibitors in cell culture

For the genome-wide DNA methylation studies, different inhibitors have been used in cell culture experiments. N-oxalylglycine (N-OG) was applied for the simultaneous inhibition of TET enzymes, as it is an inhibitor of α KG-dependent dioxygenases (Marholz et al. 2016; Peng et al. 2016; Sudhamalla et al. 2017; Sudhamalla et al. 2018). In the experiment aiming to reduce the global H3K27me3 levels, EPZ-6438 (AduoQ Bioscience), also known as ‘Tazemetostat’, was used as it was reported to selectively inhibit the HMTase EZH2 (Knutson et al. 2014). Both inhibitors were diluted in DMSO. HEK293R (TRE3G-ZnF-3AC-IRES-GFP—PGK-Neo) cells were treated with either

500 μ M N-OG or 10 μ M EPZ-6438 in DMEM⁺ for 3 days before the induction of ZnF-3AC. During the whole experiment, the cells were constantly treated with the inhibitors. Every 2 - 3 days, the medium containing the inhibitors was exchanged. As a negative control, cells treated with the same volume of DMSO were included to rule out side-effects caused by the solvent of the inhibitors.

3.3 SDS PAGE and western blot

Protein samples were obtained from whole cell lysates of treated and untreated HEK293R (TRE3G-ZnF-3AC-IRES-GFP—PGK-Neo) cells. For the experiments aiming to quantify the protein levels of ZnF-3AC before and after dox treatment, a stepwise description is available in Appendix 1. Moreover, the changes of global H3K27me3 levels were investigated in cells treated with either EPZ-6438 or DMSO. In order to achieve a better resolution of the relatively small histone proteins, the samples were run on a 15-% tricine-SDS gel according to the protocol published by Schagger (2006). Subsequently, Western Blot was performed as described in Appendix 1. An annotated list of all antibodies employed for the experiments is shown in Table 4.

Table 4: List of antibodies used for Western Blot experiments.

Antibody	Catalogue number	Lot #	Antibody dilution
α -DNMT3A	NB120-13888 (Novus Biologicals)	-	1 : 750
α -Lamin B1	sc-6217 (Santa Cruz Biotechnology)	F1014	1 : 500
α -H3K27me3	#39155 (Active Motif)	31917019	1 : 1000
α -H4	#61299 (Active Motif)	13912001	1 : 1000
α -Rabbit IgG HRP-linked whole Ab	NA934 (GE Healthcare)	9739638	1 : 5000
α -Mouse IgG HRP-linked whole Ab	NXA931V (GE Healthcare)	973269	1 : 5000
α -Goat IgG (whole molecule)–Peroxidase antibody	A4174 (Sigma-Aldrich)	071M4767	1 : 10,000

3.4 Nucleic acid analysis

3.4.1 MBD2-pulldown

For the analysis of genome-wide DNA methylation states, MBD-pulldown followed by next-generation sequencing (NGS) has emerged as a powerful and cost-efficient alternative to commonly used techniques such as antibody-based methylated DNA immunoprecipitation sequencing (MeDIP-seq) or Whole-Genome Bisulfite Sequencing (WGBS) (Jeltsch et al. 2019). A self-developed protocol has been applied which employs the human full-length Methyl-CpG-binding domain protein 2 (MBD2) fused to the C-terminus of a Glutathione-S-transferase (GST) tag. MBD2 has been shown to display a high binding affinity for 5mC, and does not bind to 5hmC or unmodified cytosine (Mellén et al. 2012). The general workflow of MBD-pulldown experiments is shown in Figure 27A.

In the first step, genomic DNA (gDNA) was isolated from the cells pellets, using the QIAamp DNA Mini Kit (QIAGEN) according to the manufacturer's protocol, and was eluted in 200 µl ddH₂O, pre-warmed to 70°C. The gDNA was then sheared with the EpiShear™ Probe Sonicator (Active Motif) using the 2-mm microtip probe (25 % amplitude, 20 x 20-s pulse / 30-s pause). 200 ng of the sonicated DNA was loaded on a 1 % agarose gel to validate whether the fragment sizes ranged from 100 – 1000 bp (Figure 27B). The fragmented gDNA was then re-purified and concentrated using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel), eluted in 50 - 100 µl ddH₂O pre-warmed to 70°C and the concentration was determined with the NanoDrop 1000 (Thermo Fisher Scientific).

Per reaction, 1 µg of sonicated DNA and 1000 nM GST-MBD2 were mixed in a total volume of 250 µl pre-cooled to 4°C PB150 buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.5 % IGEPAL® CA-630, 2 mM DTT). Immobilization of the MBD2 domain to methylated DNA was enabled by rotation the mixture on a turning wheel at 4°C overnight. The next morning, per sample 50 µl of Glutathione Agarose beads (Macherey-Nagel) were equilibrated by resuspending in 200 µl of cold PB150 buffer, and centrifuging for 1 min at 2000 g. After a total of four washes with cold PB150 buffer, the supernatant was removed and the DNA-GST-MBD2 mixture was transferred to the beads. The sample was rotated for 2 h on a turning wheel at 4°C, and subsequently, the beads were spun down in a pre-cooled centrifuge at 2000 g for 2 min. Then, the

supernatant containing the unbound DNA was removed, and 200 μ l of ice-cold PB500 wash buffer (50 mM Tris–HCl pH 8.0, 500 mM NaCl, 1 mM EDTA, 0.5 % IGEPAL[®] CA-630, 2 mM DTT) was added. After 5 min of rotation on a turning wheel at 4°C, the sample was centrifuged for 2 min at 2000 g. A total of three washing steps with PB500 were performed. Elution was carried out by resuspending the beads in 150 μ l of unchilled PB2000 elution buffer (10 mM Tris–HCl pH 8.0, 2 M NaCl, 1 mM EDTA) and rotating the sample on a turning wheel for at least 15 min at room temperature. The sample was centrifuged and the supernatant, containing the precipitated methylated DNA, was transferred into a 2-ml reaction tube. After a second elution step was performed, the elution fractions were pooled and the DNA was purified using the ChIP DNA Purification Kit (Active Motif) according to the manufacturer’s protocol. The DNA was eluted in 50 – 100 μ l of the elution buffer provided by the kit and stored at either 4°C (short-term) or -20°C (long-term).

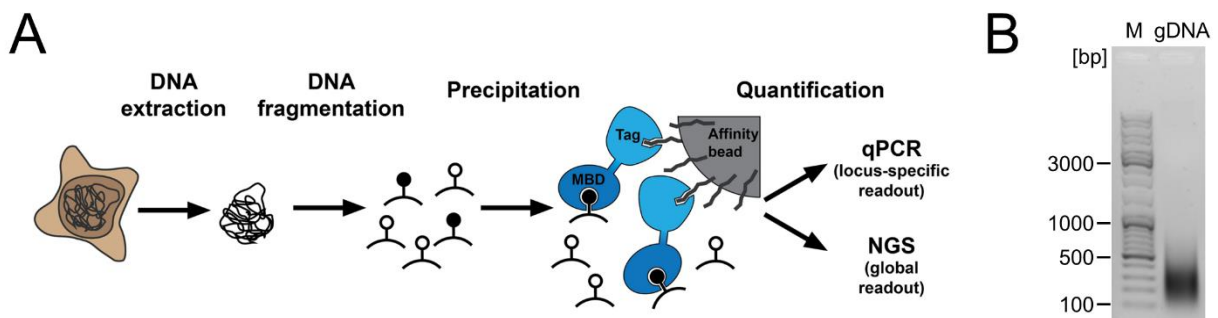


Figure 27: MBD-pulldown as a tool for the enrichment of methylated DNA. A) Schematic depiction of the experimental workflow of MBD-pulldowns. Taken from Jeltsch et al. (2019). B) 1-% agarose gel with a DNA molecular weight marker ('M', GeneRuler, 1 kb DNA ladder) loaded on lane 1, and 200 ng sonicated genomic DNA (gDNA) on lane 2.

3.4.2 Cross-linked ChIP (XChIP)

In order to detect the genomic distribution of DNA-bound proteins or their direct interaction partners, cross-linked chromatin immunoprecipitation (XChIP) is a widely used technique. In this approach, the cells are fixed by formaldehyde, allowing to create cross-links between the protein and the DNA-residues within a distance of 2 Å. The chromatin is then sheared by sonication to obtain fragments in the range of 200 – 1000

bp, and an antibody specific for the desired protein is added. The antibody-bound protein/DNA complexes are then precipitated by the addition of Protein A/G-coupled beads, enabling the specific enrichment of the protein-associated genomic loci. After extensive washing steps, the cross-links are reversed, and the DNA is purified. The relative enrichment at different loci can then be quantified by qPCR or NGS (Nelson et al. 2006; Orlando 2000).

In the experiments aiming to analyze the genome binding of the ZnF-3xHA construct in HEK293 cells, a protocol was utilized which is described in detail in Appendix 1. In brief, the cells were transiently transfected with the pAdTrack-CMV-ZnF-3xHA--CMV-eGFP plasmid using Polyethylenimine (PEI) MAX 40k (Polysciences). 3 days post-transfection, around 1×10^7 cells were harvested, cross-linked in a final concentration of 1 % (v/v) formaldehyde (Sigma-Aldrich, Cat. No. 25,254-9) and the reaction was quenched using glycine in a final concentration of 125 mM. After cell and nuclei lysis, the chromatin was sheared using the EpiShear™ Probe Sonicator (Active Motif) with the 2-mm microtip probe (40 % amplitude, 24 x 20-s pulse / 30-s pause). In a next step, the chromatin was pre-cleared using a isotype control antibody of the same species, followed by overnight incubation with an antibody raised against the triple-HA tag (Table 5). Then, Protein A/G magnetic beads (Thermo Fisher Scientific) were added to enable binding of antibody-chromatin complexes. After a total of six washing steps using wash buffers with different stringency, the crosslinking was reversed by incubation at 65°C overnight with the addition of 20 µg RNase A (Machery-Nagel). The sample was supplemented with 2.4 U Proteinase K (New England Biolabs), and further incubated for 2 h at 45°C. Finally, the DNA was purified using the ChIP DNA Purification Kit (Active Motif) and analyzed by qPCR and NGS.

For the ChIP experiments analyzing the binding of CTCF at the *H19/IGF2* imprinting control region after targeted DNA demethylation, the cells were transiently transfected with the multi-sgRNA1-5(ICR) plasmid, and the dCas9-5xSunTag/Ab-TET1(CD) dual vector. After 3 days, the cells were harvested as described in section 3.2.2, and single live cells showing green and red fluorescence were enriched using the Cell Sorter SH800S (Sony Biotechnology). After sorting 0.7×10^6 cells, they were re-seeded into the well of a 6-well plate for recovery. 6 days after transfection, a fraction of the cells was split for further cultivation for additional 21 days, while around 2×10^6 cells were fixed with 1 % (v/v) formaldehyde as described in Appendix 1 and stored at -80°C after

snap-freezing with liquid nitrogen. The same procedure was applied for the cells harvested 27 days after transfection and for untransfected control cells. Different from the ZnF-ChIP experiment, the Magna ChIP™ HiSens Chromatin Immunoprecipitation Kit (Merck Millipore) was utilized due to the requirement of substantially less cells compared to the protocol described above. The ChIP was performed following the instructions of the manufacturer's protocol, and the chromatin was sheared using the same settings as described for ZnF-ChIP. Out of the 500 µl of sample volume, 50 µl were used for immunoprecipitation with either 3 µg of CTCF-antibody, or 3 µg of a non-specific IgG control from the same species (Table 5). After elution, no further purification step was necessary. The samples were diluted 1:1 with ddH₂O and could be used directly for qPCR analysis.

Table 5: List of antibodies used for XChIP experiments

Antibody	Catalogue number	Lot number	Amount per reaction
α-HA tag	ab9110 (Abcam)	GR261166-2	5 µg
Normal rabbit IgG	sc-2027 (Santa Cruz Biotechnology)	ER1619081	3 - 5 µg
α-CTCF	#2899 (Cell Signaling Technology)	2	3 µg

3.4.3 Native ChIP (NChIP)

Native chromatin immunoprecipitation (NChIP) is a method which is mostly used for the localization of histone proteins in the genome. Due to their positive charge, they tightly bind to the negatively charged phosphate backbone of DNA and thus do not require a cross-linking step, which could impair antibody binding to the epitope. Instead, the chromatin is fragmented by enzymatic digestion of the linker DNA using micrococcal nuclease (MNase) (O'Neill 2003). For this work, a protocol has been developed which uses 1×10^6 cells per reaction and utilizes whole-cell lysates without a nuclei isolation step to avoid loss of chromatin. The fragmentation of native chromatin into mono- and oligonucleosomes was validated by agarose gel electrophoresis (Figure 28A) or LabChip GX II Touch HT (Perkin Elmer) (Figure 28B). For a detailed description of the NChIP procedure, refer to Appendix 1. An overview of the antibodies

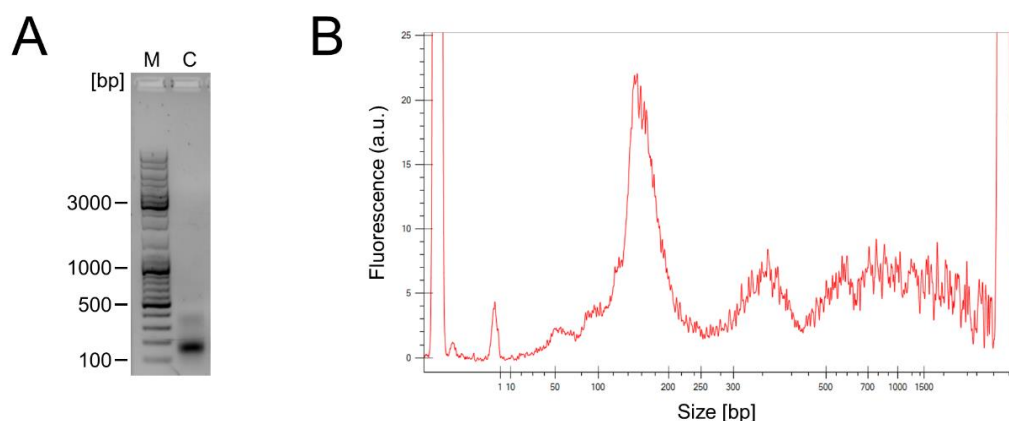


Figure 28: Fragmentation of native chromatin using micrococcal nuclease (MNase). A) Representative example of 1-% agarose gel with a DNA molecular weight marker ('M', GeneRuler, 1 kb DNA ladder) loaded on lane 1 and 250 ng digested chromatin ('C') on lane 2. B) Exemplary run of purified DNA on LabChip GX II Touch HT (Perkin Elmer), originating from digested chromatin.

and the respective amounts used for the experiments is provided in Table 6. Afterwards, the NChIP samples were tested by qPCR to analyze the locus-specific enrichment of the PTMs at selected loci. Furthermore, they were used to prepare libraries for NGS to globally investigate the dynamics of different modifications after genome-wide DNA methylation editing.

Table 6: List of antibodies used for NChIP experiments

Antibody	Catalogue number	Lot number	Amount per reaction
α -H3K4me3	ab8580 (Abcam)	GR3203653-1	2.5 μ g
α -H3K9me3	ab8898 (Abcam)	GR3176466-3	2.5 μ g
α -H3K27me3	#39155 (Active Motif)	31917019	5 μ g
α -H3K36me3	ab9050 (Abcam)	GR105541-1	3.5 μ g
α -H3K27ac	ab4729 (Abcam)	GR3216173-1	3 μ g

3.4.4 Quantitative PCR (qPCR)

The relative amount of DNA enriched in MBD2-pulldown or XChIP experiments was analyzed by real-time quantitative polymerase chain reaction (qPCR). In these experiments, a fluorescent dye such as SYBR Green I is quenched in its free state, but shows fluorescence when bound to double-stranded DNA (Dragan et al. 2012). As the

amount of DNA usually doubles during each PCR cycle, the cycle number in which the threshold of fluorescence signal is intersected by the amplification curve (C_t value) is proportional to the logarithm of the starting concentration of the DNA (Ponchel et al. 2003).

A separate mastermix was prepared for every primer set using the commercially available 2X ORA™ SEE qPCR Green Rox L Mix (highQu) as indicated in Table 7. In order to normalize the individual MBD2-pulldown or XChIP samples, a three-step 1:5 dilution series of the corresponding input samples was prepared. This also allowed the generation of a standard curve, and thus provided information about the PCR efficiency. All PCR reactions were conducted in technical triplicates. 14 μ l of mastermix was pipetted into the well of a low-profile 96-well plate (Genaxxon Bioscience). Subsequently, 1 μ l of the respective sample DNA was added to the mix, and a no-template control was included. The plate was then sealed using the adhesive Microseal 'B' PCR Plate Sealing Film (Bio-Rad Laboratories), followed by a centrifugation step at 1000 g for 1 min. Finally, the qPCR was performed employing the CFX96 Real-Time PCR detection system (Bio-Rad Laboratories) and using the cycling conditions shown in Table 8. After the PCR run, the purity of the qPCR reaction was verified by examination of the melting curve. For data analysis, the threshold was at usually 200 RFU, and the starting quantities relative to the input were exported in XLSX data format. An overview of all tested regions in MBD2-pulldown or XChIP experiment and the appropriate primer sequences is shown in Table 9.

Table 7: Mastermix prepared per qPCR reaction

Reagent	Volume
ORA™ SEE qPCR Green Rox L Mix (2X)	7.5 μ l
Forward primer (10 μ M)	0.4 μ l
Reverse primer (10 μ M)	0.4 μ l
ddH ₂ O	5.7 μ l

Table 8: Cycling conditions used for qPCR experiments

Temperature	Time	40x
95°C	3 min	
95°C	3 s	
T _A	20 s	
72°C	4 s	
Measurement		
65 - 95°C (0.5°C steps)	5 s/step (melt curve)	

Table 9: List of primers used for qPCR after MBD2-pulldown and XChIP experiments. Additional information about the genomic location and size of the amplicon is provided.

Experiment	Region	Coordinates amplicon (hg19)	Amplicon size	Forward primer	Reverse primer
ZnF-ChIP	ZnF target (VEGFA CGI)	chr6:43738352-43738469	118 bp	ACAGGGGCAAAG-TGATGAC	GCGGTGTCTGTC-TGTCTGT
ZnF-ChIP	ZnF off-target (FZD10 CGI)	chr12:130646580-130646676	97 bp	CCGAACCTCCCG-TAACCTC	TTGCCTCTCGCT-ATCCTCTC
MBD2-pulldown	SLC6A3 CGI	chr5:1446238-1446335	98 bp	GCACTCGCCTAA-GAAAACCA	GGAAGGAAAGCC-TCGGAGT
MBD2-pulldown	VEGFA CGI	chr6:43738171-43738372	202 bp	GCTTGCCATTCC-CCACTTGAATCG	GGTCACTCACTTT-GCCCCTGTC
MBD2-pulldown	NRP1 CGI	chr10:33623705-33623814	110 bp	AAATCCGGCTTG-TTTCTGG	TGTCTCCCGCTC-ATCTTTTC
MBD2-pulldown	CCND2-AS CGI	chr12:4378943-4379016	74 bp	CGCCTTCTTAAC-TCACGCCT	CTTCGGCGAATT-TCGGCTTG
MBD2-pulldown	SMURF1 CGI	chr7:98741344-98741417	74 bp	CGTACCTGTCAG-ACGGATCTT	ATCGTTGGCGGG-GATGT
MBD2-pulldown	ISG15 CGI	chr1:948717-948862	146 bp	CGGTCATTCCGGT-TTTGTTTC	CACCGGCCCTAT-TATTAAGC
MBD2-pulldown	NOTCH2 CGI	chr1:120611383-120611452	70 bp	GTGGGGTTGGAG-TGTCAAAA	GGTGTGTGGGCT-TGGTTT
MBD2-pulldown	SHH CGI	chr7:155601912-155601982	71 bp	ACTGCTCCGAAA-GTTCCACT	TCTCTAATGTGA-CTGCCGCC

MBD2-pulldown	IGF2 CGI	chr11:2161342-2161465	124 bp	GGGAAACACAGC-TCAAATCC	AAGTCCAACGCA-CTGAGGAC
CTCF-ChIP	CTS1	chr11:2024181-2024300	120 bp	TGGCGCAGAATC-GGCTGTAC	GAGACCTGGGAC-GTTTCTGT
CTCF-ChIP	CTS2	chr11:2023760-2023902	143 bp	GCCCCGATGGTG-CAGAATC	CGGCACCTAGCT-TGCGT
CTCF-ChIP	CTS3	chr11:2023404-2023496	93 bp	GATGGCACAGAA-TCGGTTGTAAG	CATCCAGGGAG-GGCTTGG
CTCF-ChIP	CTS4	chr11:2021929-2022068	140 bp	ACCCGGATGGTG-CAGAATTG	CGCCTGGCTTGC-GGGA
CTCF-ChIP	CTS5	chr11:2021525-2021643	119 bp	GGTTGTAGTTGT-GGAATCAGAGG	CCCGGCTTGGA-TGACCT
CTCF-ChIP	CTS6	chr11:2021105-2021259	155 bp	ACCGCCTGGATG-GCACG	TGAACCTGCG-ACGCGT
CTCF-ChIP	CTS7	chr11:2020241-2020345	105 bp	TTGGGTACCT-TCAGACTGTGAT	TGAGCCCATCTC-CCAGCGAT
CTCF-ChIP	CTS8	chr11:2019689-2019759	71 bp	CCTGTCTGAAGA-CCGCATGT	CCCACCGCTTGT-CAGTAGAG
CTCF-ChIP	CTCF control region 1	chr1:157465977-157466089	113 bp	CTCATTCTCAGC-CCTCACGC	CCTGTTTTTCTTT-GAAATCGTCCAC
CTCF-ChIP	CTCF control region 2	chr2:161851015-161851163	149 bp	TGCCGAGAACGT-GTGACTC	CTGCGGTATTTG-CAGCAGTA
CTCF-ChIP	2 kb upstream of ICR	chr11:2017566-2017685	120 bp	GTCATGTCCTGC-TTGTACAG	TTCTCCCCACAC-GA-CTCTCT

3.4.5 Reverse transcription quantitative PCR (RT-qPCR)

In the genome-wide DNA methylation studies, reverse transcription quantitative PCR (RT-qPCR) was conducted to investigate the efficiency of the shRNA-mediated TET1-3 knockdown efficiency. Furthermore, the transcript levels of the inducible ZnF-3AC construct, and the expression dynamics of selected genes after the introduced promoter methylation was analyzed. RT-qPCR is a highly sensitive technique to determine the relative levels of specific transcripts by normalization to a housekeeping gene and enables to quantify the expression changes after an experimental treatment (Bustin 2000).

Prior to RT-qPCR, total RNA was isolated from cell pellets and applied for complementary DNA (cDNA) synthesis as explained in detail in Appendix 1. The data was normalized as proposed by Pfaffl (2001). For more information about the RT-qPCR setup, refer to Appendix 1. The cycling conditions used for all PCR runs are displayed in Table 10, and a list of the utilized primers can be found in Table 11.

Table 10: Cycling conditions used for RT-qPCR experiments

Temperature	Time	
95°C	3 min	
95°C	15 s	40x
57°C	20 s	
Measurement		
95°C	10 s	
65 - 95°C (0.5°C steps)	5 s/step (melt curve)	

Table 11: List of primers used for RT-qPCR in genome-wide methylation studies. Primers marked with an asterisk were taken from Rawluszko-Wieczorek et al. (2015).

Gene	NCBI reference sequence	Fwd. primer	Rev. primer
RAB13	NM_002870.5	CTCGAGAGCATGGAATCCGA	AGGTCAGTACTGGGAGGCTT
ZnF-3AC	Synthetic construct	TAGACACCAACGTACTCACACCGGTC	GCAATGTAGCGGTCCACTTGG
VEGFA	NM_001025366.3	AGAAGGAGGAGGGCAGAATCA	ATGGCTTGAAGATGTACTCG
NRP1	NM_003873.7	CCTTCTGCCACTGGGAACAT	TTGCCATCTCCTGTGTGATCC
TET1*	NM_030625.3	ATACAATGGGCACCCTACCG	GGGCTTGGGCTTCTACCAA
TET2*	NM_001127208.3	GCTGACAACTCTACTCGG	CTTCTGGCAAACCTTACATCC
TET3*	NM_001287491.2	CCCAAAGAGGAAGAAGTG	GCAGTCAATCGCTATTTTC

3.4.6 Library preparation of MBD2-pulldown, NChIP and XChIP samples for next-generation sequencing

The libraries of all MBD2-pulldown samples (besides 3AC) of the genome-wide DNA methylation studies, as well as the XChIP and NChIP (H3K4me3, H3K9me3, H3K27me3, H3K36me3) samples were generated by the Max Planck-Genome-centre Cologne (MPGC). Own libraries were constructed for the 3AC and the H3K27ac samples, and the MBD2-pulldowns of the dCas9-based studies using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England Biolabs) in combination with the NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1/Dual Index Primers Set 1) (New England Biolabs) following the instructions of the manufacturer. The appropriate fragment sizes and the DNA concentration of the libraries were analyzed utilizing the LabChip® GXII Touch™ HT system (Perkin Elmer) (Figure 29). Illumina dye sequencing was conducted by the MPGC sequencing facility employing the Illumina HiSeq2500 or Illumina HiSeq3000 platform. For precise information about the respective sequencing conditions and the processing of the received data files, refer to Appendix 1 and Appendix 2.

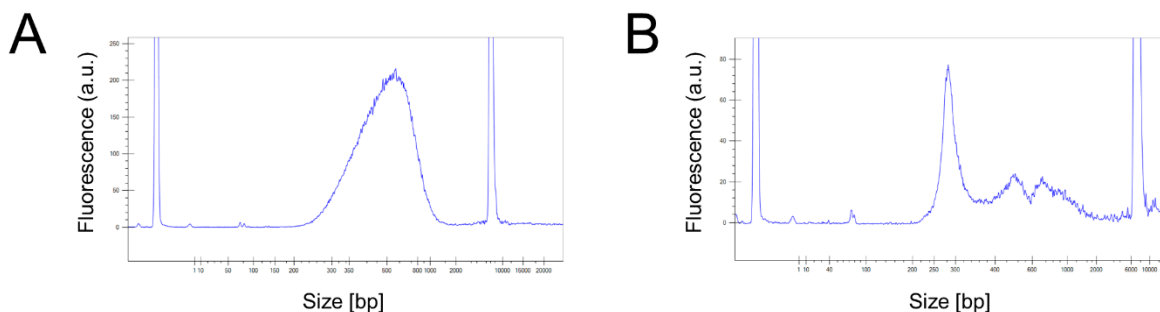


Figure 29: Quality control of constructed DNA libraries. Capillary electrophoresis was performed using the LabChip® GXII Touch™ HT system (Perkin Elmer). Exemplary electropherograms generated with the Lab-Chip GX Reviewer software are shown for A) a MBD2-pulldown sample, and B) a NChIP sample.

3.4.7 Library preparation for RNA-seq

For the genome-wide DNA methylation studies, RNA-seq libraries were prepared as described in detail in Appendix 1. In brief, 100 ng of total RNA extracted with the RNeasy® Plus Mini Kit (Qiagen) was applied for each RNA library. The NEBNext® Single Cell / Low Input RNA Library Prep Kit for Illumina® (New England Biolabs) together with the NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1) (New England Biolabs) were used for library preparation according to the manufacturer's instructions.

The quantity and quality of the in between synthesized cDNA, as well as the final RNA libraries were analyzed using the LabChip® GXII Touch™ HT system (Perkin Elmer) (Figure 30). The samples were sent to the MPGC sequencing facility for Illumina dye sequencing. A detailed description of the sequencing conditions, and the data processing of the obtained FASTQ files is provided in Appendix 1.

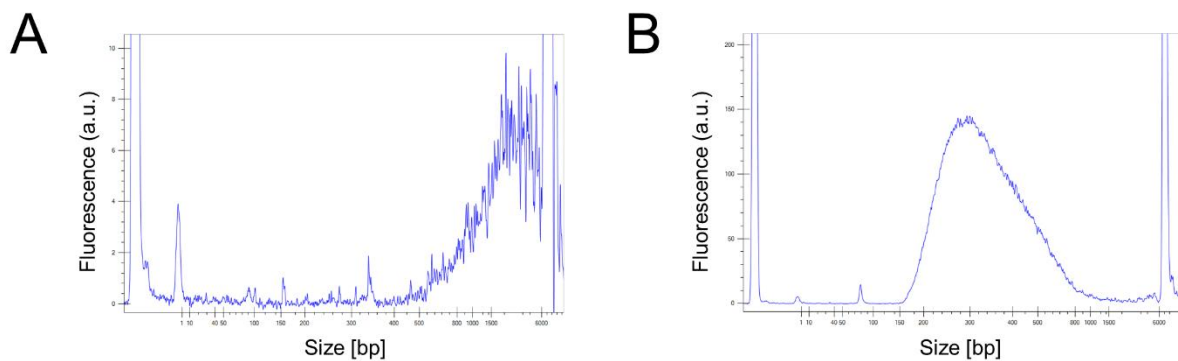


Figure 30: Quality control of constructed RNA libraries in genome-wide methylation studies. Capillary electrophoresis was performed using the LabChip® GXII Touch™ HT system (Perkin Elmer). The electropherograms were generated with the LabChip GX Reviewer software and show representative examples of A) the in between synthesized cDNA, and B) the final RNA library used for Illumina dye sequencing.

3.4.8 Bisulfite sequencing

Bisulfite sequencing is a method which has been established by Frommer et al. (1992). It is regarded as the gold standard for the determination of cytosine methylation, independent of its sequence context. Unmethylated cytosine is converted to uracil by bisulfite treatment and then recognized as a ‘thymine’ during PCR amplification and subsequent sequencing. In contrast, 5mC is protected against the sodium bisulfite-induced conversion, and thus is read as a ‘cytosine’ (Gouil and Keniry 2019).

In the genome-wide DNA methylation studies, 500 ng of sonicated gDNA, originally prepared for the MBD2-pulldown experiments (chapter 3.4.1), was converted using the EZ DNA Methylation Lightning Kit (Zymo Research) according to the manufacturer’s instructions. This was followed by the amplification of the regions of interest, as well as next-generation sequencing performed by the MPGC, and bioinformatic downstream analysis of the FASTQ data files on the European Galaxy web platform (Afgan et al. 2018). For a detailed overview about the individual steps of the experimental workflow, refer to Appendix 1.

4. Results

4.1 Dynamics of the epigenome network after global DNA methylation editing

The main project of this doctoral thesis was based on preliminary work of Dr. Goran Kungulovski. He successfully established a targeted DNA methylation delivery system which can methylate the promoter of the Vascular Endothelial Growth Factor A (*VEGFA*) with high efficiency (Kungulovski et al. 2015). In general, targeted DNA methylation has been described as a useful approach for the editing of the epigenome in regard to potential future therapeutic applications (Kungulovski and Jeltsch 2016). Consequently, it is of great importance to understand the underlying mechanisms of the epigenetic network allowing or preventing the stable propagation of the edited modification, and moreover, the associated downstream effects. For this purpose, a stable HEK293R cell line was generated, which enabled the expression of a fusion protein for DNA methylation editing in the presence of doxycycline (dox). The employed fusion protein consisted of the catalytic domain of murine DNMT3A (3AC), serving as effector domain, and a zinc finger (ZnF) targeting module, designed to bind a 9-bp motif. The short target motif of the ZnF enabled the fusion protein to bind promiscuously within the genome, followed by the introduction of DNA methylation at numerous genomic sites. Thus, the dynamics of artificially introduced DNA methylation as well as the interplay with histone PTMs, and its implications on gene expression could be investigated (Figure 31A).

To identify the actual binding sites of the ZnF within the genome, XChIP was conducted. Subsequently, ZnF binding could be correlated with the levels of *de novo* DNA methylation introduced by the ZnF-3AC fusion protein at these regions. After the setting of this modification, its stability differential stability could be investigated at several experimental time points and on a genome-wide scale (Figure 31B). This was accomplished by the establishment of a MBD2 pulldown assay which provides quantitative information about methylation changes on a genome-wide scale (Aberg et al. 2020). TET enzymes have been described as demethylases involved in the active removal of the 5mC modification. Thus, their effects on DNA methylation stability was examined in knockdown experiments of individual TET enzymes by employing small hairpin RNAs and, as a different strategy, by the inhibition of all three enzymes in parallel.

Since DNA methylation has been connected to processes involved in gene regulation, RNA-seq experiments were conducted. By this means, the dynamics of artificially introduced promoter DNA methylation could be correlated to the concomitant alterations in gene expression of associated genes. In a next step, the effects of DNA methylation on the levels of selected histone post-translational modifications (PTMs) were analyzed in a time-resolved manner by performing NChIP experiments. Moreover, the chromatin states of specific genomic loci could be determined which potentially permitted the stable preservation of DNA methylation over multiple cycles of cell division. The main findings of the aforementioned experiments will be summarized in the following chapters. This work has also been published in the Journal 'Nucleic Acids Research' and the publication has been attached to this thesis as Appendix 1.

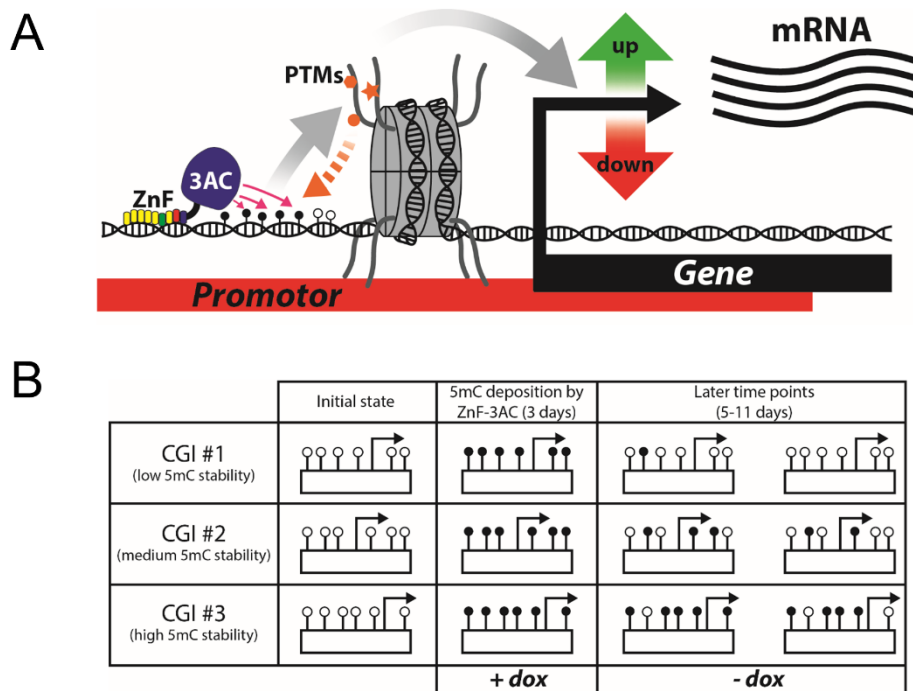


Figure 31: The stability and downstream effects of ZnF-3AC-mediated DNA methylation. A) Schematic representation of the effects of DNA methylation editing on the epigenetic network investigated in this work. The catalytic domain of DNMT3A (3AC) is recruited to a promoter through the fusion to a zinc finger (ZnF) module. Previously unmethylated CpGs (white lollipops) are getting methylated (black lollipops) by nearby ZnF-3AC (indicated by purple arrows). Effects of DNA methylation on histone post-translational modifications (PTMs) and gene expression are indicated as grey arrows and opposed effects of PTMs on DNA methylation stability as orange arrow. The potentially induced changes in gene expression are displayed as green and red arrows. B) Differential 5-methylcytosine (5mC) stability at three types of CGIs after targeted DNA methylation editing. ZnF-3AC expression is induced for 3 days by the addition of doxycycline (+dox) and then removed (-dox) for up to 11 days. Unmethylated CpG sites are indicated as white, methylated as black lollipops. The black arrow indicates the transcriptional start site.

4.1.1 Analysis of the ZnF binding properties

The ZnF utilized in this study was originally designed to bind a 9-bp motif (5'-GGGGGTGAC-3') located within the *VEGFA* promoter (Liu et al. 2001). Due to the short length of this motif, thousands of potential additional binding sites for the ZnF were expected. In order to investigate the binding properties of the ZnF module, which include the specificity, genomic localization, and the respective binding strength, XChIP was selected as an accurate and quantitative method (Carey et al. 2009). Therefore, the ZnF was fused to a triple hemagglutinin (HA)-tag, and cloned into a pAdTrack-CMV-based mammalian expression vector in which the ZnF is under the control of a constitutively active CMV promoter. HEK293 cells were then transiently transfected with the vector, harvested after three days, and prepared for XChIP as described in detail in Appendix 1. Subsequent qPCR analysis showed a 19-fold enrichment of the ZnF at the targeted *VEGFA* promoter compared to the *FZD10* promoter which does not contain the ZnF motif (Figure 32A).

The ChIP samples were sent to the MPGC sequencing facility for next-generation sequencing (NGS) in order to analyze ZnF-binding on a genome-wide scale. The ChIP-seq data tracks showed strong and defined signals with low background at different promoters carrying the binding motif of the ZnF (Figure 32B). In the case of the *NRP1* promoter, a second smaller peak was observed containing a truncated version of the ZnF motif at its summit (5'-GGGGGTG-3'). The same applies for the region upstream of the *VEGFA* promoter, where a motif similar to the designed motif (5'-GGGGTCAC-3') is present at a smaller peak. This did indicate some off-target binding of the ZnF with reduced affinity. Overall, 15,279 ZnF peaks were called which were highly enriched at CpG islands (CGIs) as well as at 5'UTRs and promoters (Figure 32C). Sequence analysis of the peaks with the DREME tool led to the identification of the originally designed ZnF motif with relaxed specificity at position 2 and 7 (E -value = 6.8×10^{-902}) (Figure 32D) (Bailey et al. 2009). Moreover, a truncated version of the designed motif (5'-HGGGGTG-3') was highly enriched as well (E -value = 2.5×10^{-241}).

These data demonstrate the capability of the ZnF to bind its intended motif with reduced specificity. Since the ZnF was chosen as a targeting module with promiscuous binding properties, the lack of specificity was rather a desirable feature than an issue.

It is interesting to note that the ZnF showed strong enrichment at gene regulatory elements close the transcriptional start sites of the genes, which are important regions for epigenome editing approaches. Based on these findings, targeted DNA methylation could be performed with the critical background knowledge about potential binding sites for the ZnF-fused methyltransferase.

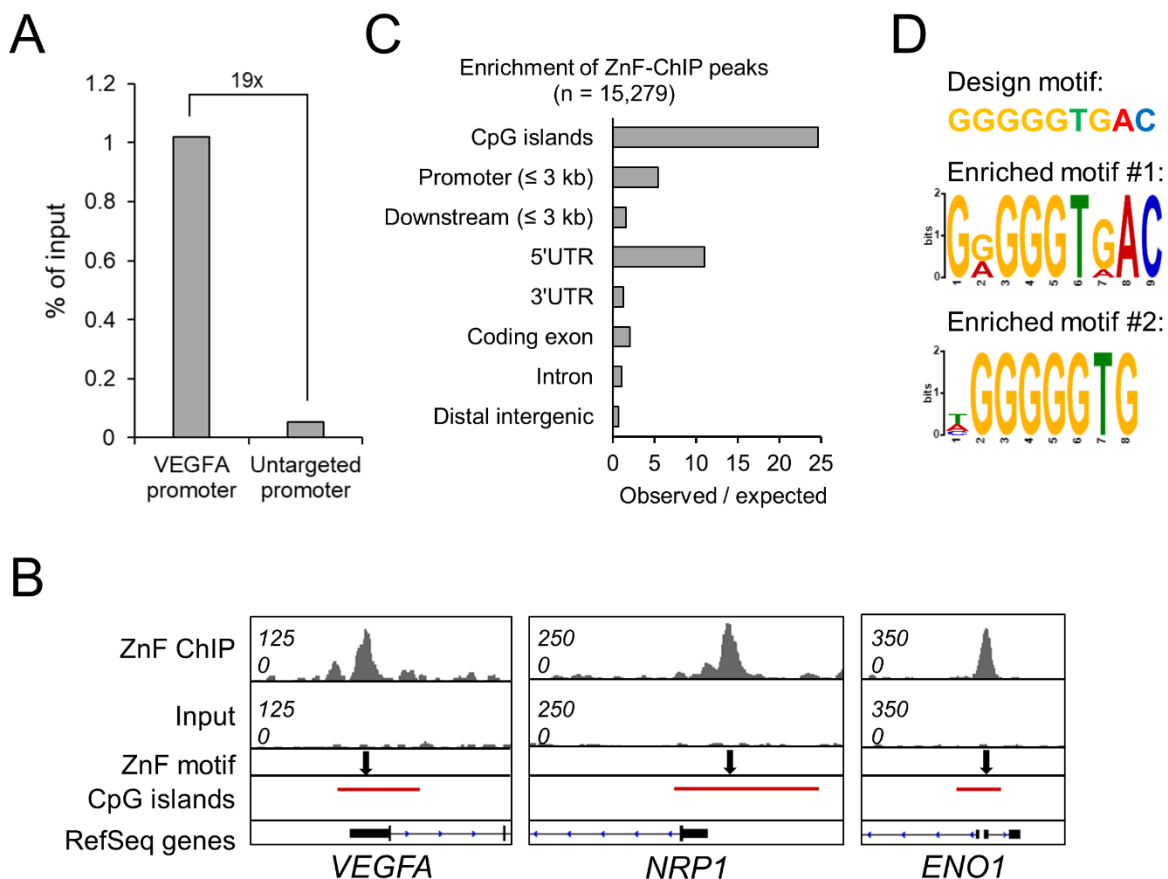


Figure 32: Cross-linked ChIP of the ZnF fused to a triple hemagglutinin (HA)-tag reveals genome-wide binding to its designed sequence with reduced specificity. A) ChIP-qPCR results of the targeted VEGFA promoter and the untargeted promoter of the FZD10 gene. ChIP signals were normalized relative to input (n = 1 repli-cates). B) Representative browser views of the ChIP-seq data of HA-fused ZnF and the corresponding input. Browser views generated using the Integrative Genomics Viewer (IGV) and the promoter regions of the VEGFA (chr6:43735154-43742227, hg19), NRP1 (chr10:33619838-33627164, hg19) and ENO1 (chr1:8935063-8941166, hg19) genes are shown. The position of the ZnF motif (5'-GGGGGTGAC-3') is indicated with a black arrow. C) ZnF-ChIP peaks were called using the MACS2 tool and their enrichment in genomic elements was determined by the CEAS tool (<http://cistrome.org/>) (Liu et al. 2011). D) The peaks of the ZnF-ChIP were analyzed for motif enrichment using DREME. The motifs showing strongest enrichment are compared to the originally designed motif of the ZnF. Panel C) and D) were adapted from Broche et al. (2021).

4.1.2 Genome-wide methylation dynamics after induced ZnF-3AC overexpression

Targeted DNA methylation for the analysis of the dynamics of DNA methylation has already been applied by various groups. However, the technical approach for the delivery of the mark, as well as the cellular systems and the respective target regions for DNA methylation editing strongly differed between these groups. Hence, it is no surprise that contradictory observations were made regarding the *de novo* methylation efficiency and its stability after installation. While some studies showed that the introduced DNA methylation was not maintained over longer periods of time (Amabile et al. 2016; Kungulovski et al. 2015; Vojta et al. 2016), others reported its stable establishment or persistent proficiency to silence a gene of interest for several weeks (Bintu et al. 2016; Stolzenburg et al. 2015). These diverging findings left many questions unanswered, and the studies often focused on a low number of target genes or artificial reporter systems. Therefore, a more global view on DNA methylation dynamics was of great interest.

In various studies, targeted DNA methylation was introduced by transiently transfecting the cells with plasmids coding for the respective constructs using cationic polymers or lipids (Stepper et al. 2017; Vojta et al. 2016). However, this vector delivery method displays multiple undesired side-effects, as it causes severe changes in the genome, transcriptome or the cellular phenotype (Stepanenko and Heng 2017). For this reason, the generation of a stable HEK293 cell line was selected as an alternative approach which allows the doxycycline (dox)-inducible expression of ZnF-3AC through a TRE3G promoter (Figure 33A). In this highly sensitive Tet-On gene expression system, a rtTA3 protein is constantly expressed within the cell. By addition of dox, a conformational change occurs within the protein enabling it to bind to the tetracycline responsive promoter (TRE), which, in turn, is comprised of a minimal cytomegalovirus (CMV) promoter and seven TetO repeats (Das et al. 2016; Markusic et al. 2005). As a consequence, ZnF-3AC expression could be controlled in a time-dependent manner. The expression of the construct could be monitored by the co-expression of GFP over an internal ribosomal entry site (IRES). The experiments were designed to induce ZnF-3AC expression for three days ('3d dox') by the addition of dox (Figure 33B). At day three, cells were enriched by fluorescence-activated cell sorting (FACS). A fraction of

them was harvested and further processed, while another fraction was re-cultured under dox-removal condition and harvested at later time points ('5d/8d/11d off'). Additionally, a sample without dox treatment was harvested as a reference for the initial epigenetic state of the cells ('no dox'). Dox induction of ZnF-3AC enabled the fusion protein to globally bind its intended targets and introduce DNA methylation for up to three days. In the further time course of the experiment, the *de novo* methylation could either be maintained or get lost.

There is a wide variety of methods available for the genome-wide analysis of methylation levels. Amongst the most commonly used ones is whole-genome bisulfite sequencing (WGBS) which provides information of methylation levels at single-CpG resolution, albeit it relies on a high sequencing coverage which involves a lot of resources. Besides methylated DNA immunoprecipitation sequencing (MeDIP-seq) utilizing a specific antibody against methylated single-stranded DNA, methyl-CpG binding domain sequencing (MBD-seq) has been employed as a cost-efficient method with medium resolution and an analysis workflow similar to ChIP-seq (Aberg et al. 2020; Jeltsch et al. 2019). In the course of this work, a protocol has been established which is described in detail in chapter 3.4.1. For this technique, the MBD domain of Methyl-CpG-binding domain protein 2 (MBD2) was utilized due to its high specificity and affinity for 5mC (Mellén et al. 2012). Fusion to a glutathione S-transferase (GST) tag facilitated the enrichment and purification of MBD2-bound methylated DNA. Afterwards, the DNA was either analyzed by quantitative PCR (qPCR) or submitted to the MPGC facility for library generation and NGS. Subsequently, the retrieved data files were processed as described in Appendix 1.

In the first step of the analysis, the MBD2-seq data derived from untreated HEK293 cells ('no dox') were compared side-by-side with ZnF-ChIP-seq data. The analysis focused on the signals in CpG islands (CGIs), as they are supposedly the most relevant targets for DNA methylation in this experimental setting. Conspicuously, the strongest ZnF enrichment was observed in unmethylated CGIs whereas heavily methylated CGIs were fully depleted in ZnF binding (Figure 33C). This moderate to strong negative relationship between the MBD2-seq and ZnF-ChIP-seq signals is expressed by a Spearman's correlation factor of $\rho = -0.49$. In respect of the goal of this study to increase methylation at unmethylated loci, the lack of ZnF binding to pre-methylated CGIs did not constitute a serious constraint. Kungulovski et al. (2015) previously

demonstrated the successful introduction of DNA methylation at close proximity to the ZnF binding site in the *VEGFA* promoter using the same ZnF-3AC construct, but using an adenoviral delivery system. For the purpose of comparability, the *VEGFA* locus was also tested in the stable HEK293 cell line setting with dox-inducible ZnF-3AC expression. Therefore, MBD2-pulldown experiments followed by qPCR (MBD2-qPCR) were

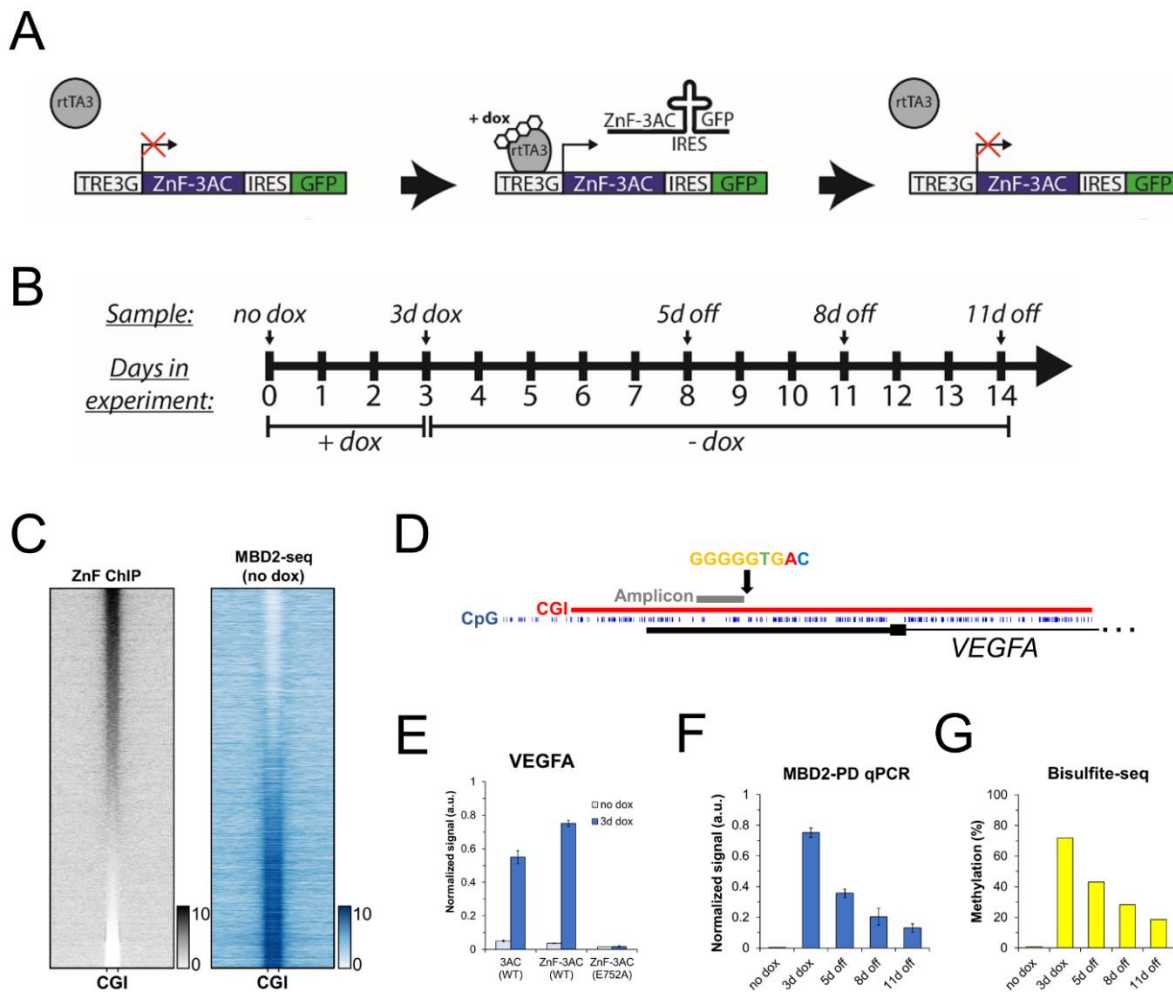


Figure 33: Targeted DNA methylation by the zinc finger-fused catalytic domain of DNMT3A in HEK293R cells. A) Stable HEK293R cell line containing a doxycycline (dox)-inducible TRE3G promoter controlling the expression of the catalytic domain of murine DNMT3A (3AC) fused to a zinc finger (ZnF). Green fluorescent protein (GFP) is co-expressed over an internal ribosomal entry site (IRES). In presence of dox, the reverse tetracycline-controlled transactivator 3 (rtTA3) binds the TRE3G promoter and induces ZnF-3AC expression together with GFP. In absence of dox, rtTA3 is released from TRE3G promoter and ZnF-3AC expression is disabled. B) Schematic depiction of the experimental workflow for the induction of ZnF-3AC. The cells are treated for 3 days with dox ('+dox'), followed by the withdrawal of dox ('-dox') for up to 11 days. C) Heatmaps of ZnF-ChIP, and MBD2-seq data in CGIs with 5 kb flanks, sorted by ZnF-ChIP signals. D) Schematic depiction of the *VEGFA* promoter region, and the corresponding CGI (red). Individual CpG dinucleotides are displayed as blue lines, the region of the amplicon used for Bis-Seq and MBD2-qPCR analysis in grey, and the position of the ZnF target motif by a black arrow. E) and F) Relative MBD2-qPCR signals of 3AC-based constructs at the *VEGFA* promoter before and after the induction with dox. Signals were internally normalized to the methylated *SLC6A3* promoter. Error bars indicate the standard deviation of two replicates. G) Bis-seq data showing the average methylation of all CpGs in the region displayed in D). Panel C) and E) were adopted from Broche et al. (2021).

conducted and the same region upstream of the ZnF target motif was amplified as used by Kungulovski et al. (2015) (Figure 33D). The *SLC6A3* promoter was chosen as a reference region for internal normalization of the pulldown in qPCR, based on publicly available reduced representation bisulfite sequencing (RRBS) data, showing it to be heavily methylated (Wang et al. 2012).

Compared to untreated cells ('no dox'), three days of dox induction ('3d dox') of ZnF-3AC led to a strong increase in the MBD2-qPCR signal (Figure 33E). Control experiments using the catalytically inactive E752A mutant of 3AC showed no increase in the signal, highlighting the requirement of a catalytically active domain for *de novo* methylation. Even 3AC without ZnF fusion caused a strong increase in methylation after induction, indicating strong activity of the untargeted domain. Nevertheless, MBD2-qPCR signals were still 1.37-fold higher for ZnF-3AC than for 3AC alone, despite of the even higher expression levels of the untargeted 3AC (Appendix 1 Figure S7B). Regarding the stability of the introduced methylation at the *VEGFA* promoter, Kungulovski et al. (2015) observed a rapid and gradual loss within 10 days with final methylation levels close to the initial state. A similar trend also appeared in the time course experiments of this study. Compared to '3d dox', the MBD2-qPCR signals were reduced by 53 % after 5 days of dox removal ('5d off'), by 73 % after 8 days ('8d off') and by 83 % after 11 days ('11d off') (Figure 33F). Since bisulfite sequencing (Bis-seq) is regarded as the gold-standard technique for methylation analysis, the dynamics of DNA methylation at the *VEGFA* promoter was also tested with this method. Starting with uninduced cells, an average methylation of below 1 % was detected (Figure 33G). Induction of ZnF-3AC lead to 72 % DNA methylation at '3d dox', followed by 43 % at '5d off', 28 % at '8d off', and 19 % at '11d off'. This means a relative reduction of 74 % between '3d dox' and '11d off', which is in good agreement with the results obtained by MBD2-qPCR.

With the confirmation that the self-developed MBD2-pulldown protocol enables the quantification of relative methylation levels at a defined locus, the next analysis focused on the genome-wide methylation dynamics within CGIs. Data from all previously mentioned experimental time points of ZnF-3AC as well as 3AC at '3d dox' were internally normalized as described in Appendix 1 Figure S3A. To assess the quality of normalization, the ZnF-3AC data were compared to Bis-seq data at selected loci (Appen-

dix 1 Figure S1A-C). Like in the case of MBD2-qPCR, the MBD2-seq data closely reflected the dynamics observed by Bis-seq. Based on the signals in the dataset of 'no dox', all CGIs were split into two groups and designated as either 'methylated' ($n = 16,730$) or 'unmethylated' ($n = 10,910$). Since pre-methylated CGIs were not regarded as suitable targets for potential *de novo* methylation, the further analysis was only based on unmethylated CGIs which were expected to be more susceptible for major changes in methylation. It has to be noted that induction of 3AC was already sufficient to produce around 9300 new MBD2 peaks compared to uninduced cells. However, a supportive effect of ZnF fusion to 3AC on methylation efficiency could be ascertained by comparing MBD2-seq data of ZnF-3AC and 3AC with the respective ZnF-ChIP signals (Appendix 1 Figure 1F). For ZnF-3AC, a moderate Spearman's correlation factor ($\rho = 0.43$) was calculated compared to ZnF-ChIP, while 3AC displayed only a weak correlation ($\rho = 0.24$).

K-means clustering was applied to all initially unmethylated CGIs, and was based on the MBD2-seq signals of the samples after dox treatment and -removal. In this approach, four clusters displaying differential methylation dynamics could be identified (Figure 34A). In cluster A, *de novo* methylation levels at day 3 were comparatively low and completely lost after five days. Around 7000 CGIs (clusters B-D) showed strong *de novo* methylation after three days of dox treatment (Figure 34A+B). Cluster B contained 3257 CGIs in which methylation was almost completely lost at '5d off'. A gradual loss of methylation could be observed in cluster C ($n = 2297$). Relative to '3d dox', two thirds of the signal were already lost at '5d off', and 90 % at '11d off' (Figure 34A-C). Strikingly, cluster D ($n = 1411$) showed a relatively high methylation stability. After five days of dox removal, 56 % of the signal was still present, but it was further depleted to 26 % at '11d off'. For the different clusters, half-lives between 1.5 and 8.2 days were calculated (Figure 34B).

Control experiments showed that transcript- and protein levels of ZnF-3AC at '5d off' were comparable to those observed before dox treatment (Appendix 1 Figure S8A+B). Hence, the relatively stable methylation of multiple CGIs even at later time points could not be explained by residual ZnF-3AC within the cells. Furthermore, no major differences in GC- and CpG content were identified across all clusters (Appendix 1 Figure S9).

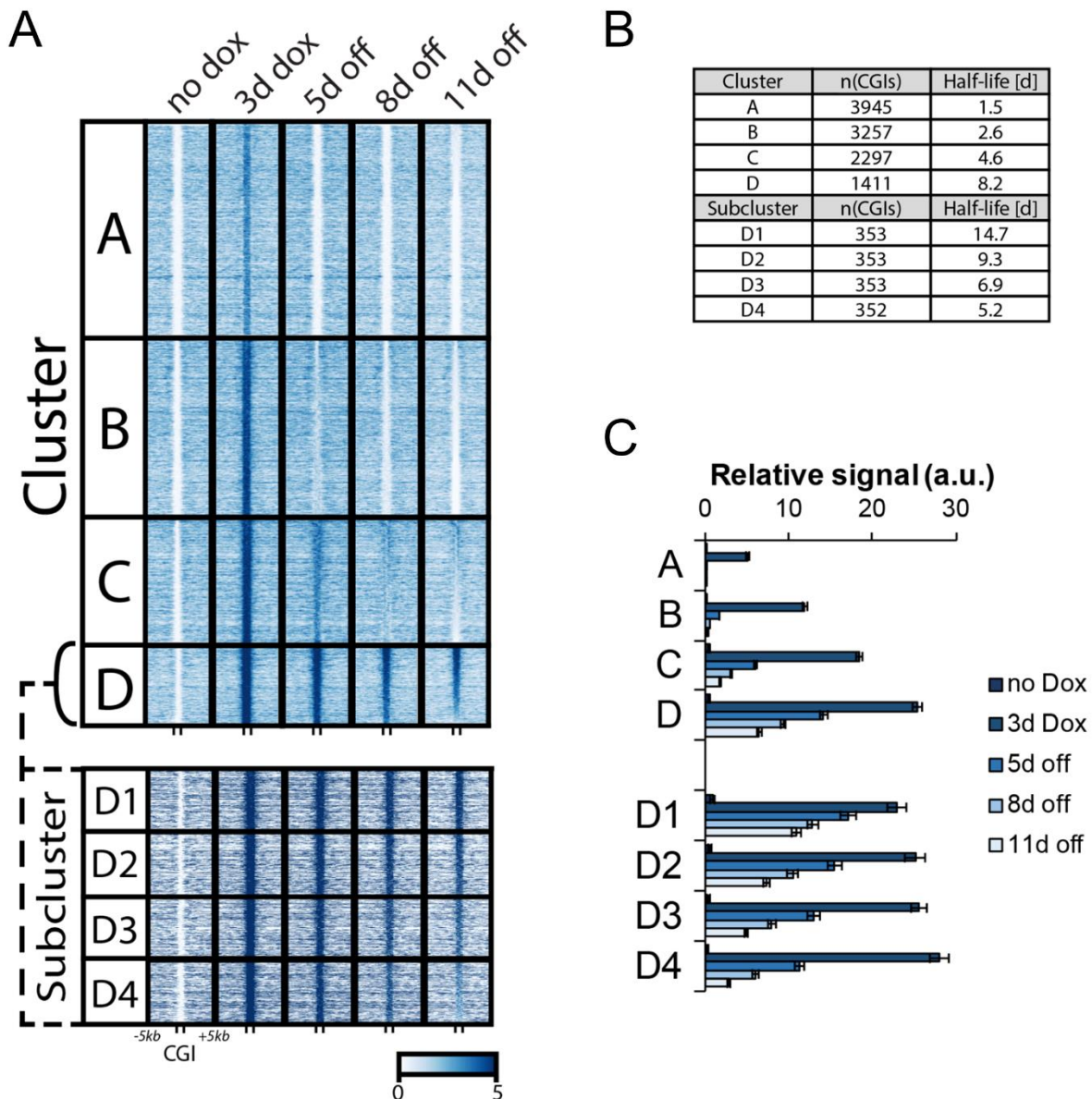


Figure 34: Clustering of CpG islands based on methylation stability after ZNF-3AC editing. A) Heatmaps of MBD2-seq data in CGIs unmethylated in initial state ('no dox') with 5 kb flanks. Expression of ZNF-3AC was induced with doxycycline (dox) for three days ('3d dox'). Differential methylation stability was tested 5, 8 and 11 days after dox removal ('5d/8d/11d off'). Clusters A-D were obtained by K-means cluster using the ChAsE software (Younesy et al. 2016). Subclusters D1-D4 were derived from cluster D and sorted according to the relative stability at 11d off to 3d dox. B) Summary of cluster sizes and methylation half-lives of cluster A-D and sub-clusters D1-D4. C) Average MBD2-seq signals of all clusters at the different experimental time points. The promoters indicate the confidence intervals calculated in Excel using an alpha value of 0.01. Panels A) and B) were adapted from Broche et al. (2021).

Overall, a rapid loss of DNA methylation was observed for around 87 % of all CGIs (clusters A-C). Consequently, cluster D with its robust methylation over long time periods was of particular interest to understand the underlying mechanisms leading to stable methylation. In order to obtain a more defined picture of cluster D, four sub-clusters (D1-D4) of equal size were generated, based on the signals at '11d off' relative to '3d

dox' (Figure 34A). Thereby, strong differences regarding the half-lives of methylation signals were observed, ranging from 14.7 days (subcluster D1) to 5.2 days (subcluster D4) (Figure 34B). The most stable subcluster D1 showed a signal which was only reduced by 21 % at '5d off' and 46 % at '11d off' relative to '3d dox' (Figure 34C). While the main clusters A-D showed a trend towards more stable methylation when the CGIs were methylated more efficiently at '3d dox', the opposite pattern is observed in the subclusters D1-D4 (Figure 34C).

4.1.3 Knockdown and inhibition of TET enzymes

The loss of the artificially introduced DNA methylation can generally be explained by different models. In the course of DNA replication during cell division, a new DNA strand is synthesized and the DNA is present in a hemimethylated state. The maintenance methyltransferase DNMT1 is able to re-establish the initial methylation pattern, but if this process does not occur, methylation is passively depleted. However, DNA methylation can also be actively removed by enzymes of the so-called ten-eleven translocation (TET) family, which is comprised of three members (TET1, TET2, and TET3) (Jeltsch and Jurkowska 2014). In regard of the results of the previous chapter, showing that the vast majority of CGIs lost *de novo* methylation rapidly, the potential influence of TET enzymes on these effects was of great interest.

In a first effort, knockdowns of individual TETs were attempted using short-hairpin RNAs (shRNAs). A control shRNA was included, targeting the *Renilla luciferase* transcript which was not present in this experimental setup. The stable ZnF-3AC cell lines were reinfected with MSCV vectors containing the respective shRNAs. After the antibiotic selection and expansion of the cells, which took ten days, the targeted DNA methylation time course experiment was conducted as previously (Figure 35A). Samples of cells were taken before and after 3 days of dox treatment as well as after 5 days and 8 days of dox removal. As first step, the knockdown efficiency of the various shRNAs was tested at the 'no dox' time point by RT-qPCR. The expression of all three TET enzymes was tested in all of these samples. For normalization, the *RAB13* transcript was chosen as it is not under the control of a CGI promoter and hence, presumably less affected by potential methylation changes caused by TET knockdown. Expression levels of *RAB13* also turned out to be very constant across all samples (data

not shown), qualifying it as a suitable gene for normalization. The confirmation of a sufficient efficiency of the knockdowns was carried out with only one replicate since it was considered as a pilot study prior to MBD2-pulldown experiments. Presence of the TET1.2851 shRNA lead to a 61- % reduction of *TET1* expression relative to Ren.660 control, while the expression of *TET2* (17 % reduction) and *TET3* (10 % reduction) remained largely unaffected (Figure 35B). Similar effects could be observed for the TET2.617 shRNA, which reduced *TET2* transcript levels by 70 %, whereas *TET1* (20- % reduction) and *TET3* (25 % reduction) showed only a mild loss in expression. In comparison, the TET3.9887 shRNA displayed weaker effects on *TET3* expression (44 % reduction), but higher specificity in regard of *TET1* (5 % gain) and *TET2* (9 % reduction) expression. Overall, in the Ren.660 control cell line, the relative expression of *TET2* to *RAB13* was 4.2-times higher than *TET1* and 89.7-times higher than *TET3*, which may suggest that TET3 plays a minor role in this cell line.

Next, the effects of the TET knockdowns on the stability of DNA methylation introduced by ZnF-3AC was investigated using MBD2-qPCR. The analysis focused on two different groups of CGI promoters: The first group (*NRP1* and *CCND2*) showed a slow and the second group (*SMURF1* and *ISG15*) a fast loss of DNA methylation signals in the previous experiments, and hence different mechanisms could be potentially responsible for the methylation loss. In the case of the first group, no significant difference in relative methylation stability could be observed between TET1-3 knockdown samples compared to Ren.660 control samples (Figure 35C). The same applies for the *ISG15* locus, where no significantly increased stability was visible. Only at the *SMURF1* promoter, a small, but statistically significant increase in signals was observed at the time point '5d off' for cells expressing TET3.9887 ($P = 0.01$). For TET2.617, a 2.4-times higher signal was determined at the same locus, which, however, was statistically insignificant ($P = 0.15$).

In summary, the knockdown of individual TET enzymes did not contribute to major changes in DNA methylation stability at the analyzed loci. Since it was unclear if this observation can be attributed to residual TET expression and/or compensatory effects of the respective other two TET family members, an alternative approach was tested.

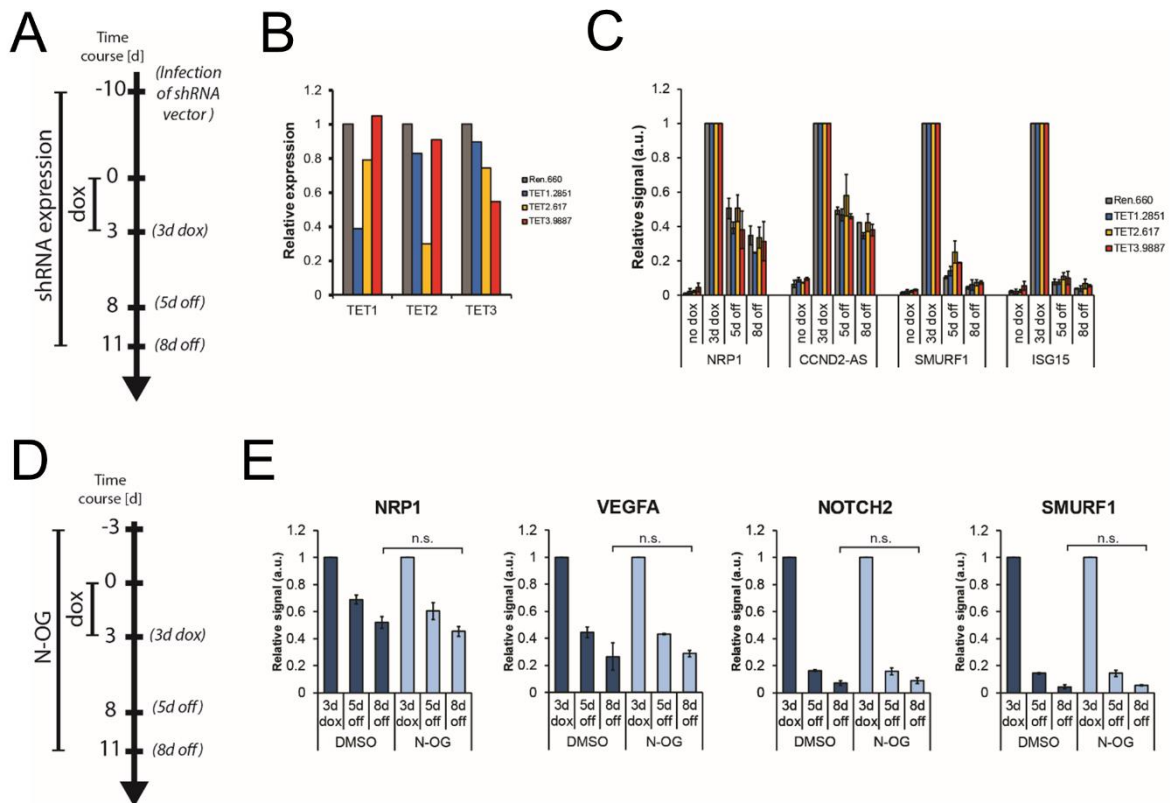


Figure 35: Examination of the role of TET enzymes on DNA methylation stability by knockdown and inhibition experiments. A) Schematic depiction of the experimental time course for the knockdown of individual TET enzymes. Infection with a vector constitutively expressing a specific small hairpin RNA (shRNA). The numbering of the shRNA refers to its binding site within the CCDS. 10 days post-infection, the time course experiment is started by induction of ZnF-3AC with doxycycline (dox) followed by removal of dox for up to 8 days. B) Expression levels of all three TET enzymes (TET1-3) before dox induction after shRNA-mediated knockdown of individual TETs. Ren.660 targets the *Renilla luciferase* transcript which is not present in this setup and serves as reference. *TET1-3* expression is normalized to *RAB13* expression and to the Ren.660 control using the Pfaffl method (Pfaffl 2001)) ($n = 1$ replicate). C) MBD2-qPCR results at different promoters after knockdown of individual TETs. The signal is normalized internally to the fully methylated *SLC6A3* locus. Error bars represent the standard deviation of two technical replicates. D) Schematic depiction of the experimental time course for the inhibition of all TET enzymes with N-oxalylglycine (N-OG). Prior to dox induction of ZnF-3AC, cells were pre-treated with 500 μM N-OG or DMSO, and treatment was continued during the whole experimental time course. E) MBD2-qPCR results of cells treated with TET inhibitor N-OG or DMSO respectively at selected promoters after induction and subsequent removal of dox ($n = 2$ technical replicates). Data representation is identical as described in C). 'no dox' = before dox treatment, '3d dox' = 3 days of dox treatment, '5d/8d off' = 5 or 8 days after dox removal.

Thereby N-oxalylglycine (N-OG) was used as an inhibitor to abolish the catalytic activity of all TET enzymes in parallel. N-OG has been reported to be an inhibitor of α -Ketoglutarate (αKG)-dependent dioxygenases, which also include the TET enzyme family, and potent TET inhibition by N-OG has also been demonstrated by several studies (Marholz et al. 2016; Peng et al. 2016; Sudhamalla et al. 2017; Sudhamalla et al. 2018).

The ZnF-3AC cell line was pre-treated with either 500 μ M N-OG or with DMSO for three days, and the treatment was continued throughout the entire time course of the experiment (Figure 35D). Samples were taken at time points '3d dox', '5d off' and '8d off' and MBD2-qPCR was conducted as described before. Four promoter CGIs, which showed differential methylation dynamics in previous experiments, were analyzed. Neither the *VEGFA* promoter, nor the *NRP1* promoter which exhibited medium to high stability before, showed a significant change in methylation stability after N-OG treatment when compared to the DMSO control (Figure 35E). The same was obtained for the *NOTCH2* and *SMURF1* promoters, where TET inhibition did not lead to a significantly higher methylation stability at later time points. Altogether, an active role of TET enzymes in the demethylation of aberrantly methylated CGIs could not be demonstrated using these experimental setups.

4.1.4 Implications of global *de novo* methylation on gene expression

Methylation of CGI promoters is generally associated with the silencing of the respective gene (Jones 2012). Thus, the direct effects on gene expression caused by global *de novo* methylation, and the dynamics of expression changes in respect to varying methylation levels was investigated. In a first step, the transcription of two heavily methylated genes (*VEGFA*, *NRP1*) was analyzed by RT-qPCR at all experimental time points. In view of the strong *de novo* methylation observed at multiple promoter CGIs of regularly-used normalizer genes (e.g. *GAPDH*, *SDHA*; data not shown), *RAB13* served as a reference gene instead, due to its lack of a CGI at its promoter. Surprisingly, *VEGFA* expression was not significantly altered after 3 days of ZnF-3AC expression (Figure 36A) despite of the highly efficient introduction of DNA methylation at its promoter (Figure 36B). This, however, did not apply to the *NRP1* gene whose expression was reduced by a factor of 6.4 after 3 days of dox treatment (Figure 36A). Moreover, the expression of *NRP1* gradually recovered after removal of dox and even surpassed the initial expression levels by 26 % at '11d off', which was not statistically significant though. To get an insight into the global changes of gene expression after DNA methylation editing, RNA-seq was conducted. RNA libraries of the experimental time points 'no dox', '3d dox', and '5d off' were generated, submitted to the MPGC

sequencing facility, and the received data files were processed as described in more detail in Broche et al. (2021).

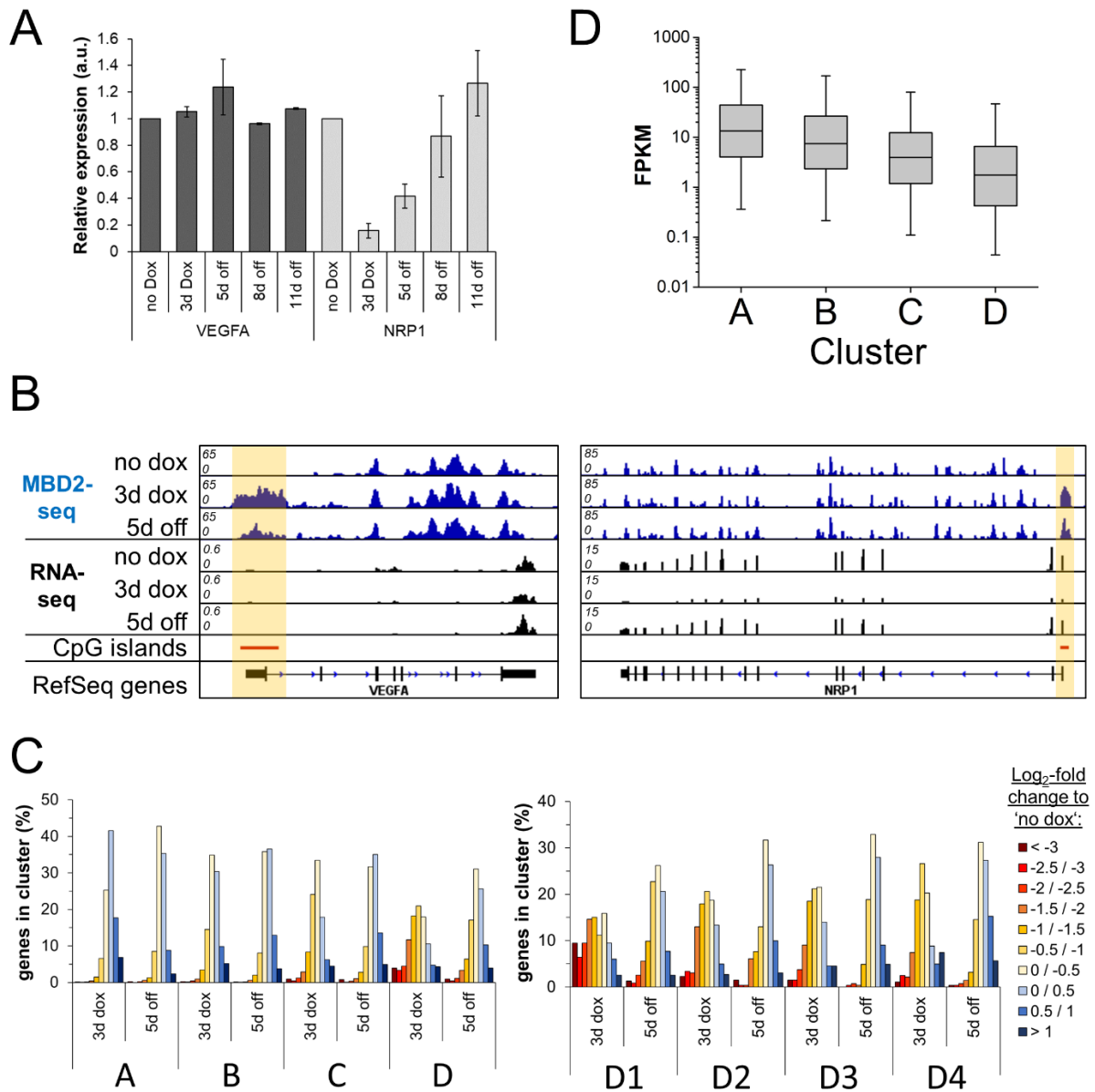


Figure 36: Implications of genome-wide methylation by ZnF-3AC on gene expression. A) RT-qPCR of the *VEGFA* and *NRP1* genes before and after ZnF-3AC induction with doxycycline (dox). Expression was normalized to the *RAB13*, and uninduced cells using the Pfaffl method (Pfaffl 2001). Error bars represent the standard deviation of $n = 2$ biological replicates. B) Browser views of the *VEGFA* (chr6:43735552-43755377, hg19) and *NRP1* (chr10:33460002-33629297, hg19) genomic loci showing the MBD2-seq and RNA-seq data at different experimental time points. The RNA-seq data consists of $n = 2$ biological replicates merged together, and is normalized to counts per million (CPM). Yellow boxes highlight the *de novo* methylated promoter regions of *NRP1* and *VEGFA*. C) Bar diagrams showing the log₂-fold expression changes within the different clusters A-D and sub-clusters D1-4 at different time points. D) Boxplots showing the expression levels of the genes associated to the CGIs from cluster A-D. Expression levels were calculated as fragments per kilobase of exon model per million reads mapped (FPKM). The line in the box indicates the median FPKM, the box indicates the 25th and 75th percentile, and the whiskers the 5th and 95th percentile. Panels C) and D) were adopted from Broche et al. (2021).

The relative abundance of transcripts was calculated in fragments per kilobase of exon model per million reads mapped (FPKM). The dynamics obtained by RT-qPCR are well reflected in the RNA-seq data, showing no major changes in *VEGFA* expression (1.2-fold reduction in FPKM) after ZnF-3AC induction (Figure 36B). However, a drastic loss of *NRP1* expression (4.9-fold reduction) after promoter methylation was observed at '3d dox' and it was still reduced by 22 % at '5d off'. For the global analysis of DNA methylation-induced expression changes, CGIs were associated to their closest gene. Afterwards, the methylation of the previously obtained clusters A-D and sub-clusters D1-4 could be correlated to the respective dynamics in gene expression. In cluster A, where only weak *de novo* methylation was detected, 91 % of the genes remained largely unaffected and showed less than a 2-fold expression change at '3d dox' (Figure 36C). Cluster B genes displayed almost the same pattern, with 90 % of genes altered less than 2-fold in expression at '3d dox' (93 % at '5d off') despite the higher efficiency in methylation at their CGI promoters. In cluster C, a slightly stronger effect was visible, as 14 % of the genes were downregulated at least 2-fold at '3d dox', and still 5 % at '5d off'. In contrast, a strong impact of DNA methylation on expression levels was observed for cluster D, where 42 % of the genes were downregulated at least 2-fold at '3d dox', and 12 % more than 4-fold. Even 5 days after dox removal, the repression was at least 2-fold for 12 % of the genes and even 4-fold for 2.4 % of them.

The most pronounced effects were noticed in sub-cluster D1, where 55 % of the genes were reduced in expression by at least 2-fold, 25 % of them more than 4-fold, and 9 % of them even more than 8-fold. Relatively stable gene repression could be achieved for a fraction of the genes, with 5 % of them being downregulated more than 4-fold even 5 days after dox removal, and 20 % more than 2-fold. Interestingly, there is a clear trend towards higher methylation stability at lowly expressed genes whereas highly expressed genes appear to be protected better against aberrant methylation (Figure 36D). This observation was also supported by chromatin state segmentation analysis (Appendix 1 Figure 3A). Taken all together, the genome-wide introduction of DNA methylation by ZnF-3AC led to extensive changes in the transcriptome, with more than 900 genes showing an at least 2-fold downregulation after 3 days. Most of the genes returned to their initial expression levels after losing the DNA methylation at their promoters. In contrast, genes with more stable promoter CGI methylation also exhibited a more lasting reduction in expression.

4.1.5 Changes in the epigenetic landscape after global DNA methylation editing

In the previous chapter, a strong influence on gene expression was observed after global DNA methylation editing. Thus, the downstream effects after ZnF-3AC-mediated DNA methylation on the epigenetic landscape were of great interest. Many reports have described the complex interplay of different histone post-translational modifications (PTMs) and chromatin-modifying proteins with DNA methylation. Some of these PTMs show a strong anti-correlation with DNA methylation (e.g. H3K4me3, H3K79me3). However, some others either co-localize with DNA methylation (e.g. H3K9me3, H3K36me3, H4K20me3), or show an ambivalent pattern (e.g. H3K27me3) (Jeltsch et al. 2018).

In this respect, the response of various histone post-translational modifications towards the introduction of DNA methylation at the same locus was investigated. In addition, the potential properties of the epigenome network facilitating the stable preservation of DNA methylation was analyzed. This could be achieved by conducting NChIP experiments for several histone marks that were either associated with active transcription (H3K4me3, H3K27ac, H3K36me3) or with gene repression (H3K27me3, H3K9me3). Cells were tested in untreated state, after 3 days of ZnF-3AC induction with dox, and up to 11 days after dox removal. For a global insight into the chromatin dynamics after DNA methylation editing by ZnF-3AC, NGS libraries were generated and processed as described in Appendix 1.

The NChIP signals were displayed in the same clusters and order as the MBD2-seq experiments. For H3K4me3 and H3K27ac, no noteworthy changes could be observed in CGIs with low DNA methylation efficiency (cluster A) (Figure 37A+B). In cluster B and C, where stronger MBD2-seq signals were detected, H3K4me3 was reduced by 15 % and 27 % respectively, but almost returned to initial levels after 5 days. The same trend occurred for H3K27ac in cluster B and C, where the signal decreased by 14 % and 21 % and then returned to the initial levels as at 'no dox'. Intriguingly, H3K4me3 signals dropped by 46 % in cluster D at '3d dox', and only gradually recovered to 87 % after 11 days of dox removal. Likewise, for H3K27ac the highest loss was observed in cluster D (39 % reduction), but the acetylation was already restored to 91 % after 5 days. Furthermore, in sub-cluster D1, where the highest relative DNA methylation sta-

bility was determined, also the strongest H3K4me3 loss was detected (56 % reduction). Concomitantly, the depletion of this modification was partially stable, and it was still reduced by 23 % after 11 days in comparison to the initial signal. Similarly, H3K27ac showed the strongest changes in sub-cluster D1, with a depletion to 50 % at '3d dox' and a partial recovery to 81 % at '5d off'.

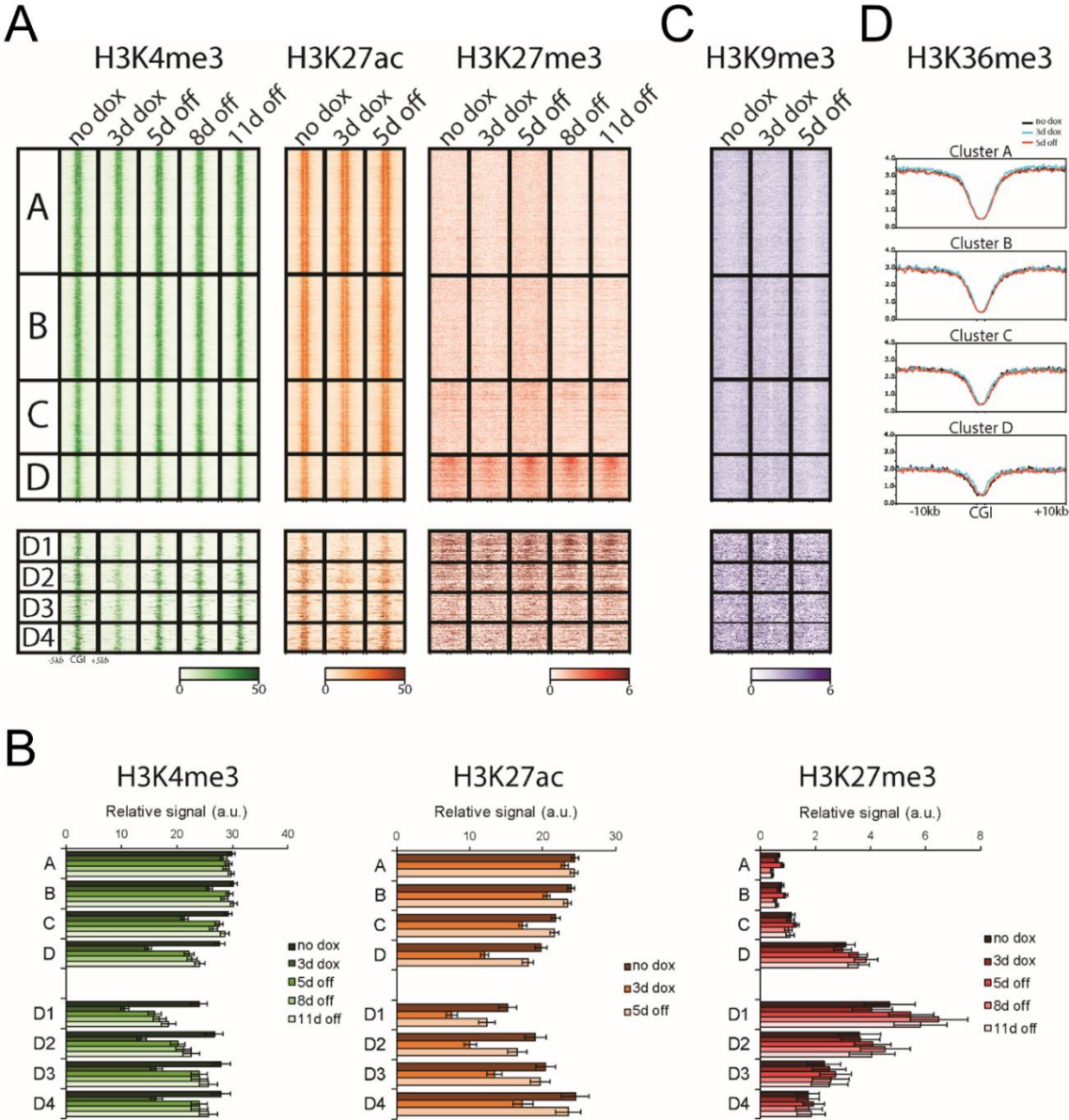


Figure 37: Enrichment and dynamics of diverse histone post-translational modifications in clusters with differentially stable DNA methylation after ZnF-3AC induction. A) Heatmaps of H3K4me3, H3K27ac, H3K27me3, and C) H3K9me3 ChIP data showing the respective signals in the previously obtained main clusters A-D and sub-clusters D1-4, after doxycycline (dox)-induction of ZnF-3AC. Heatmaps are centered around CGIs with 5 kb flanks. B) Average ChIP-seq signals of all clusters at the different experimental time points. The error bars indicate the confidence intervals calculated in Excel using an alpha value of 0.01. D) Profiles of H3K36me3 ChIP data in clusters A-D, centered around CGIs with 10 kb flanks. 'no dox' = before dox treatment, '3d dox' = 3 days of dox treatment, '5d/8d/11d off' = 5, 8 or 11 days after dox removal. Panels A), C) and D) were adapted from Broche et al. (2021).

In the case of H3K27me3, only low signals were visible in the unstably methylated clusters A-C (Figure 37A). However, a strong enrichment was observed in cluster D, without showing significant changes over time (Figure 37A+B). Moreover, sub-clusters D1-2 clearly contained the highest H3K27me3 signals. This is interesting, since these sub-clusters also previously displayed the highest relative DNA methylation stability, thus indicating a link between H3K27me3 and robust DNA methylation. Surprisingly, for H3K9me3, no enrichment could be found in any of the clusters and at any of the time points of the experiment (Figure 37C). Similarly, H3K36me3 was depleted within the CGIs across all clusters, and showed no changes after DNA methylation (Figure 37D). However, H3K36me3 was enriched in the flanking gene bodies, and higher signals were observed in the clusters that also showed higher expression levels of the associated genes at 'no dox' (Figure 37D and Figure 36D). Despite of the large alterations in expression of hundreds of genes after DNA methylation editing, gene body H3K36me3 levels remained unaffected.

Overall, the global introduction of DNA methylation by ZnF-3AC was followed by pronounced changes in the density of the activating marks H3K4me3 and H3K27ac. A clear trend between high methylation levels at different experimental time points and concomitant depletion of these histone modifications could be demonstrated. In contrast, H3K9me3 and H3K36me3 remained unaffected after DNA methylation editing. While increased stability of methylation could not be connected to the presence of the H3K9me3 at the investigated loci, a strong enrichment of H3K27me3 was observed at the most stably methylated CGIs.

4.1.6 The implications of bivalency or H3K27me3 on DNA methylation stability

In the previous chapter, the high DNA methylation stability which was observed in the sub-clusters D1-2 could be connected to strong signals of H3K27me3 in these regions. At the same time, H3K4me3 was enriched in the same sub-clusters. This raises the question, whether a bivalent state of the CGIs leads to the protection against demethylation or if H3K27me3 alone is sufficient for this effect. To clarify this question, an alternative clustering strategy was applied. CGIs were split into two groups, based on their methylation state at 'no dox'. Methylated (M1-3) and unmethylated CGIs (U1-4) were sub-clustered, taking into account the H3K27me3- and H3K4me3 signals at the

initial state (Figure 38A). The MBD2-seq signals of all time points were presented in the mentioned clusters. Interestingly, a large cluster (M1, $n = 2591$) was identified which showed high levels of H3K27me3 and DNA methylation in the initial state, indicating that the co-existence of both marks is permitted in these cells. Regarding the initially unmethylated CGIs, a bivalent cluster (U1, $n = 592$), containing strong signals of H3K4me3 and H3K27me3, and a cluster with high H3K27me3, but low H3K4me3 (U2, $n = 506$), were obtained. They were complemented by two additional clusters, displaying high levels of H3K4me3 together with either medium (U3, $n = 4107$) or low (U4, $n = 7483$) H3K27me3. The heatmaps (Figure 38A) and the quantitative analysis of the MBD2-seq signal depletion (Figure 38B) document the rapid loss of DNA methylation in clusters U3 and U4, showing similar half-lives of 5.1 days or 4.5 days respectively. In the case of the bivalent cluster (U1), a higher stability of DNA methylation could be detected ($t_{1/2} = 9.4$ days). However, an even more stable MBD2-seq signal was observed in cluster U2, showing a slightly longer half-life of 10.3 days. The fact that this cluster did not contain high levels of H3K4me3 may suggest that H3K27me3 is the determining factor supporting stable DNA methylation at these CGIs.

To further investigate the role of H3K27me3 in preserving the DNA methylation introduced by ZnF-3AC, the goal was to globally reduce H3K27me3 using the highly potent EZH2 inhibitor EPZ-6438 (Knutson et al. 2014). This was achieved by pre-treating the cells with either 10 μ M EPZ-6438 or equivalent volumes of DMSO for three days before dox induction of ZnF-3AC. Western blot analysis of whole-cell lysates at 'no dox' confirmed the strong inhibitory effect of EPZ-6438, showing 8.3-times reduced H3K27me3 levels compared to control cells (Figure 38C+D). The targeted DNA methylation dynamics experiment was carried out as previously, but under constant EPZ-6438 or DMSO treatment. MBD2-qPCR experiments were conducted to investigate the potential effects of H3K27me3 on DNA methylation stability. For this purpose, three CGIs (*CCND2-AS*, *SHH*, and *IGF2*) were analyzed which all contained high levels of H3K27me3, and showed stable DNA methylation in the previous experiments (Figure 38E). Strikingly, after 11 days of dox removal, the relative DNA methylation signals were slightly, but significantly reduced in cells treated with the inhibitor compared to control cells at the *CCND2-AS* and *IGF2* loci. A slightly reduced stability was also found at the *SHH* CGI, which however was not statistically significant. However, in this context, it has to be noted that the *SHH* locus-specific qPCR amplicon was located at the

edge of the H3K27me3 peak. These findings strengthen the previous observation, that presence of H3K27me3 at certain CGIs can support the stable introduction of DNA methylation over longer time.

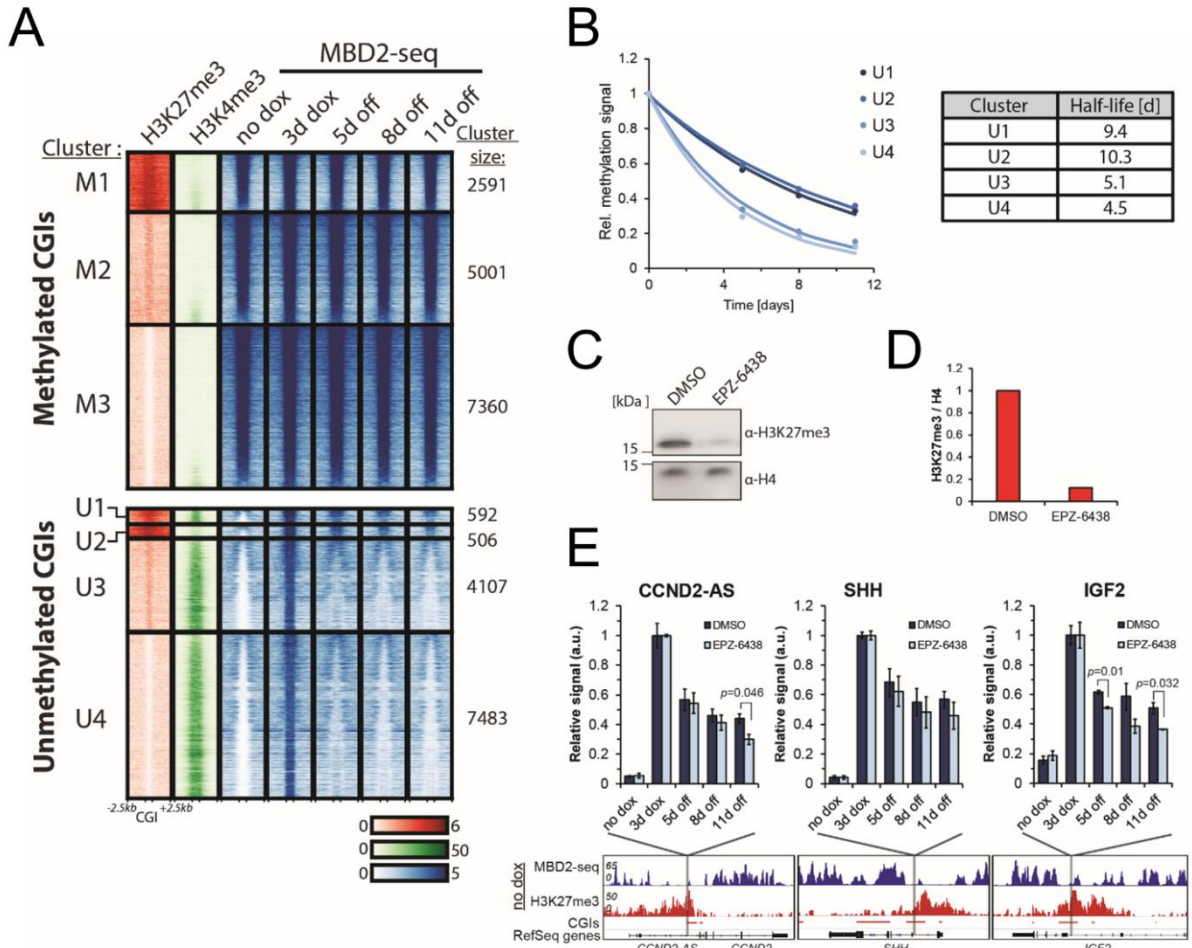


Figure 38: The influence of H3K27me3 on DNA methylation stability. A) Heatmaps were centered around CGIs flanked by 2.5 kb. Signals of H3K27me3 and H3K4me3 before dox treatment as well as MBD2-seq signals at all time points are shown. CGIs were split in two groups ('methylated' and 'unmethylated' CGIs) based on their DNA methylation state before ZnF-3AC induction with dox. *K*-means clustering was performed in both clusters, based on H3K27me3 and H3K4me3. Three clusters in methylated CGIs (M1-3) with different levels of H3K27me3 were obtained as well as four clusters in unmethylated CGIs with high (U1-2), medium (U3) and low (U4) H3K27me3 levels. B) Average methylation levels of clusters U1-4 fitted to a single exponential decay curve. The corresponding half-lives of the methylation are displayed in the table. C) Western Blot image showing the content of H3K27me3 and histone H4 in whole cell lysates originating from the stable HEK293R cell line before induction of ZnF-3AC expression. Cells were treated with either DMSO or the EZH2 inhibitor EPZ-6438 (10 μ M) for 3 days. D) Quantification of the intensity of the H3K27me3 bands relative to the histone H4 bands from C) after treatment with EPZ-6438 or DMSO using the ImageJ software. E) MBD2-qPCR results of the *CCND2-AS*, *SHH* and *IGF2* promoter CGIs after treatment with 10 μ M EPZ-6438 or DMSO. The browser views show the localization of the qPCR amplicon. The ChIP-seq data of H3K27me3 and MBD2-seq data from previous experiments at 'no dox' is displayed. Error bars represent the standard deviation of two technical replicates. The *P*-values are calculated with Excel, using the one-tailed t-Test assuming equal variance.

4.2 Engineering of an EpiEditor with improved specificity for targeted DNA methylation editing

The development of epigenome editing tools (EpiEditors) with high target specificity and efficiency, is of great interest for clinical applications and basic research. With these tools, the chromatin state of a region of interest can be reprogrammed without altering the genome sequence (Kungulovski and Jeltsch 2016). For instance, many diseases are associated with hypomethylation of various gene promoters, leading to the aberrant expression of disease-driving genes (Jin and Liu 2018). Targeted re-methylation of these promoters using EpiEditors might consequently enable the reverse the undesired epimutations. Stepper et al. (2017) could demonstrate that a single-chain construct consisting of the catalytic domain of Dnmt3a (3AC) and the C-terminal domain of Dnmt3L (3L) fused to dCas9 showed stronger methylation at various promoters than a construct lacking the Dnmt3L domain. Meanwhile, Huang et al. (2017) utilized the dCas9-SunTag system to recruit the antibody (Ab)-fused 3AC to the *HOXA5* locus, and thereby observed low levels of global off-target activity with modest to high methylation efficiency.

The aim was to develop an improved EpiEditor, which on the one hand has the potential to introduce high levels of DNA methylation at a specific target locus, and on the other hand displays a minimum of off-target activity at other genomic loci. The technical design of the project was developed in collaboration with Dr. Pavel Bashtrykov, and the obtained samples were split into two separate experimental pipelines. First, the MBD2 pulldown-based approach was performed together with the Bachelor's student Laura Laistner, and allowed a fast screening of *de novo* methylation, as well as the analysis of genome-wide methylation changes. The obtained data will be illustrated in this chapter. Second, Bis-seq was conducted by the Master's student Daniel Hofacker in order to get information about methylation efficiency at single-base resolution. Eventually, the Bis-seq results and the MBD2-seq data were published in the 'International Journal of Molecular Sciences' (Hofacker et al. 2020), which is attached as Appendix 2 in this thesis.

At the beginning, the already available targeting platforms were compared with different combinations of effector domains. Specifically, the dCas9 direct fusion system and the dCas9-SunTag system were both tested with either 3AC or single-chain 3AC-3L.

The dCas9 protein forms a complex with a single guide RNA (sgRNA) and binds to its target sequence by the formation of an RNA:DNA hybrid structure. In the case of directly fused 3AC-3L (dCas9-3AC-3L), a second EpiEditor has to be recruited through the RD interface of 3AC in order to form the catalytically active DNMT3AC/3L tetramer (Figure 39A). For dCas9-3AC, even three further EpiEditors need to be recruited through the RD and FF interfaces of 3AC to assemble an active DNMT3A tetramer at the target locus (Rajavelu et al. 2012). In comparison, the dCas9-SunTag system contains ten GCN4 peptide sequences which enable the recruitment of multiple individual Ab-fused 3AC(-3L) fusion proteins (SunTag/ Ab-3AC(-3L)) (Figure 39B). Active tetramers can, hence, either form as a consequence of Ab-mediated recruitment of two 3AC-3L dimers or by 3AC-mediated recruitment through its multimerization interfaces.

HEK293 cells were transiently transfected with plasmids coding for the EpiEditors, and the sgRNA targeting the *ISG15* promoter. Three days after transfection, positive cells, co-expressing different fluorescent markers, were enriched by FACS. Untransfected cells were harvested in parallel and served as a reference. MBD2-qPCR was conducted, and the signals were normalized to the fully methylated *SLC6A3* promoter as before. Targeted *de novo* methylation was analyzed using primers that bind in proximity to the dCas9 target sequence at the *ISG15* promoter. Furthermore, potential off-target methylation was also investigated at the untargeted *VEGFA* promoter.

Both loci showed very low initial methylation in untransfected cells (Figure 39C). All EpiEditors were able to introduce methylation at their target region. The strongest signals could be observed for the systems containing 3AC-3L, with relative MBD2-qPCR signals of 0.80 for dCas9-3AC-3L, and even 0.85 for SunTag/Ab-3AC-3L. However, for dCas9-3AC-3L the off-target methylation (0.70) was almost as high as the target locus methylation. A similar effect was observed for SunTag/Ab-3AC-3L, with a relative MBD2-qPCR signal of 0.64 at the *VEGFA* locus. In comparison, dCas9-3AC displayed only medium methylation efficiency at the *ISG15* promoter (0.44) with still relatively high off-target methylation (0.27). Interestingly, the SunTag/Ab-3AC system showed by far the lowest signals at the *VEGFA* promoter, but targeted methylation was also the lowest (0.33) compared to the other systems.

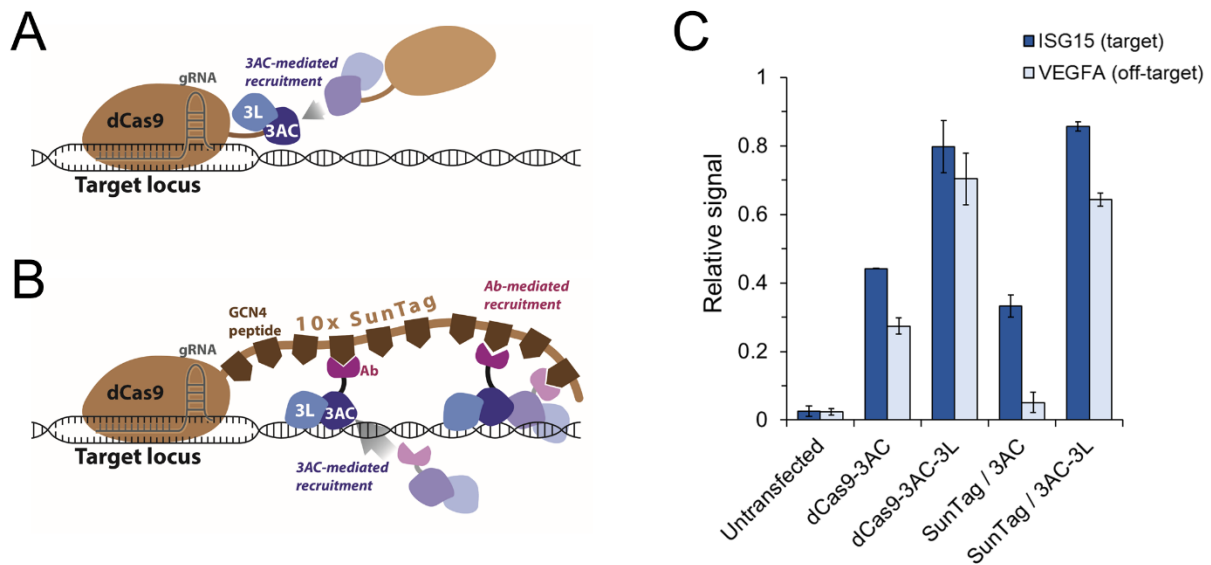


Figure 39: Comparison of different EpiEditing systems for targeted DNA methylation. A) dCas9 fused to the catalytic domain of DNMT3A (3AC) and the C-terminal domain of DNMT3L (3L). The fusion protein forms a complex with a single guide RNA (gRNA) and binds to its target locus. B) dCas9 fused to a SunTag, comprised of ten GCN4 peptides that are separated by 5 aa linkers. An antibody (Ab) is fused to single-chain 3AC-3L for recruitment to the GCN4 peptide. C) Relative MBD2-qPCR signals at the targeted *ISG15* and untargeted *VEGFA* promoters. Signals were normalized to the fully methylated *SLC6A3* promoter. Error bars represent the standard deviation of two or three biological replicates.

Based on this result, the SunTag system was chosen for further optimization due to the high methylation efficiency observed for SunTag/Ab-3AC-3L and the very low off-target activity of SunTag/Ab-3AC. Since an increase in 3AC activity appeared to be rather difficult to achieve, the focus was set on the improvement of 3AC-3L specificity without a major loss of methylation efficiency at the target locus. Therefore, the crystal structure of the catalytic domain of DNMT3A (pdb 5XY2; Zhang et al. 2018) was analyzed, and multiple basic residues on its surface were identified. These residues form a positively charged patch which can interact with the negatively charged DNA phosphodiester backbone (Figure 40A). To abolish the undesired off-target activity of Ab-3AC-3L, the rationale of the design was to mutate individual arginines (R) or lysines (K) to glutamic acid. By this, a reversal of the charge was introduced, leading to a repulsion of the effector domain from DNA (Figure 40B). In the presence of a targeting module, Ab-3AC(-3L) would be forced into a DNA contact due their spacial proximity. In a different approach, the R885 residue, responsible for the formation of the RD interface, was mutated to alanine (A). The idea behind this was to artificially bring two Ab-3AC-3L dimers together at the SunTag by Ab-mediated recruitment, joining them

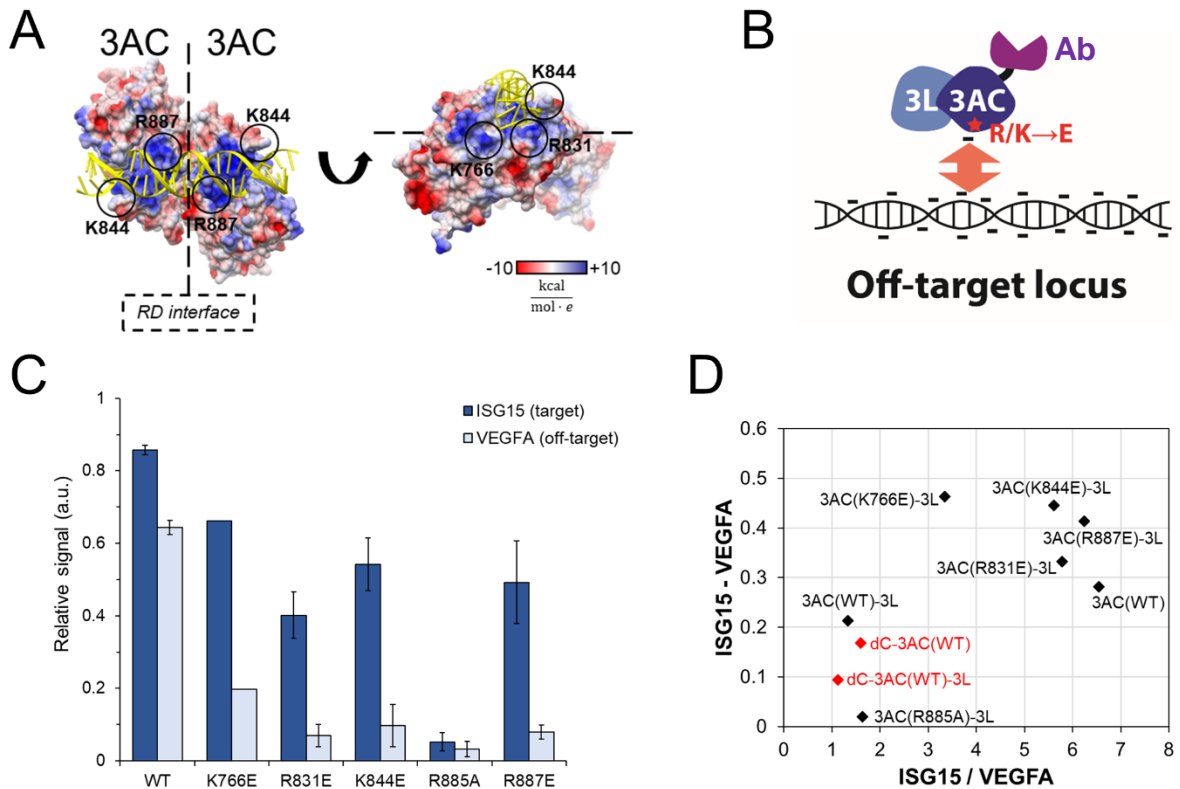


Figure 40: Engineering of a SunTag/3AC-3L system for improved specificity. A) Crystal structure of two catalytic domains of DNMT3A (3AC) (pdb 5YX2, (Zhang et al. 2018)) colored based on their surface charge (blue = positive, red = negative). The DNA is depicted in yellow, the basic residues substituted by mutagenesis (K766, R831, K844, R887) are highlighted by circles. The dashed line indicates the interaction interface between two 3ACs. Adapted from (Hofacker et al. 2020). B) Schematic depiction of the experimental design. By mutagenesis of individual arginine or lysine residues to glutamic acid, the charge of the residue within the basic patch is reversed. The affinity to the DNA is reduced due to the negative charge of the DNA phosphate diester backbone, leading to less off-target binding of the effector. C) Relative MBD2-qPCR signals of the different 3AC-3L variants at the targeted *ISG15* promoter, and the untargeted *VEGFA* promoter. Signals were normalized to the fully methylated *SLC6A3* promoter. Experiments were conducted in one (K766E), two (WT and R885A), three (K844E9 or four (R831E) biological replicates. Error bars represent the standard deviation. D) The MBD2-qPCR signal of *ISG15* was divided by *VEGFA* (x-axis), and *VEGFA* was subtracted from *ISG15* (y-axis) to assess relative the specificity and absolute gain in methylation of different EpiEditors. The EpiEditors tested in the context of dCas9 direct fusion are written in red letters, and in the context of the dCas9-SunTag system in black letters.

into a potentially active tetramer. At the same time, freely diffusing Ab-3AC-3L dimers would not be able to form a stable tetramer at off-target sites.

Transfections and downstream experiments were carried out as for the wildtype (WT) constructs previously. For the RD interface mutant Ab-3AC(R885A)-3L, no noteworthy increase in relative MBD2-qPCR signal was detected at the *ISG15* promoter (Figure

40C). However, all mutants which were designed for the reduction of DNA binding still showed medium to high residual activity at the target locus with strongly decreased off-target methylation. The Ab-3AC(K766E)-3L construct exhibited the highest activity of all mutants with a relative MBD2-qPCR signal of 0.66 at the target site, but also mild levels (0.20) of untargeted methylation at the *VEGFA* promoter. The Ab-3AC(K844E)-3L variant still showed medium to high on-target methylation (0.54), but strikingly, off-target methylation was very low (0.10). In the case of Ab-3AC(R887E)-3L, similar on-target (0.49) and off-target (0.08) methylation was observed. Finally, a large impact on the overall activity was displayed by the Ab-3AC(R831E)-3L mutant, with a relative methylation of 0.40 at the *ISG15* locus and 0.07 at the untargeted *VEGFA* promoter.

In order to evaluate the specificity and efficiency of the different EpiEditors, the ratios as well as the deltas of the methylation signals of *ISG15* compared to *VEGFA* were calculated (Figure 40D). The relative preference for on-target methylation of Ab-3AC(K844E)-3L (5.6x), Ab-3AC(R831E)-3L (5.8x), and Ab-3AC(R887E)-3L (6.2x) were almost as high, as of Ab-3AC(WT) (6.5x). However, the absolute gain in methylation was clearly greater for all of these three mutants, thus combining both, high methylation efficiency and specificity. The other variants tested in the context of the SunTag system as well as the dCas9 direct fusion system were all substantially less specific or active.

In the next step, the genome-wide off-target activity of the most promising constructs, Ab-3AC(K844E)-3L, and Ab-3AC(R887E)-3L was investigated by MBD2-seq, and compared to Ab-3AC(WT)-3L, and Ab-3AC(WT). For this purpose, DNA libraries were generated for all MBD2 pulldown samples, and sent to the MPGC sequencing facility for NGS. While the *ISG15* promoter CGI was unmethylated in untransfected cells, a major increase in methylation could be observed for all tested constructs, analogously to the differences in the signal obtained by MBD2-qPCR for this locus (Figure 41A). The same applies for the untargeted *VEGFA* promoter, which gained high levels of methylation by Ab-3AC(WT)-3L, and only low methylation by Ab-3AC(K844E)-3L, Ab-3AC(R887E)-3L, and Ab-3AC(WT). To get a more global view into off-target activity, CGIs were tested for changes

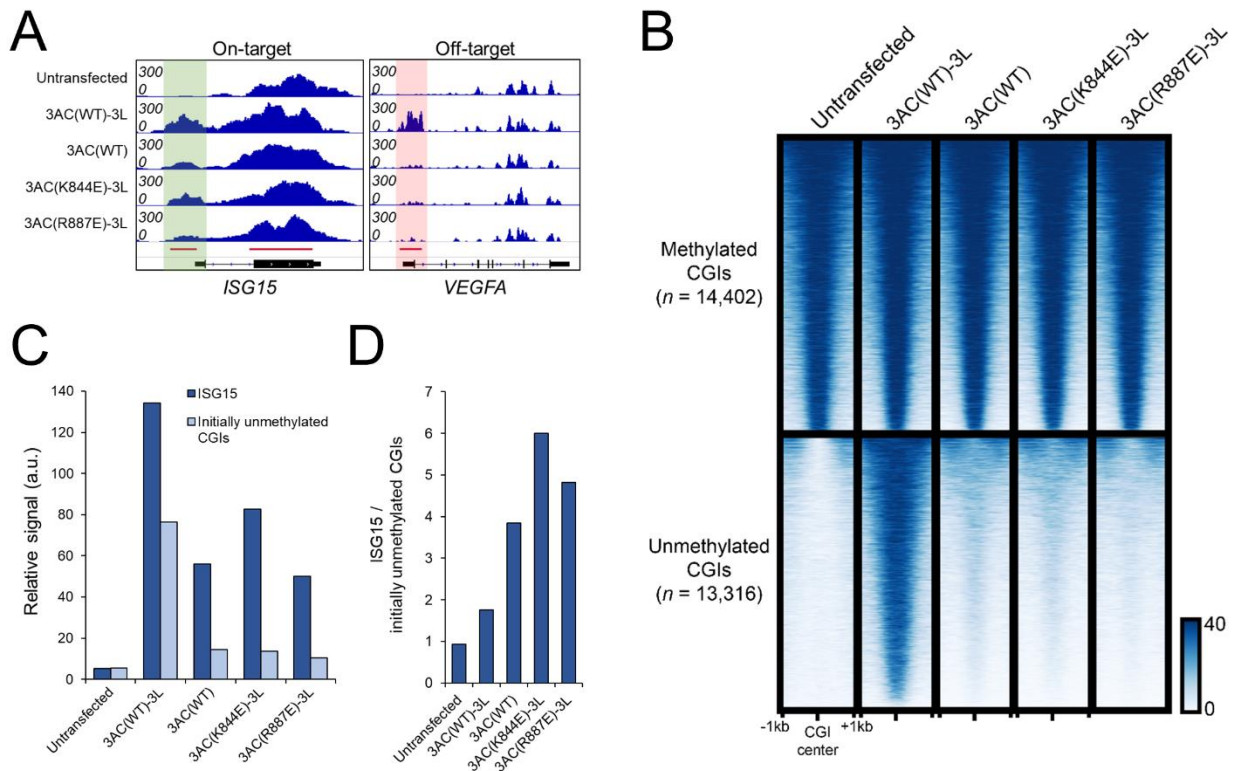


Figure 41: Reduced global off-target methylation by the SunTag/3AC-3L mutants shown by MBD2-seq analysis. A) Browser views of MBD2-seq data tracks generated with the Integrative Genomics Viewer (IGV). The targeted *ISG15* (chr1:948441-950302, hg19), and the untargeted *VEGFA* (chr6:43736149-43754992, hg19) regions are shown. The targeted CpG island (CGI) of *ISG15* is highlighted by a green, the untargeted CGI of *VEGFA* by a red shade. B) Heatmaps showing the signals of MBD2-seq data within CGIs. 1-kb flanks are displayed around the center of the CGIs. The CGIs are clustered in two groups ('Methylated' and 'Unmethylated'), based on their signals in untransfected cells. C) The signals at the *ISG15* CGI, and the average signals of all CGIs unmethylated in untransfected cells are displayed for the different samples. Signals were obtained using the multiBigwigSummary tool (Ramírez et al. 2016). D) Signals at the *ISG15* CGI relative to the average signals of unmethylated CGIs in untransfected cells calculated in C). Panels A) and B) were adapted from Hofacker et al. (2020).

in methylation after transfection, as they were regarded as important genomic regulatory regions amenable for DNA methylation. Therefore, the CGIs were subdivided into two groups, 'methylated' and 'unmethylated', based on their signals in untransfected cells (Figure 41B). Ab-3AC(WT)-3L displayed massive off-target activity at most of the previously unmethylated CGIs (Figure 41B+C). Intriguingly, Ab-3AC(K844E)-3L and Ab-3AC(WT) showed very low *de novo* methylation at these sites, and an even slightly lower average off-target signal was observed for Ab-3AC(R887E)-3L. At the target locus, methylation introduced by Ab-3AC(K844E)-3L was medium to high, while Ab-3AC(R887E)-3L and Ab-3AC(WT) showed similar, but reduced methylation efficiency at medium level. When comparing on-target (*ISG15* CGI) versus off-target (all initially unmethylated CGIs) methylation, Ab-3AC(K844E)-3L showed the largest ratio (6.0-

fold), followed by Ab-3AC(R887E)-3L (4.8-fold), and Ab-3AC(WT) (3.9-fold) (Figure 41D). In comparison, Ab-3AC(WT)-3L globally exhibited only a low preference for the target locus (1.8-fold). These data indicate that the design of the study was successful, offering a novel EpiEditing system with improved off-target activity, and adequate efficiency at the target region.

4.3 Reprogramming of the *H19/IGF2* imprinting control region by targeted demethylation

Genomic imprinting is an epigenetic mechanism for the control of genes which are involved in the regulation of growth, and is established during oocyte- and sperm development in early embryogenesis. At imprinted loci a so-called imprinting control region (ICR) is differentially methylated on the maternal and paternal alleles, mostly leading to monoallelic expression of associated genes (Zeng and Chen 2019). The *H19/IGF2* ICR is a well-studied example, in which the paternally methylated ICR is located 2-4 kb upstream of the TSS of the *H19* lncRNA. This methylation state prevents CTCF from binding to the ICR and allows an enhancer, located downstream of *H19*, to access the *IGF2* promoter and activate its expression (Figure 42A). At the same time, *H19* expression is downregulated. In contrast, the maternal allele is unmethylated at the ICR, thus enabling CTCF to bind and block enhancer accessibility to the *IGF2* promoter, while *H19* expression is activated (Nordin et al. 2014).

This project was designed in collaboration with Dr. Pavel Bashtrykov. It was also part of the Master's thesis of Claudia Albrecht, who participated in cell culture experiments, and downstream experiments equally. The ultimate goal of the project was to reprogram the *H19/IGF2* imprinting control region (ICR) by targeted DNA demethylation, facilitating CTCF to bind the previously methylated allele, and induce the loss of imprinting (LOI) of the *IGF2* gene (Figure 42A). It needs to be mentioned that LOI would not be desirable in potential therapeutic application, as it reflects the epigenetic state of the Silver-Russell Syndrome (Gicquel et al. 2005; Netchine et al. 2007). However, it serves as a proof-of-principle and could theoretically be applied for patients displaying

hypermethylation of both alleles. If these patients carried a heterozygous single-nucleotide polymorphism (SNP) at the ICR, this locus could be specifically targeted by an EpiEditor, leading to the erasure of the aberrant methylation and recruitment of CTCF.

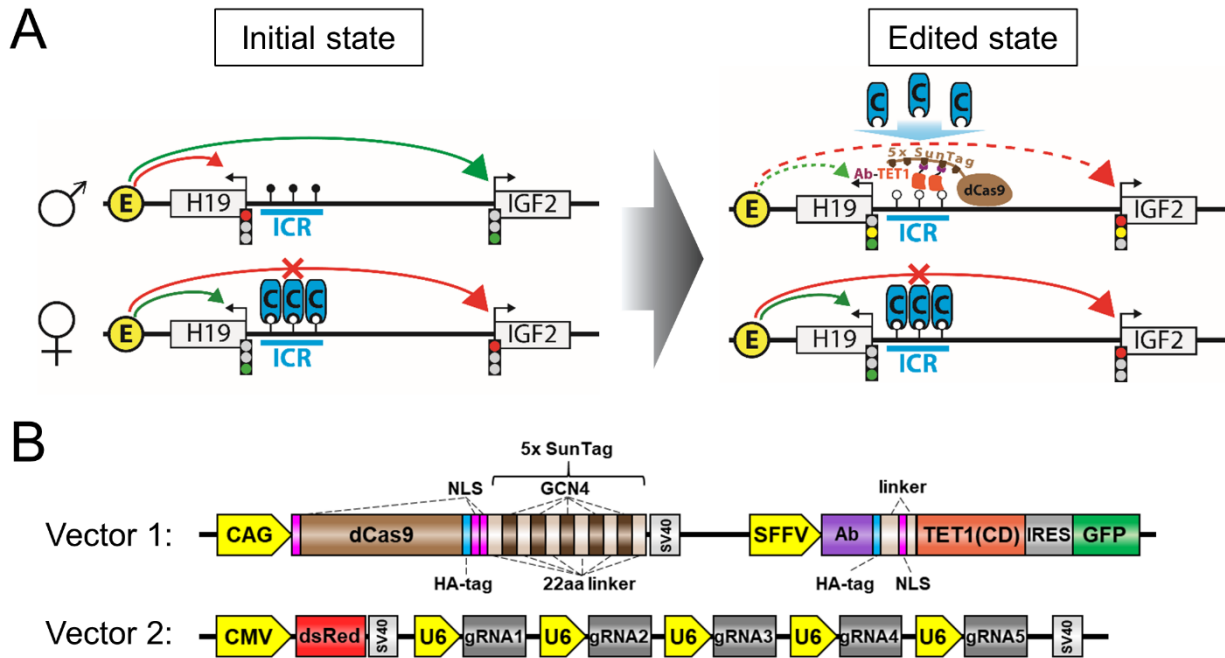


Figure 42: Allele-specific regulation by the *H19/IGF2* imprinting control region (ICR), and tools for the targeted demethylation of the ICR. A) Schematic representation of the *H19/IGF2* ICR. At the initial state, the ICR is unmethylated (white lollipops) at the maternal allele, allowing CTCF (C) to bind and block (red cross) enhancer (E) accessibility to the *IGF2* promoter. The paternal allele is methylated (black lollipops), preventing CTCF to bind and enabling the enhancer to only access the *IGF2* promoter. In edited state, the dCas9-5xSunTag/Ab-TET1 system demethylates the ICR at the paternal allele, allowing the recruitment of CTCF (blue arrow). Enhancer accessibility is indicated by a green arrow, inaccessibility by a red arrow. Changes in accessibility are shown as dashed lines, expression states of *H19* and *IGF2* by traffic lights. B) Design of the vectors used for targeted demethylation of the ICR. Yellow boxes = promoters; NLS = nuclear localization sequence; HA-tag = hemagglutinin tag; GCN4 = peptide targeted by the antibody (Ab) fused to the catalytic domain (CD) of human TET1; IRES = internal ribosomal entry site; SV40 = terminator.

For this purpose, the dCas9-SunTag system was utilized together with an antibody (Ab)-fused catalytic domain (CD) of TET1. The design of the construct was based on that of Morita et al. (2016), who developed a SunTag with five GCN4 peptides instead of ten and introduced longer linkers (22 aa) between these peptides to enable a higher flexibility of the system. Furthermore, they cloned the Ab-TET1(CD) construct, as well as the sgRNA into the same vector, by which all required components for epigenome editing are present in each individual cell after transfection. However, for this project, the technical design was adjusted by cloning only the dCas9-5xSunTag and the Ab-TET1(CD) into the same plasmid, creating a dual vector (Figure 42B). This allowed higher experimental versatility, since the sgRNA plasmid could be easily exchanged

for different targets. Moreover, GFP was co-expressed with Ab-TET1(CD) over an IRES instead of a direct fusion to the construct, to reduce potential steric hindrance and allow more flexibility of the effector domain.

In order to confirm the capability of the novel EpiEditing system to reprogram the *H19/IGF2* ICR, HEK293 cells were transiently transfected with either the dual vector containing the Ab-TET1(CD) EpiEditor, or an empty vector (EV) control without an effector domain (Ab-EV). Moreover, the cells were co-transfected with a multi-sgRNA vector coding for five sgRNAs which recruit the dCas9-5xSunTag targeting module to multiple CTCF target sites (CTS) within the ICR (Figure 43A). Three days after transfection, the cells were sorted by FACS, in order to enrich living single cells that display green (dual vector) and red (multi-sgRNA) fluorescence. A fraction of these cells was then harvested for follow-up experiments, and another fraction was re-cultivated for analysis at later time points. MBD2-seq data obtained in the course of the project from chapter 4.2 showed relatively high methylation at the ICR in untransfected HEK293 cells (Figure 43A). Furthermore, based on publicly available ChIP-seq data sets, at least eight potential CTSs with strong CTCF enrichment could be identified within the ICR. Copy number variant (CNV) analysis revealed the presence of four copies of the 11p15.5 locus in the HEK293 cell line (Lin et al. 2014).

In a first step, MBD2-qPCR was conducted to investigate the efficiency of Ab-TET1(CD) mediated demethylation at the eight CTSs (CTS1-8). As a reference, untransfected HEK293 cells and cells transfected with the Ab-EV dual vector were analyzed as well. MBD2-qPCR signals were internally normalized to the fully methylated *SLC6A3* promoter region as in the previous experiments. In untransfected cells, the normalized MBD2-qPCR signals in CTS1-8 were between 0.33 and 0.71, with a combined average of 0.54. This could be interpreted to indicate, that two out of four alleles are methylated (Figure 43B). This assumption is supported by preliminary Bisulfite-seq data obtained by Dr. Pavel Bashtrykov (data not shown). Compared to this, the signal of an untargeted control region 2 kb downstream of the ICR showed was around 1.3, which is more than two times higher than the ICR signal. Three days post-transfection with the dual dCas9-5xSunTag/Ab-TET1(CD) vector, a drastic loss of DNA methylation within the ICR could be observed (Figure 43B). The methylation at CTS1-8 decreased to relative MBD2-qPCR signals between 0.11 and 0.38. Interestingly, the weakest change was observed for CTS8, which also had the furthest distance to a sgRNA target

site (Figure 43A). On average, the DNA methylation was reduced to a signal of 0.18, thus being exactly 3x less than in parental cells, and demonstrating the highly efficient demethylation activity of the EpiEditor. It has to be noted that in the control region the methylation was also reduced by 60 % compared to untransfected cells (Figure 43B). Six days after transfection, the methylation levels in the ICR were very constant, reaching an average signal of 0.17. Interestingly, the demethylation in the control region, which is located 2 kb downstream in the *H19* promoter, was only transient and the MBD2-qPCR returned to 72 % compared to untransfected cells. Even 27 days after transfection, the low DNA methylation levels at the ICR barely changed, and showed a relative signal of 0.22. This value is still 2.5x lower than in untransfected cells. The methylation in the control region did not return to initial levels and stagnated at 68 %. Transfection with the Ab-EV control did not show drastic changes in comparison to untransfected cells suggesting that loss of methylation was not caused by dCas9-binding to its target site, but the catalytic activity of the TET1(CD) was necessary to obtain the effect.

Next, the aim was to ascertain whether the successful DNA methylation reprogramming at the ICR subsequently enabled the binding of CTCF to its target sites. Therefore, XChIP experiments with an α CTCF antibody were conducted with either untransfected cells, or cells transfected with the EpiEditing system. The CTCF occupancy was investigated 6 days, and in addition, 27 days after transfection by ChIP-qPCR to find out whether the stable reprogramming was also reflected in an increased CTCF binding over long periods of time. In order to improve the validity of the obtained qPCR results, the data of the ChIP samples were first normalized to the input, followed by internal normalization to the average of two unrelated CTCF reference sites. Both reference sites are characterized by low CpG density and strong CTCF enrichment based on the publicly available data set.

For untransfected cells, an average relative ChIP-qPCR signal of 0.42 at CTS1-8 was determined (Figure 43C). The relative signals of 0.28 to 0.58 at the individual CTSs might indicate that either one or two alleles out of the four total alleles were occupied by CTCF. No noteworthy signal was detected at the negative control region, located 2 kb downstream of the ICR. Strikingly, six days after transfection with the dCas9-5xSun-Tag/Ab-TET1(CD) EpiEditing system, the average ChIP-qPCR signal at CTS1-8 more

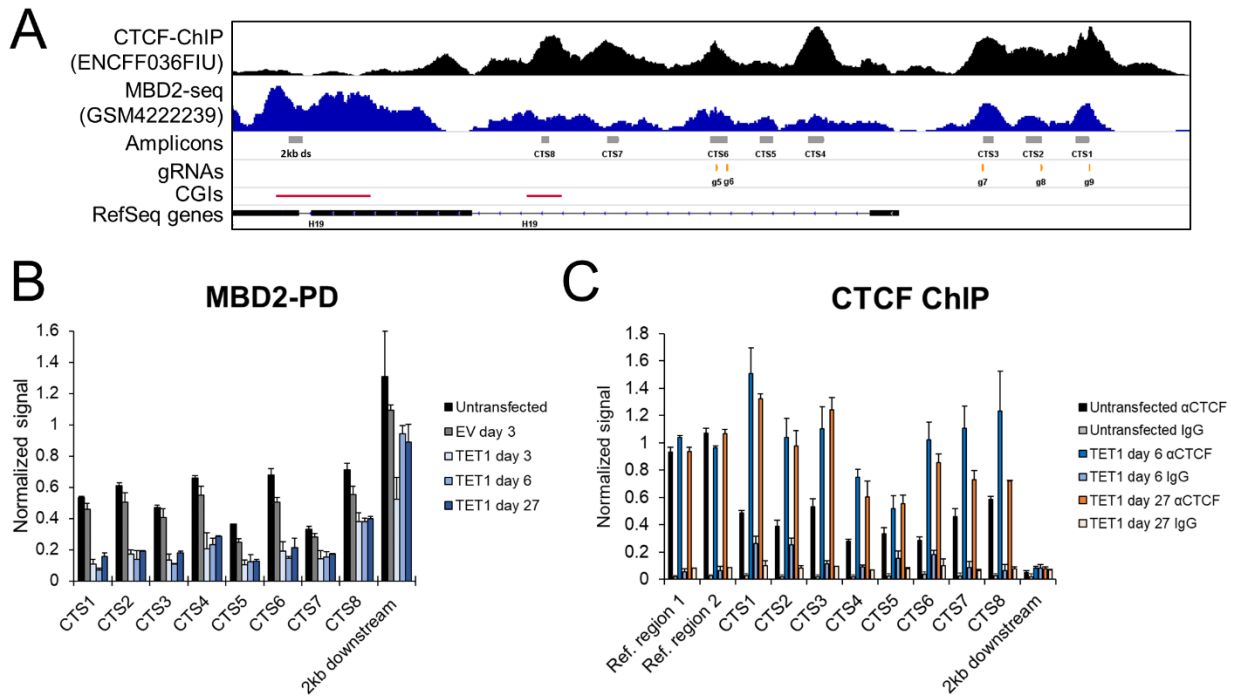


Figure 43: Reprogramming of the *H19/IGF2* imprinting control region (ICR) by targeted demethylation. The ICR contains at least eight CTS target sites (CTS1-8). A) Browser view (chr11:2017176-2025162, hg19) showing the DNA methylation and CTCF occupancy at the *H19/IGF2* ICR. The MBD2-seq data track was taken from Hofacker et al. (2020), the CTCF-ChIP data track (experiment ENCSR617IFZ) from Zhang et al. (2020). The multi-sgRNA target sequences (g5-9) are indicated by a red line. B) MBD2-qPCR data from either untransfected cells, or cells transiently transfected with TET1 or an empty vector control lacking an effector domain. Co-transfections were carried out with the novel dual vector system and a multi-sgRNA plasmid targeting five sequences within the ICR. The bar diagrams show the methylation status before and after transfection. The data was first normalized to input, then internally to the fully methylated *SLC6A3* promoter. Error bars indicate the standard deviation of 2-3 replicates. C) XChIP-qPCR data of CTCF binding to the ICR before and after reprogramming. The data were first normalized to input, then internally to the average signal of two unrelated reference (Ref.) regions. Error bars indicate the standard deviation of 2-3 replicates.

than doubled (2.5x) and ranged from 0.52 to 1.51. Even 27 days after transfection, the average relative CTCF signal at CTS1-8 was 2.1-fold higher than in untransfected cells, and remained in a similar range of 0.55 to 1.32. The observation, that elevated CTCF binding at the ICR was retained almost four weeks after transfection perfectly agrees with the detection of stably reduced methylation levels at this locus.

Further experiments are still ongoing in collaboration with Claudia Albrecht, and will be part of her Master's thesis. Thereby, the changes of *H19* and *IGF2* expression after ICR reprogramming will be investigated. Furthermore, the potential off-target effects caused by Ab-TET1(CD) expression will be analyzed by MBD2-seq. In order to gain insight into the methylation and hydroxymethylation state at single-base resolution, Bis-seq as well as oxidative Bis-Seq will be carried out, and supported by Nivethika Rajaram.

5. Discussion

DNA methylation is a major epigenetic signal and involved in a multitude of regulatory processes. The accurate establishment and maintenance of DNA methylation patterns is critical in mammalian development, and for the preservation of cell identity. Aberrations in DNA methylation are linked to numerous types of cancer and other diseases. Despite being mostly described as a ‘silencing’ mark, its function is highly context-dependent. Usually, the presence of DNA methylation at CGI promoters is associated with the silencing of the underlying gene. Besides its physical properties directly affecting the local chromatin, DNA methylation requires the crosstalk with additional factors for the induction of a repressed epigenetic state. In mammals, the enzymes of the DNMT family were identified to be responsible for the deposition of the 5mC mark. *Vice versa*, DNA methylation can also be actively removed by TET dioxygenases, or passively lost by DNA replication (Greenberg and Bourc'his 2019; Jones 2012).

To date, numerous groups have employed the catalytic domains of DNMTs and TET enzymes, and fused them to a DNA binding domain (DBD) in order to modify the DNA methylation states at defined regions. However, these DBDs strongly vary in specificity and flexibility (Kungulovski and Jeltsch 2016). The variable specificity of the targeting devices was taken as an advantage in the main project using a ZnF fusion, because it was aimed to introduce DNA methylation at a large number of genomic sites to get a more global insight into DNA methylation features. In contrast, the dCas9-based projects were directed at highly specific editing of DNA methylation states. In addition, the application of a self-developed MBD2-pulldown assay enabled the rapid screening of the modulated DNA methylation levels via qPCR, and offered the possibility to analyze the accompanied genome-wide effects by subsequent next-generation sequencing (Aberg et al. 2020). In this chapter, the findings of the diverse epigenome editing approaches will be discussed in detail, and a light will be shed on the dynamic interplay of the epigenome network with *de novo* DNA methylation. Moreover, the general pitfalls and challenges of stable and/or efficient DNA methylation editing will be discussed.

5.1 Dynamics of the epigenome network after global DNA methylation editing

Genome-wide deposition and dynamics of DNA methylation

Targeted DNA methylation is a widely used approach in basic science to gain insight into the regulatory effects of this modification on the epigenetic landscape, and on gene regulation. With its presumably high silencing potential at gene promoters, DNA methylation editing is also a promising tool for potential future therapeutic application, for instance in various types of cancer. However, a lot of progress is still required in order to implement this technique in patients in a safe and eligible way. The main concerns about epigenome editing generally include the concomitant off-target effects, the efficiency of the induced changes, the inheritability of the modified mark, and the generalizability of the approach for different targets (Kungulovski and Jeltsch 2016).

By employing a ZnF protein in fusion with 3AC, the intention of this project was to introduce DNA methylation at thousands of genomic sites. Statistically, the utilized ZnF designed to target a 9-bp motif had a binding probability of $p = 4^{-9}$, which should have resulted in around 12,000 binding sites in the haploid human reference genome Hg19 ($n = 3.14 \times 10^9$ bp). Interestingly, even more than 15,000 ZnF peaks were identified in ChIP-seq experiments, indicating a relaxed specificity of the ZnF, which was also demonstrated in Figure 32D. This observation is in good agreement with the finding of other groups, where individual fingers of two naturally occurring C2H2 ZnF proteins did not display perfect specificity (Bulyk et al. 2001; Chou et al. 2017). Hence, small mismatches in the target sequence could be tolerated by the ZnF applied in this work, and consequently enabled the ZnF-3AC construct to be directed to even more genomic sites. Albeit, the binding of the ZnF towards DNA appeared to be impaired at regions with repressive chromatin states, as the anti-correlation of ZnF-ChIP data with the MBD2-seq data at CGIs clearly indicated (Figure 33C). At CGIs with high DNA methylation levels, the chromatin is packed more tightly, which presumably affected the binding of the ZnF to the target DNA (Blackledge and Klose 2011). For this reason, the utilized ZnF might not be the best choice for projects aiming to remove repressive

chromatin marks. Nonetheless, it was a well suited DBD for genome-wide DNA methylation editing, especially due to its strong enrichment at CGI promoters (Figure 32C).

The establishment of a stable HEK293R cell line containing a dox-inducible promoter in front of the construct allowed the temporally-controlled coexpression of ZnF-3AC together with a fluorescence reporter. By this means, it was ensured that only cells with sufficient expression of the construct were enriched, without experiencing undesired side-effects from transfection reagents (Stepanenko and Heng 2017). Strikingly, the ZnF-3AC construct was exceptionally efficient in methylating its target regions, with individual CpG sites at the *VEGFA* promoter approaching up to 96 % of DNA methylation, as shown by Bis-Seq (Appendix 1 Figure S1A). Despite using the same EpiEditor construct, these values even surpassed the highest methylation signals of 76 % within the *VEGFA* promoter, obtained in the preliminary study from Kungulovski et al. (2015). One may speculate that the increased methylation efficiency can be ascribed to differential dynamics or levels in ZnF-3AC expression, as the previous study used adenoviral vectors as a delivery system for the construct (Kungulovski et al. 2015).

Importantly, DNMT3A requires the formation of a homotetramer for DNA methylation (Gowher and Jeltsch 2018). This however would imply that the introduction of DNA methylation by the DNA-bound ZnF-3AC was dependent on the further recruitment of freely-diffusing fusion proteins via its RD- and FF interfaces. Notably, also 3AC without ZnF fusion was capable to globally methylate DNA, but with less relative efficiency at ZnF-ChIP peaks than ZnF-3AC (Appendix 1 Figure 1F). This observation indicated a synergistic binding of 3AC to the DNA backbone with its active site, together with the ZnF targeting module, leading to an improved methylation efficiency compared to 3AC alone. Actually, it was not very surprising that induction of 3AC expression was already sufficient to induce genome-wide DNA methylation changes. Several studies previously showed a global increase of DNA methylation after overexpressing full-length DNMT3A (Jia et al. 2016; Zheng et al. 2017). In conclusion, the fusion of the rather unspecific ZnF domain to 3AC was beneficial to improve the methylation efficiency at thousands of CGIs.

One major goal of this work was to answer the highly controversial question, whether the artificial introduction of DNA methylation can be stably maintained over longer time

periods. Answering this question would, in general, be highly relevant for potential clinical applications utilizing targeted DNA methylation techniques. Diverging observations were made in various studies, characterizing the 5mC dynamics after targeted DNA methylation editing. In these studies, the introduced DNA methylation at genes promoters was either rapidly reduced within only a couple of days, or reported to be stable for several weeks. Similar differences were also obtained in the dynamics of reporter gene expression after the deposition of DNA methylation at the corresponding gene promoters. However, the studies considerably differed in the design of the construct, the cellular systems, the choice of the target loci, and the technical readout. For instance, several groups only indirectly analyzed the inheritance and silencing potential of DNA methylation based on the expression of a fluorescence reporter in an artificial setting (Amabile et al. 2016; Bintu et al. 2016; Kungulovski et al. 2015; Rivenbark et al. 2012).

By investigating the DNA methylation changes on a more global scale and with temporal resolution, this work provides an important basis to understand the differential stability of endogenous CGI promoter methylation. With its more than 16,000 methylated CGIs in uninduced state, the utilized HEK293R cell line reflected the CpG island methylator phenotype, which has also been described for various types of cancers or cell lines (Smiraglia et al. 2001; Toyota et al. 1999). Intriguingly, despite of the efficient introduction of DNA methylation at thousands of CGIs, ZnF-3AC-mediated DNA methylation was unstable at the vast majority of CGIs (Figure 34A), while the already pre-existing CGI methylation was completely maintained (data not shown). The observed lack of 5mC stability is supported by the previous finding about CGIs being refractory to DNA methylation due to their sequence content (Long et al. 2016). In addition, further studies reported the transient nature of *de novo* DNA methylation, or the concomitant short-term silencing of promoters after DNA methylation editing. Like in this work, demethylation of CGIs already occurred only within a few days (Amabile et al. 2016; Kungulovski et al. 2015). However, around 10 % of all methylated CGIs also displayed a medium to high 5mC stability, even two weeks after editing. This is a striking finding, encouraging for future clinical applications relying on the long-term propagation of DNA methylation. It also underlines the strong dependence on the local features of the CGIs, affecting the stability of the 5mC mark. Sequence analysis of different clusters of CGIs formed based on their level and stability of introduced DNA methylation indicated that

neither the CpG content, nor the GC density were closely related to 5mC stability (Appendix 1 Figure S9). Again, our data were supported by the results of other groups, showing the stable maintenance of DNA methylation at selected promoters for several weeks (Bintu et al. 2016; Rivenbark et al. 2012; Saunderson et al. 2017).

It remains intricate to clearly determine the cause for the differential loss of DNA methylation. In theory, passive demethylation accompanied by complete absence of the maintenance methylation machinery would lead to the halving of 5mC after each round of replication. With a DNA methylation half-life of 1.5 days, and a doubling time of around 24 h reported for HEK293 cells, cluster A could principally be assigned to this category (Figure 34B) (Cervera et al. 2011). However, as the refined stochastic DNA methylation model proposes, the varying local levels of TET enzymes and DNMTs might be the factors, defining the stability of the 5mC mark (Jeltsch and Jurkowska 2014). Thus, TET enzymes might compete with DNMT1 to either remove or maintain the introduced DNA methylation. Indeed, TET1-3 are enriched at active, CpG-rich promoters, and thereby preserve their hypomethylated state (Melamed et al. 2018). The results in Figure 36D illustrate that especially the genes associated to clusters A and B displayed high expression levels, and were thus probably protected against aberrant deposition of DNA methylation by TET enzymes.

A further interesting finding in this work was that the relative DNA methylation stability was not necessarily dependent on the introduced DNA methylation levels (Appendix 1 Figure S11A). This was in good agreement with the study from Ginno et al. (2020), showing that the local chromatin environment and TFs strongly affect the turnover of DNA methylation, instead of the initial methylation levels. They also found that loss of methylation was more attributable to TET-mediated active demethylation, than on DNMT1 infidelity (Ginno et al. 2020).

We also sought to address the potential influence of TET1-3 on the stability of the 5mC modification by conducting knockdown (KD) experiments of individual TETs. Despite of the relatively high KD efficiencies of TET1 and TET2, no strong effects on DNA methylation stability were observed. One simple explanation could be an inadequate choice of investigated loci by MBD2-qPCR. Hence, a genome-wide screening via MBD2-seq might potentially have revealed a group of CGIs displaying differential stability. However, it seems more likely that the KD efficiency was not sufficient, and the

local levels of all combined TET enzymes were still sufficiently high to induce demethylation. This argument is supported by a previous report, showing that TET2 and TET3 both consistently co-occurred at active promoters (Deplus et al. 2013). In retrospect, an unintentional silencing of the shRNAs resulting from ZnF-3AC-mediated DNA methylation cannot be ruled out, which consequently would have abrogated the KD effect during the experiment.

To overcome the problem of inefficient KDs, the simultaneous inhibition of all TET enzymes was carried out using a chemical inhibitor. Again, no significant differences in DNA methylation stability were observed at multiple loci after continuous inhibitor treatment. The unvaried dynamics of DNA methylation might be caused by an insufficient inhibition of the TET enzymes. However, the most basic explanation for the lack of effects after TET-KD or TET inhibition would be that the observed demethylation actually occurred passively, at least at the investigated promoters. To ultimately clarify the role of TET enzymes in the removal of aberrant methylation, the employment of TET-knockout cell lines could be a more suitable approach. Moreover, the detection of 5hmC, 5fC, or 5caC resulting from TET-mediated oxidation of 5mC would be a strong evidence for active demethylation after DNA methylation editing (An et al. 2017).

Dynamic changes in gene expression and the epigenome network after genome-wide DNA methylation editing

The previously described observations about differential DNA methylation stability at thousands of CGI promoters indicated that the dynamics of this mark also had to be examined in view of the local chromatin context. Moreover, it was of great interest to investigate the pleiotropic effects arising from genome-wide DNA methylation editing. In general, the presence of 5mC at CpG-rich gene promoters is associated with the silencing of the underlying gene (Jones 2012). Thus, the goal was to determine whether the ZnF-3AC-mediated DNA methylation was already sufficient to induce major changes in gene expression. In a preprint from Ford et al. (2017), which however did not pass the peer review process and never was published in a journal, a low repressive capacity of DNA methylation was reported after globally introducing the mark at gene promoters (Ford et al. 2017). Despite using a similar epigenome editing system and experimental setup, our data essentially lead to a very different conclusion. With

more than 900 genes being downregulated at least 2-fold (Figure 36C), it became obvious that DNA methylation did actually have a strong and direct effect on gene expression. The extent as well as the dynamics of gene silencing were strongly connected to the 5mC levels present at the promoter CGIs. Hence, indirect effects such as a stress response of the cells or the dysregulation of chromatin-associated factors are very unlikely to explain the observed transcriptional changes. Moreover, no strong downregulation was observed for the genes associated to cluster A, which only gained very low levels of DNA methylation.

The observed gene silencing effects can be explained by various mechanisms. For instance, the binding of multiple transcription factors (e.g. CTCF, SP1, ETS) is impaired when a CpG within the recognition motif is methylated. As a consequence, TF-dependent recruitment of RNAPII to the methylated promoters is abolished (Hashimoto et al. 2017; Rozenberg et al. 2008). In addition, several very important pathways have been described, which are all based on the recruitment of MBD protein-containing repressive complexes to methylated CpG sites, followed by chromatin remodeling (Du et al. 2015). Some of the potentially involved mechanisms will be discussed later in this chapter. However, it has to be noted that a large fraction of genes did not show a considerable response to promoter methylation, as the RT-qPCR data of *VEGFA* (Figure 36A), and the RNA-seq data (Figure 36C) showed. A potential reason might be that additional epigenetic signals were required to induce the silencing of these genes. Moreover, it is possible that longer induction of ZnF-3AC might have led to a more robust establishment DNA methylation at the promoters, and thus to a more effective downregulation of the genes, as Bintu et al. (2016) previously showed (Bintu et al. 2016).

Our transcriptome data also unveiled that long-term silencing of genes by targeted DNA methylation editing still poses a tough challenge. Even some of the genes from the most stably methylated subcluster D1, which retained half of their DNA methylation at '11d off', partially returned to their initial expression levels. Importantly, the highly expressed genes (cluster A and B) were less susceptible for 5mC-dependent downregulation than lowly expressed genes (cluster C and D) (Figure 36D, Appendix 1 Figure 4C). As gene ontology analysis revealed, these genes were merely involved in processes that could be summarized as 'housekeeping function', (Appendix 1 Figure S12). One may hypothesize that these housekeeping genes possessed a better protection against aberrant DNA methylation, as their expression levels are usually fine-

tuned across different cell types in order to maintain basal cellular functions (She et al. 2009). In contrast, the more stably methylated promoters of the less expressed genes were connected to processes associated with differentiation and development, and these types of genes are typically highly dynamic in expression (Uyehara et al. 2017). Taken all together, gene expression and promoter methylation were strongly dependent of one another. The levels of gene expression to *de novo* DNA methylation displayed an inverse dynamic, and were strongly target gene-related.

DNA methylation does not function as an isolated epigenetic signal, but is also tightly interconnected with the highly complex epigenome network, including post-translational modifications of histone tails (Jeltsch et al. 2018). More than 500 different histone PTMs have been identified to date, which together form the histone code and control a large variety of chromatin-based processes (Huang et al. 2015; Strahl and Allis 2000). Thus, one of our main goals was to gain a better understanding of the dynamic interplay between DNA methylation and five selected histone PTMs.

Intriguingly, we were able to demonstrate that the deposition of DNA methylation caused a drastic depletion of the activating mark H3K4me3 at thousands of CGIs. The extent of this antagonistic effect was strongly correlated with the 5mC levels present at each time point of the experiment. On the other hand, the consistently high levels of H3K4me3 in cluster A and B were apparently not beneficial for the stable maintenance of DNA methylation. This is supported by the previous observation, that H3K4me3-marked CGIs were exclusively protected against off-target methylation caused by dCas9-3AC (Galonska et al. 2018). Moreover, the loss of DNA methylation at clusters A-C was accompanied by the recovery of H3K4 trimethylation, indicating that residual DNA methylation presumably prevented the complete return to initial H3K4me3 levels (subcluster D1-3) (Figure 37A). Mechanistically, the reduction of H3K4 trimethylation might be ascribed to the inability of the CXXC domains of MLL1/2 or CFP1 to bind to methylated DNA. As a consequence, the H3K4 trimethylation mark could not be efficiently re-established once being lost. *Vice versa*, by utilizing *Dnmt1*-KO cells, it was shown that the reduction of DNA methylation was accompanied by the establishment of H3K4me3, which well reflects the observations of our work (Hashimoto et al. 2010; Sharifi-Zarchi et al. 2017). The H3K4me3 mark is recognized by the reader protein TAF3, enabling formation of the transcription preinitiation complex. Accordingly, RNAPII colocalizes with trimethylated H3K4 at active promoters (Heintzman et al.

2007; Lauberth et al. 2013). Moreover, there is a positive crosstalk between H3K4me3 and histone acetylation, as trimethylation at H3K4 serves as an anchor for histone acetyltransferase (HAT) complexes (Zhang et al. 2015). The DNA methylation-dependent reduction of H3K4me3 may thus explain the observed gene repression, and their overall similar dynamics by inducing a downstream cascade of events.

In line with this, we found that H3K27ac was strongly reduced after DNA methylation editing (Figure 37B). This effect can also be explained by a mechanism directly linking DNA methylation to lysine deacetylation. MBD2 is a variable subunit of NuRD complexes, and can recruit HDACs to methylated CpG sites (Basta and Rauchman 2015). The combination of these both effects, MBD2-mediated recruitment of HDACs and the reduced association of HATs due to H3K4me3 depletion, may thus have contributed to the acetylation loss. Interestingly, the H3K27ac levels in all clusters and at each experimental time point well resembled the corresponding expression levels of the associated genes even more closely than H3K4me3 (Appendix 1 Figure 4C). One may speculate that after 5mC deposition, H3K27ac responded to the progressing DNA demethylation more rapidly, as acetylation is generally a highly dynamic modification occasionally changing within minutes (Weinert et al. 2018). Furthermore, H3K27 acetylation was shown to have a strong stimulatory effect on RNAPII activity, and to be tightly correlated with transcript levels (Stasevich et al. 2014). In summary, our work provides a comprehensive and time-resolved insight into the antagonistic effects between DNA methylation, and both activating marks H3K27ac and H3K4me3 at CGIs. By this, the immediate implications of DNA methylation on gene expression, and the predominantly transient silencing can be explained mechanistically.

A major focus of this work was set on the identification of features of the epigenome network that support the maintenance of *de novo* DNA methylation over long periods of time. The chromatin state segmentation analysis of cluster D clearly indicated that DNA methylation was more stable Polycomb-repressed chromatin (Appendix 1 Figure 3A), a chromatin state defined by the presence of H3K27me3, and known to regulate the expression of developmental genes (Filion et al. 2010). As mentioned before, cluster D could also be associated with the promoters of genes involved in development or differentiation, and therefore it was suspected to observe H3K27me3 at these sites. Strikingly, a marked enrichment of this modification could be identified exclusively in

cluster D. Within this cluster, a strong relationship between DNA methylation stability and H3K27me3 levels was determined (Figure 37A).

This finding was remarkable, as the co-occurrence of both modifications is highly debated. Various studies showed that H3K27me3 and 5mC act antagonistically, and are mutually exclusive at different genomic elements, such as CGIs or DNA methylation canyons (Brinkman et al. 2012; Zhang et al. 2020). This is in stark contrast to the results of our work, demonstrating that H3K27me3 and 5mC did even coexist in more than 2500 CGIs before DNA methylation editing (Figure 38A, cluster M1). Another group observed that a depletion of the 5mC mark was followed by a gain of H3K27me3, and *vice versa* (Hagarman et al. 2013). Again, our data do not confirm this trend, as the H3K27me3 levels remained unchanged in cluster D after the introduction of DNA methylation, even after later time points (Figure 37A). These diverging effects might be explained by the involvement of different variants of PRC2 complexes, some of them containing the subunits MTF2 and PHF1, which are unable to bind their target motif in the presence of CpG methylation. Other PRC2 variants however contain the AEBP2 subunit, which has a strong preference for methylated CpG sites, thus enabling the maintenance of the H3K27me3 mark after DNA methylation editing (Li et al. 2017; Wang et al. 2017).

There were also other findings supporting the simultaneous genome-wide enrichment of both H3K27me3 and DNA methylation, and particularly at CGIs. In various types of cancer, H3K27me3-marked promoters were susceptible for the acquisition of aberrant DNA methylation (Schlesinger et al. 2007; Takeshima et al. 2015; Widschwendter et al. 2007). Moreover, *Dnmt3b* KO-experiments in mice showed that the catalytic activity of DNMT3B was associated with the enrichment of H3K27me3 at a subset of DMRs, while the enzyme only functioned as an accessory factor at DMRs marked by activating histone PTMs (Nowialis et al. 2019). Further studies revealed that monovalent and bivalent promoters containing H3K27me3 are permissive for the accumulation of DNA methylation by a DNMT3A mutant associated with microcephalic dwarfism (Heyn et al. 2019; Sendžikaitė et al. 2019). Our data illustrate that the H3K4me3/H3K27me3 bivalency was not the determining factor allowing the stable propagation of the 5mC marks over long periods of time, but instead, the presence of H3K27me3 was already sufficient at these CGIs (Figure 38A, clusters U1 and U2). A very recent study from Weinberg et al. (2021) revealed the mechanism for the recruitment of DNMT3A to

H3K27me3-marked CGIs. These authors showed that DNMT3A1 binds to H2AK119ub via a newly discovered ubiquitin-dependent recruitment region (UDR) domain (Weinberg et al. 2021). H2AK119ub is deposited by PRC1 complexes, recruits and stimulates PRC2 by JARID2/AEBP2 binding, and consequently co-localizes with H3K27me3 (Zhang et al. 2015).

Finally, the effects on gene expression were the strongest in subcluster D1 (Figure 36C), where strong DNA methylation, and the highest H3K27me3 levels co-occurred. These regions also contained H3K4me3, and thus represented a bivalent state. Hence, it is tempting to speculate that the deposition of 5mC induced the resolution of the bivalent state in parts, and consequently reinforced the repressive capacity of H3K27me3. This may also explain why the H3K4me3 did not fully return to its initial levels (Figure 37A). Furthermore, the efficient long-term repression of genes by combinatorial DNA methylation and H3K27me3 enrichment at the corresponding promoters was also observed in the study by O'Geen et al. (2019). Hence, the co-delivery of 5mC and H3K27me3 is a promising strategy for the induction of a stably repressed chromatin state, and further studies will be required to validate its universality at different genomic targets.

A rather surprising observation of this work were the low levels of H3K9me3 at the already pre-methylated CGIs (Appendix 1, Figure S4A). This observation was in line with a lack of H3K9me3 after DNA methylation editing, even at the most stably methylated regions (Figure 37C). In principle, diverse pathways could have led to the enrichment of H3K9me3 at methylated CGIs. For instance, MBD1 was shown to recruit either SETDB1 or the SUV39H1-HP1 complex to regions with CpG methylation, followed by the trimethylation of H3K9, and the induction of heterochromatin formation (Al-Sady et al. 2013; Ichimura et al. 2005). Similarly, MeCP2 was reported to direct H3K9 HMTases to regions containing DNA methylation (Fuks et al. 2003a). An inverse crosstalk between H3K9me3 and DNA methylation has also been described, as HP1 was shown to direct DNMT3B to major satellite repeats (Lehnertz et al. 2003). Potentially, the low H3K9me3 levels at methylated CGIs have been a cell line-specific feature, which is supported by the finding that both modifications were mutually exclusive in the HCC1954 and IMR90 cancer cell lines (Hon et al. 2012). By this, one may conclude that 5mC-mediated deposition of H3K9me3 does not occur efficiently at CGIs in

various cell types. Nevertheless, Nuñez et al. (2021) recently published their CRIS-PRoff system, by which they could show that the co-delivery of DNA methylation and H3K9me3 enabled the stable silencing of a reporter gene in HEK293T cells. In contrast, the deposition of only one modification only lead to a transient and less effective gene repression (Nuñez et al. 2021). This finding indicates a synergistic effect of DNA methylation and H3K9 trimethylation in maintaining a silent chromatin state, even in a cell line closely related to ours.

Interestingly, no global effects of DNA methylation on the levels of H3K36me3 at CGIs or gene bodies were observed (Figure 37D). In the case of CGIs, it was rather unsurprising to obtain no H3K36me3 increase, as this modification is inherently depleted at transcriptional start sites, independent of the expression level of the underlying gene. However, a strong correlation between gene expression levels, and the enrichment of H3K36me3 within the respective gene bodies was reported previously (Barski et al. 2007). Thus, it is difficult to elucidate the unaltered H3K36me3 levels, especially in cluster D, where strong transcriptional repression was observed. Despite of the similar mechanism between yeast and humans regarding the deposition of H3K36me3 at gene bodies, there might still be a different dynamic in setting or removing the modification after silencing. At least, no direct crosstalk between the deposited 5mC mark and H3K36me3 was anticipated, as DNA methylation acts downstream of H3K36me3 at gene bodies by the recruitment of DNMT3A/B via their PWWP domain (Li et al. 2019).

Altogether, our work provided valuable insights into the complex interplay of DNA methylation with various histone PTMs. We were able to directly connect the dynamics of the activating marks H3K4me3 and H3K27ac to the changes in gene expression after DNA methylation editing. Moreover, we identified PcG chromatin to be a suitable target for the stable integration of the 5mC mark for longer periods of time, and could reveal a marked antagonistic response against aberrant DNA methylation at the majority of CGIs. Lastly, we were able to demonstrate the strong silencing capacity of promoter methylation at almost one thousand genes, and thus refined the understanding about the regulation of gene expression by the epigenome network.

5.2 Engineering of an EpiEditor with improved specificity for targeted DNA methylation editing

Unlike the main project of this PhD thesis, the goal of the second project was to deliver DNA methylation with the highest possible specificity, i.e. as little methylation as possible at off-target sites, but without compromising the on-target methylation. This feature of epigenome editing systems is of central importance for potential therapeutic applications, in which the efficient reprogramming of chromatin states is intended, while undesired side-effects originating from off-target activity have to be avoided. For this reason, we sought to compare and engineer different EpiEditors based on CRISPR/dCas9 targeting, as this system displays an improved specificity and flexibility compared to TALE- and ZnF proteins (Kungulovski and Jeltsch 2016).

In this process, we were able to reproduce the observations from Stepper et al. (2017), showing that the direct fusion of the chimeric 3AC-3L effector domain to dCas9 introduced DNA methylation with markedly higher efficiency at a target promoter than dCas9-3AC lacking 3L fusion (Figure 39C). Different from their results, we observed a very high off-target methylation at the untargeted *VEGFA* promoter. This off-target activity of both constructs might be attributed to the unintended binding of dCas9 to other genomic loci. Using high-speed atomic force microscopy, Shibata et al. (2017) showed that the Cas9 apoenzyme randomly bound and subsequently slid along the DNA (Shibata et al. 2017). Consequently, a possible explanation might be the binding of dCas9-3AC(-3L) to DNA in the absence of sgRNA. This way, one may speculate that dCas9 brought its directly fused effector domains into close contact to the DNA, and thereby facilitated the delivery of untargeted DNA methylation. However, the higher specificity of dCas9-3AC compared to dCas9-3AC-3L (Figure 39D) indicated that the off-target activity presumably was caused by the employed effector domains. This assumption is supported by a previous study showing that dCas9 or the utilized sgRNA did not contribute to the off-target methylation of dCas9-3AC (Galonska et al. 2018). A lower specificity of the dCas9-3AC system has also been previously confirmed by other studies, and thus contradicts the results from Stepper et al. (2017) (Galonska et al. 2018; Pflueger et al. 2018; Stepper et al. 2017). Notably, in order to be catalytically active, dCas9-3AC has to form an active 3AC homotetramer, even at the off-target sites (Gowher and Jeltsch 2018). In contrast, dCas9-3AC-3L only needs to dimerize.

Based on our data obtained with the dCas9 direct fusion systems, we considered necessary to employ the novel dCas9-SunTag system in combination with either Ab-3AC or Ab-3AC-3L. Using dCas9-SunTag/Ab-3AC, we observed a medium gain in DNA methylation at the *ISG15* promoter (40 %), and almost no increase at the untargeted locus (5 %, Bis-Seq data from Laura Laistner's Bachelor's thesis). In comparison, Huang et al. (2017) obtained a higher gain in methylation at the *HOXA5* promoter (80 %), and Pflueger et al. (2018) a lower gain at the *UNC5C* promoter (up to 19 %) (Huang et al. 2017; Pflueger et al. 2018). However, the data from the previous ZnF-3AC project showed, that the efficiency of *de novo* DNA methylation strongly varies among different target regions. Furthermore, the gain in methylation is also depended on the concentration of the effector domain within the cell (Pflueger et al. 2018). This also explains why in the main project of this thesis, we observed high levels of genome-wide *de novo* methylation after the overexpression of 3AC, as the construct was under the control of the strong TRE3G promoter (Qin et al. 2010). Overall, our data indicated that the 3AC homotetramer formation was not highly efficient in our experimental setting, leading to unsatisfying low *de novo* DNA methylation levels at the target locus.

By utilizing the dCas9-SunTag/3AC-3L system, we aimed to overcome this problem, as only two 3AC-3L dimers had to interact at the RD interface in order to form an active heterotetramer. We also wanted to make use of the stimulatory effect of 3L on 3AC activity to boost the efficiency of the effector domain (Gowher et al. 2005). Indeed, in this setting, we obtained the highest on-target methylation levels of all tested systems (Figure 39C). Though, similar to the direct fusion of 3AC-3L to dCas9, the off-target activity at the *VEGFA* promoter was very high. But even more staggering was the enormous global off-target activity at thousands of CGIs (Figure 41B), illustrating the challenges which epigenome editing has to overcome before being applied in therapeutics.

Based on these observations, our goal was to reduce the affinity of the effector domain to DNA, and consequently reduce its global off-target binding. Our design in protein engineering aimed to reverse the positive charge of individual lysine or arginine residues in the 3AC/DNA interface by mutating them to glutamic acid. Previously, a similar effort was also successfully applied for the bacterial MTase M.SssI, which displayed a higher specificity after mutagenizing the DNA binding domain (Xiong et al. 2018). Strikingly, all of our tested DNA interface mutants of 3AC-3L displayed drastically reduced

off-target activity at the *VEGFA* promoter (Figure 40C), which also supports the previous assumption that 3AC binding to DNA was predominantly responsible for the untargeted methylation. These mutants also exhibited a medium to high residual activity at the target locus, indicating that dCas9-SunTag recruitment was efficient and rescued the reduced DNA affinity of the effector domains. Both, the Ab-3AC(K844E)-3L and Ab-3AC(R887E)-3L constructs showed an improved relative specificity compared to the already existing Ab-3AC construct at genome-wide scale, (Figure 41D), and an equal or higher on-target activity (Figure 40C and Figure 41C). An interesting side note is the observation that the R887E mutant exhibited a higher activity and slightly stronger off-target binding than the R831E mutant. Previously, *in vitro* studies using arginine-to-alanine mutations actually showed an inverse trend for both variants. Another finding of our experiments was that the RD interface mutation (R885A) could not be rescued by the simultaneous association of multiple effector domains at the SunTag (Figure 40C). This mutation was shown to cause reduced AdoMet binding, and completely abolish the catalytic activity of the enzyme (Gowher et al. 2006). Hence, the R885A mutation has turned out to be too drastic for our DNA methylation editing attempts.

In summary, the dCas9-SunTag-based system in combination with either 3AC or 3AC-3L displayed are higher specificity, when compared to the direct fusion systems. With our work, we contributed to the expansion of the toolbox of available EpiEditors, that can be employed for targeted DNA methylation editing approaches in a more advanced manner.

5.3 Reprogramming of the *H19/IGF2* imprinting control region by targeted demethylation

While many studies have focused in the targeted deposition of DNA methylation at gene promoters using DNMT3 enzymes, some progress has also been previously made in the enzyme-driven demethylation of specific genomic regions. Our goal was to efficiently reprogram the *H19/IGF2* imprinting control region (ICR), which should be achieved by the demethylation of multiple CpG-rich motifs at this locus. In healthy individuals, the ICR is methylated on one of the two alleles, and hypermethylation of this

locus is associated with various disorders (e.g. Beckwith-Wiedemann Syndrome or Wilms' tumor). Since our HEK293 cell model was methylated on half of the alleles (preliminary data from Dr. Pavel Bashtrykov), we aimed to remove the present DNA methylation as a proof-of-principle, and consequently induce the methylation state of the Silver-Russell Syndrome (Cui et al. 2001; Gicquel et al. 2005; Nativio et al. 2011). We were also interested in the stability of the reprogrammed state of the ICR. Finally, we sought to investigate whether the recruitment of the methylation-sensitive insulator protein CTCF was facilitated after demethylation.

In order to succeed, the choice of a suitable EpiEditing system was of great importance. The first studies aiming to specifically demethylate selected promoters via the direct fusion of TET enzymes to TALE targeting modules only obtained weak reductions in 5mC levels (Chen et al. 2014; Maeder et al. 2013). Similarly, the direct fusion of the TET1 catalytic domain to dCas9 led to only a rather mild reduction of 5mC levels at several gene promoters, but also only very weak off-target effects (Choudhury et al. 2016; Liu et al. 2016). In a very recent publication from Nuñez et al. (2021), a strongly improved efficiency of the direct fusion system by applying an elongated 80 aa linker between TET1 and dCas9 was shown (Nuñez et al. 2021). Since this paper was released after finishing our experiments, the tool was not included in our work. However, already in 2016, a major step was made in the development of an efficient targeted DNA demethylation tool by Morita et al. (2016). They combined the dCas9-SunTag with Ab-TET1(CD) and implemented some important modifications in the design of the construct. For instance, they cloned all components of the system into the same vector, applied longer linkers between the GCN4 peptides allowing a higher mobility of the SunTag, and reduced the overall number of GCN4 repeats to five. In the end, they demonstrated a strong demethylation activity of up to 90 % reduction in methylation at several loci (Morita et al. 2016).

By employing their novel EpiEditing system, and introducing some minor adjustments into the construct (e.g. co-expression of GFP via an IRES), we succeeded in efficiently reducing DNA methylation at the *H19/IGF2* ICR to only one third of the initial levels on average (Figure 43B). The mostly unaltered 5mC levels obtained with the empty vector control, containing all components besides the TET1 effector domain, also proved that the demethylation was executed by the catalytic activity of the enzyme. The strong 5mC loss once more confirmed the high proficiency of the utilized dCas9-SunTag/Ab-

TET1 system. Strikingly, we were able to show that the reduced DNA methylation levels were stable for almost one month, which is of great importance for potential clinical approaches (Kungulovski and Jeltsch 2016). Although the multiplexed sgRNAs only targeted four of the CTCF target sites (CTS) of the ICR directly, a marked reduction of DNA methylation was also observed at the additional CTSs as well, and even at the *H19* CGI, which was located 2 kb downstream of the ICR (Figure 43A+B). This might be explained by the presumably high flexibility of the SunTag with its long linkers, enabling the effector domains to reach genomic regions that were even several kilobases away from the dCas9 binding site, but within one local 3D region. These long-range effects of the dCas9-SunTag system were also previously reported by Huang et al. (2017), highlighting the strong potential of this EpiEditing platform to even modify larger stretches of DNA (Huang et al. 2017). In the course of our lab work, Horii et al. (2020) published a study trying to accomplish the same goals as our work. At six of their targeted CTSs, they obtained an approximately 2-fold reduction in DNA methylation five days after transfection in HEK293 cells. In comparison, at the same sites, we even achieved a 3.6-fold reduction in MBD2-qPCR signal, demonstrating that our slightly modified construct was even more efficient (Horii et al. 2020). This effect might be related to differences in the expression levels of the EpiEditor components. Another explanation might also be the direct fusion of GFP in their construct, which in parts may have sterically hindered the effector in the demethylation process.

As a side note, the control region 2 kb downstream was the only site which did not preserve its hypomethylated state (Figure 43A+B). As this site was located within a CGI, one may speculate that the local excess of endogenous DNMTs protected the CGI against erroneous demethylation (Jeltsch and Jurkowska 2014). This assumption is supported by the observation, that dCas9-SunTag/Ab-TET1-mediated demethylation of the murine *Fgf21* promoter was only transient and completely lost after six weeks (Hanzawa et al. 2020).

A further major discovery of our work was the recruitment of CTCF to the CTSs after targeted demethylation. In principle, it was expected to obtain a 2-fold gain at the maximum, as half of the alleles were assumed to be already bound by CTCF. Our average 2.5-fold increase at '6 days' (Figure 43C) indicated a very efficient recruitment of CTCF, and the excess gain might simply be attributed to experimental fluctuations. A slightly

weaker gain in CTCF signal was also reported by Horii et al. (2020) in the same experimental setting, which once more confirms the methylation-dependent binding of the insulator (Horii et al. 2020). Strikingly, the increased occupancy of CTCF was even maintained after 27 days, and was concomitant with the low DNA methylation levels at these sites (Figure 43C). Hence, our study illustrated that the stable reprogramming of the *H19/IGF2* ICR was technically possible. It also shows that the loss of imprinting can be artificially induced, and thereby probably leads to the reorganization of the chromatin topology by CTCF. Based on these results, further work is ongoing to further characterize the effects of the reprogramming of the ICR on the expression of *H19* and *IGF2*, the off-target activity of dCas9-SunTag/Ab-TET1, and the changes in the 3D chromatin structure.

6. Outlook

The results of this PhD thesis provide a detailed insight into the challenges and opportunities of targeted DNA methylation editing. We were able to show that the establishment of promoter methylation in many cases had direct implications on gene expression, but also that only a small fraction of gene promoters actually became stably methylated. The propagation of the 5mC mark over long time periods was shown to be dependent on the local chromatin context. It would be an interesting goal to analyze the DNA methylation stability in different cell lines or even in primary cells, in order to determine the universal validity of our observations. On the other hand, by using engineered EpiEditing tools, we were also able to demonstrate how DNA methylation can be delivered or removed with high efficiency at different genomic loci. At the same time, the off-target effects caused by unspecific 3AC binding were shown to be reduced to a minimum. In the near future, our currently ongoing targeted demethylation project will also unveil the specificity of the dCas9-SunTag/Ab-TET1 system on a genome-wide scale.

With the improved knowledge about the dynamics of the epigenome network, and the availability of enhanced EpiEditors, it will become more feasible in basic research or therapeutic approaches to efficiently modify the DNA methylation state at different genomic loci. However, our data also indicate that it might be necessary to determine the local chromatin state of a given genomic locus prior to DNA methylation editing, in order to predict its predisposition for the stable propagation of this mark. The overall strong protection of CGI promoters against *de novo* DNA methylation shows that the deposition of additional chromatin modifications may be unavoidable. In line with this, several studies previously revealed a more robust silencing of various example genes by the co-delivery of H3K9me3 or H3K27me3 (Nuñez et al. 2021; O'Geen et al. 2019). With our improved dCas9-SunTag/Ab-3AC(mt)3L system, a higher specificity in epigenome editing approaches is made possible, and similar studies will also be necessary for histone-modifying enzymes in order to eliminate unintended off-target effects.

For therapeutic purposes, the delivery of the EpiEditors to the unhealthy cells or tissues still remains a difficult task. For instance, the nucleic acids coding for the EpiEditors can be introduced into the patient cells by using viral vectors, nanoparticles, liposomes

or microinjections (Liu et al. 2020). One concept for the treatment of epigenetic diseases in blood cells might be the withdrawal of bone marrow or mobilized peripheral blood, followed by the isolation and cultivation of hematopoietic stem cells (HSCs). By transfecting the HSCs *ex vivo* with the respective EpiEditing tools, the dysregulated genes could be efficiently silenced or reactivated. Afterwards, the modified cells would be reinfused into the patient, and could differentiate into various cell types, showing a healthy phenotype (Morgan et al. 2017). However, still a lot of progress has to be made in basic research, in order to make epigenome editing applicable and safe for the daily therapeutic use.

References

- Aberg, Karolina A.; Chan, Robin F.; van den Oord, Edwin J. C. G. (2020): MBD-seq - realities of a misunderstood method for high-quality methylome-wide association studies. In *Epigenetics* 15 (4), pp. 431–438.
- Afgan, Enis; Baker, Dannon; Batut, Bérénice; van den Beek, Marius; Bouvier, Dave; Cech, Martin et al. (2018): The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. In *Nucleic acids research* 46 (W1), W537-W544.
- Agarwal, Noopur; Hardt, Tanja; Brero, Alessandro; Nowak, Danny; Rothbauer, Ulrich; Becker, Annette et al. (2007): MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. In *Nucleic acids research* 35 (16), pp. 5402–5408.
- Alajem, Adi; Roth, Hava; Ratgauzer, Sofia; Bavli, Danny; Motzik, Alex; Lahav, Shlomtzion et al. (2021): DNA methylation patterns expose variations in enhancer-chromatin modifications during embryonic stem cell differentiation. In *PLoS genetics* 17 (4), e1009498.
- Albert, Pál; Varga, Bence; Zsibrita, Nikolett; Kiss, Antal (2018): Circularly permuted variants of two CG-specific prokaryotic DNA methyltransferases. In *PloS one* 13 (5), e0197232.
- Al-Sady, Bassem; Madhani, Hiten D.; Narlikar, Geeta J. (2013): Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. In *Molecular cell* 51 (1), pp. 80–91.
- Amabile, Angelo; Migliara, Alessandro; Capasso, Paola; Biffi, Mauro; Cittaro, Davide; Naldini, Luigi; Lombardo, Angelo (2016): Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. In *Cell* 167 (1), 219-232.e14.
- An, Jungeun; Rao, Anjana; Ko, Myunggon (2017): TET family dioxygenases and DNA demethylation in stem cells and cancers. In *Experimental & molecular medicine* 49 (4), e323.
- Anders, Carolin; Niewoehner, Ole; Duerst, Alessia; Jinek, Martin (2014): Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. In *Nature* 513 (7519), pp. 569–573.
- Aran, Dvir; Toperoff, Gidon; Rosenberg, Michael; Hellman, Asaf (2011): Replication timing-related and gene body-specific methylation of active human genes. In *Human molecular genetics* 20 (4), pp. 670–680.
- Atlasi, Yaser; Stunnenberg, Hendrik G. (2017): The interplay of epigenetic marks during stem cell differentiation and development. In *Nature reviews. Genetics* 18 (11), pp. 643–658.
- Azuara, Véronique; Perry, Pascale; Sauer, Stephan; Spivakov, Mikhail; Jørgensen, Helle F.; John, Rosalind M. et al. (2006): Chromatin signatures of pluripotent cell lines. In *Nature cell biology* 8 (5), pp. 532–538.
- Bachman, K. E.; Rountree, M. R.; Baylin, S. B. (2001): Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. In *The Journal of biological chemistry* 276 (34), pp. 32282–32287.

- Baedke, Jan (2013): The epigenetic landscape in the course of time: Conrad Hal Waddington's methodological impact on the life sciences. In *Studies in history and philosophy of biological and biomedical sciences* 44 (4 Pt B), pp. 756–773.
- Bailey, Timothy L.; Boden, Mikael; Buske, Fabian A.; Frith, Martin; Grant, Charles E.; Clementi, Luca et al. (2009): MEME SUITE: tools for motif discovery and searching. In *Nucleic acids research* 37 (Web Server issue), W202-8.
- Bannister, Andrew J.; Kouzarides, Tony (2011): Regulation of chromatin by histone modifications. In *Cell research* 21 (3), pp. 381–395.
- Barau, Joan; Teissandier, Aurélie; Zamudio, Natasha; Roy, Stéphanie; Nalesso, Valérie; Héroult, Yann et al. (2016): The DNA methyltransferase DNMT3C protects male germ cells from trans-ponson activity. In *Science (New York, N.Y.)* 354 (6314), pp. 909–912.
- Barlow, Denise P.; Bartolomei, Marisa S. (2014): Genomic imprinting in mammals. In *Cold Spring Harbor perspectives in biology* 6 (2).
- Barneda-Zahonero, Bruna; Parra, Maribel (2012): Histone deacetylases and cancer. In *Molecular oncology* 6 (6), pp. 579–589.
- Barski, Artem; Cuddapah, Suresh; Cui, Kairong; Roh, Tae-Young; Schones, Dustin E.; Wang, Zhibin et al. (2007): High-resolution profiling of histone methylations in the human genome. In *Cell* 129 (4), pp. 823–837.
- Bashtrykov, Pavel; Jankevicius, Gytis; Smarandache, Anita; Jurkowska, Renata Z.; Ragozin, Sergey; Jeltsch, Albert (2012): Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain. In *Chemistry & biology* 19 (5), pp. 572–578.
- Basta, Jeannine; Rauchman, Michael (2015): The nucleosome remodeling and deacetylase complex in development and disease. In *Translational research : the journal of laboratory and clinical medicine* 165 (1), pp. 36–47.
- Baubec, Tuncay; Colombo, Daniele F.; Wirbelauer, Christiane; Schmidt, Juliane; Burger, Lukas; Krebs, Arnaud R. et al. (2015): Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. In *Nature* 520 (7546), pp. 243–247.
- Baulina, Natalia; Kiselev, Ivan; Favorova, Olga (2021): Imprinted Genes and Multiple Sclerosis: What Do We Know? In *International journal of molecular sciences* 22 (3).
- Beaulaurier, John; Schadt, Eric E.; Fang, Gang (2019): Deciphering bacterial epigenomes using modern sequencing technologies. In *Nature reviews. Genetics* 20 (3), pp. 157–172.
- Bell, A. C.; Felsenfeld, G. (2000): Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. In *Nature* 405 (6785), pp. 482–485.
- Berger, Shelley L.; Kouzarides, Tony; Shiekhatar, Ramin; Shilatifard, Ali (2009): An operational definition of epigenetics. In *Genes & development* 23 (7), pp. 781–783.
- Bergman, Daniel; Halje, Matilda; Nordin, Matilda; Engström, Wilhelm (2013): Insulin-like growth factor 2 in development and disease: a mini-review. In *Gerontology* 59 (3), pp. 240–249.

Bernstein, Bradley E.; Mikkelsen, Tarjei S.; Xie, Xiaohui; Kamal, Michael; Huebert, Dana J.; Cuff, James et al. (2006): A bivalent chromatin structure marks key developmental genes in embryonic stem cells. In *Cell* 125 (2), pp. 315–326.

Bhattacharya, Saikat; Levy, Michaela J.; Zhang, Ning; Li, Hua; Florens, Laurence; Washburn, Michael P.; Workman, Jerry L. (2021): The methyltransferase SETD2 couples transcription and splicing by engaging mRNA processing factors through its SHI domain. In *Nature communications* 12 (1), p. 1443.

Bintu, Lacramioara; Yong, John; Antebi, Yaron E.; McCue, Kayla; Kazuki, Yasuhiro; Uno, Narumi et al. (2016): Dynamics of epigenetic regulation at the single-cell level. In *Science (New York, N.Y.)* 351 (6274), pp. 720–724.

Bird, Adrian (2002): DNA methylation patterns and epigenetic memory. In *Genes & development* 16 (1), pp. 6–21.

Blackledge, Neil P.; Klose, Robert (2011): CpG island chromatin: a platform for gene regulation. In *Epigenetics* 6 (2), pp. 147–152.

Blake, Georgina Et; Watson, Erica D. (2016): Unravelling the complex mechanisms of transgenerational epigenetic inheritance. In *Current opinion in chemical biology* 33, pp. 101–107.

Blomen, V. A.; Boonstra, J. (2011): Stable transmission of reversible modifications: maintenance of epigenetic information through the cell cycle. In *Cellular and molecular life sciences : CMLS* 68 (1), pp. 27–44.

Blow, Matthew J.; Clark, Tyson A.; Daum, Chris G.; Deutschbauer, Adam M.; Fomenkov, Alexey; Fries, Roxanne et al. (2016): The Epigenomic Landscape of Prokaryotes. In *PLoS genetics* 12 (2), e1005854.

Bogdanove, Adam J.; Voytas, Daniel F. (2011): TAL effectors: customizable proteins for DNA targeting. In *Science (New York, N.Y.)* 333 (6051), pp. 1843–1846.

Bourc'his, D.; Xu, G. L.; Lin, C. S.; Bollman, B.; Bestor, T. H. (2001): Dnmt3L and the establishment of maternal genomic imprints. In *Science (New York, N.Y.)* 294 (5551), pp. 2536–2539.

Bourc'his, Déborah; Bestor, Timothy H. (2004): Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. In *Nature* 431 (7004), pp. 96–99.

Brenner, Carmen; Deplus, Rachel; Didelot, Céline; Lorient, Axelle; Viré, Emmanuelle; Smet, Charles de et al. (2005): Myc represses transcription through recruitment of DNA methyltransferase corepressor. In *The EMBO journal* 24 (2), pp. 336–346.

Brinkman, Arie B.; Gu, Hongcang; Bartels, Stefanie J. J.; Zhang, Yingying; Matarese, Filomena; Simmer, Femke et al. (2012): Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. In *Genome research* 22 (6), pp. 1128–1138.

Broche, Julian; Kungulovski, Goran; Bashtrykov, Pavel; Rathert, Philipp; Jeltsch, Albert (2021): Genome-wide investigation of the dynamic changes of epigenome modifications after global DNA methylation editing. In *Nucleic acids research* 49 (1), pp. 158–176.

Brockdorff, Neil (2017): Polycomb complexes in X chromosome inactivation. In *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 372 (1733).

Brown, K. W.; Villar, A. J.; Bickmore, W.; Clayton-Smith, J.; Catchpoole, D.; Maher, E. R.; Reik, W. (1996): Imprinting mutation in the Beckwith-Wiedemann syndrome leads to biallelic IGF2 expression through an H19-independent pathway. In *Human molecular genetics* 5 (12), pp. 2027–2032.

Browne, M. J.; Turnbull, J. F.; McKay, E. L.; Adams, R. L.; Burdon, R. H. (1977): The sequence specificity of a mammalian DNA methylase. In *Nucleic acids research* 4 (4), pp. 1039–1045.

Bulyk, M. L.; Huang, X.; Choo, Y.; Church, G. M. (2001): Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. In *Proceedings of the National Academy of Sciences of the United States of America* 98 (13), pp. 7158–7163.

Bustin, S. A. (2000): Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. In *Journal of molecular endocrinology* 25 (2), pp. 169–193.

Cano-Rodriguez, David; Gjaltema, Rutger A. F.; Jilderda, Laura J.; Jellema, Pytrick; Dokter-Fokkens, Jelleke; Ruiters, Marcel H. J.; Rots, Marianne G. (2016): Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. In *Nature communications* 7, p. 12284.

Cano-Rodriguez, David; Rots, Marianne G. (2016): Epigenetic Editing: On the Verge of Reprogramming Gene Expression at Will. In *Current genetic medicine reports* 4 (4), pp. 170–179.

Carey, Michael F.; Peterson, Craig L.; Smale, Stephen T. (2009): Chromatin immunoprecipitation (ChIP). In *Cold Spring Harbor protocols* 2009 (9), [pdb.prot5279](https://doi.org/10.1101/cuprot5279).

Carnell, Ammie N.; Goodman, Jay I. (2003): The long (LINEs) and the short (SINEs) of it: altered methylation as a precursor to toxicity. In *Toxicological sciences : an official journal of the Society of Toxicology* 75 (2), pp. 229–235.

Cervera, Laura; Gutiérrez, Sonia; Gòdia, Francesc; Segura, María M. (2011): Optimization of HEK 293 cell growth by addition of non-animal derived components using design of experiments. In *BMC proceedings* 5 Suppl 8, P126.

Chantalat, Sophie; Depaux, Arnaud; Héry, Patrick; Barral, Sophie; Thuret, Jean-Yves; Dimitrov, Stefan; Gérard, Matthieu (2011): Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. In *Genome research* 21 (9), pp. 1426–1437.

Chen, Hui; Kazemier, Hinke G.; Groote, Marloes L. de; Ruiters, Marcel H. J.; Xu, Guo-Liang; Rots, Marianne G. (2014): Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter. In *Nucleic acids research* 42 (3), pp. 1563–1574.

Chen, R. Z.; Pettersson, U.; Beard, C.; Jackson-Grusby, L.; Jaenisch, R. (1998): DNA hypomethylation leads to elevated mutation rates. In *Nature* 395 (6697), pp. 89–93.

Cheng, Albert W.; Jillette, Nathaniel; Lee, Phoebe; Plaskon, Dylan; Fujiwara, Yasuhiro; Wang, Wenbo et al. (2016): Casilio: a versatile CRISPR-Cas9-Pumilio hybrid for gene regulation and genomic labeling. In *Cell research* 26 (2), pp. 254–257.

Cheng, X.; Roberts, R. J. (2001): AdoMet-dependent methylation, DNA methyltransferases and base flipping. In *Nucleic acids research* 29 (18), pp. 3784–3795.

Chou, Chun-Chi; Wei, Shu-Yi; Lou, Yuan-Chao; Chen, Chinpan (2017): In-depth study of DNA binding of Cys2His2 finger domains in testis zinc-finger protein. In *PloS one* 12 (4), e0175051.

Choudhury, Samrat Roy; Cui, Yi; Lubecka, Katarzyna; Stefanska, Barbara; Irudayaraj, Joseph (2016): CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. In *Oncotarget* 7 (29), pp. 46545–46556.

Christensen, David G.; Xie, Xueshu; Basisty, Nathan; Byrnes, James; McSweeney, Sean; Schilling, Birgit; Wolfe, Alan J. (2019): Post-translational Protein Acetylation: An Elegant Mechanism for Bacteria to Dynamically Regulate Metabolic Functions. In *Frontiers in microbiology* 10, p. 1604.

Collins, Bridget E.; Greer, Celeste B.; Coleman, Benjamin C.; Sweatt, J. David (2019): Histone H3 lysine K4 methylation and its role in learning and memory. In *Epigenetics & chromatin* 12 (1), p. 7.

Cui, H.; Niemitz, E. L.; Ravenel, J. D.; Onyango, P.; Brandenburg, S. A.; Lobanenko, V. V.; Feinberg, A. P. (2001): Loss of imprinting of insulin-like growth factor-II in Wilms' tumor commonly involves altered methylation but not mutations of CTCF or its binding site. In *Cancer research* 61 (13), pp. 4947–4950.

Cui, Xiao-Long; Nie, Ji; Ku, Jeremy; Dougherty, Urszula; West-Szymanski, Diana C.; Collin, Francois et al. (2020): A human tissue map of 5-hydroxymethylcytosines exhibits tissue specificity through gene and enhancer modulation. In *Nature communications* 11 (1), p. 6161.

Dante, R.; Dante-Paire, J.; Rigal, D.; Roizès, G. (1992): Methylation patterns of long interspersed repeated DNA and alphoid repetitive DNA from human cell lines and tumors. In *Anti-cancer research* 12 (2), pp. 559–563.

Das, Atze T.; Tenenbaum, Liliane; Berkhout, Ben (2016): Tet-On Systems For Doxycycline-inducible Gene Expression. In *Current gene therapy* 16 (3), pp. 156–167.

Deplus, Rachel; Delatte, Benjamin; Schwinn, Marie K.; Defrance, Matthieu; Méndez, Jacqui; Murphy, Nancy et al. (2013): TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. In *The EMBO journal* 32 (5), pp. 645–655.

Dhayalan, Arunkumar; Rajavelu, Arumugam; Rathert, Philipp; Tamas, Raluca; Jurkowska, Renata Z.; Ragozin, Sergey; Jeltsch, Albert (2010): The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. In *The Journal of biological chemistry* 285 (34), pp. 26114–26120.

Dragan, A. I.; Pavlovic, R.; McGivney, J. B.; Casas-Finet, J. R.; Bishop, E. S.; Strouse, R. J. et al. (2012): SYBR Green I: fluorescence properties and interaction with DNA. In *Journal of fluorescence* 22 (4), pp. 1189–1199.

Du, Qian; Luu, Phuc-Loi; Stirzaker, Clare; Clark, Susan J. (2015): Methyl-CpG-binding domain proteins: readers of the epigenome. In *Epigenomics* 7 (6), pp. 1051–1073.

Dukat, Michael; Holzer, Katharina; Choudalakis, Michel; Emperle, Max; Lungu, Cristiana; Bashtrykov, Pavel; Jeltsch, Albert (2019): H3K36me2/3 Binding and DNA Binding of the DNA Methyltransferase DNMT3A PWWP Domain Both Contribute to its Chromatin Interaction. In *Journal of molecular biology* 431 (24), pp. 5063–5074.

Egger, Gerda; Liang, Gangning; Aparicio, Ana; Jones, Peter A. (2004): Epigenetics in human disease and prospects for epigenetic therapy. In *Nature* 429 (6990), pp. 457–463.

Emperle, Max; Adam, Sabrina; Kunert, Stefan; Dukatz, Michael; Baude, Annika; Plass, Christoph et al. (2019): Mutations of R882 change flanking sequence preferences of the DNA methyl-transferase DNMT3A and cellular methylation patterns. In *Nucleic acids research* 47 (21), pp. 11355–11367.

Ernst, Jason; Kheradpour, Pouya; Mikkelsen, Tarjei S.; Shores, Noam; Ward, Lucas D.; Epstein, Charles B. et al. (2011): Mapping and analysis of chromatin state dynamics in nine human cell types. In *Nature* 473 (7345), pp. 43–49.

Fatemi, Mehrnaz; Hermann, Andrea; Gowher, Humaira; Jeltsch, Albert (2002): Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. In *European journal of biochemistry* 269 (20), pp. 4981–4984.

Filion, Guillaume J.; van Bemmel, Joke G.; Braunschweig, Ulrich; Talhout, Wendy; Kind, Jop; Ward, Lucas D. et al. (2010): Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. In *Cell* 143 (2), pp. 212–224.

Ford, Ethan; Grimmer, Matthew R.; Stolzenburg, Sabine; Bogdanovic, Ozren; Mendoza, Alex de; Farnham, Peggy J. et al. (2017): Frequent lack of repressive capacity of promoter DNA methylation identified through genome-wide epigenomic manipulation (11).

Frommer, M.; McDonald, L. E.; Millar, D. S.; Collis, C. M.; Watt, F.; Grigg, G. W. et al. (1992): A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. In *Proceedings of the National Academy of Sciences of the United States of America* 89 (5), pp. 1827–1831.

Fryxell, Karl J.; Moon, Won-Jong (2005): CpG mutation rates in the human genome are highly dependent on local GC content. In *Molecular biology and evolution* 22 (3), pp. 650–658.

Fu, Kai; Bonora, Giancarlo; Pellegrini, Matteo (2020): Interactions between core histone marks and DNA methyltransferases predict DNA methylation patterns observed in human cells and tissues. In *Epigenetics* 15 (3), pp. 272–282.

Fujita, Naoyuki; Watanabe, Sugiko; Ichimura, Takaya; Tsuruzoe, Shu; Shinkai, Yoichi; Tachibana, Makoto et al. (2003): Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. In *The Journal of biological chemistry* 278 (26), pp. 24132–24138.

Fuks, Francois; Hurd, Paul J.; Wolf, Daniel; Nan, Xinsheng; Bird, Adrian P.; Kouzarides, Tony (2003a): The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. In *The Journal of biological chemistry* 278 (6), pp. 4035–4040.

Fuks, François; Hurd, Paul J.; Deplus, Rachel; Kouzarides, Tony (2003b): The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. In *Nucleic acids research* 31 (9), pp. 2305–2312.

Fyodorov, Dmitry V.; Zhou, Bing-Rui; Skoultchi, Arthur I.; Bai, Yawen (2018): Emerging roles of linker histones in regulating chromatin structure and function. In *Nature reviews. Molecular cell biology* 19 (3), pp. 192–206.

Gabory, Anne; Ripoché, Marie-Anne; Le Digarcher, Anne; Watrin, Françoise; Ziyat, Ahmed; Forné, Thierry et al. (2009): H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. In *Development (Cambridge, England)* 136 (20), pp. 3413–3421.

Galonska, Christina; Charlton, Jocelyn; Mattei, Alexandra L.; Donaghey, Julie; Clement, Kendall; Gu, Hongcang et al. (2018): Genome-wide tracking of dCas9-methyltransferase footprints. In *Nature communications* 9 (1), p. 597.

Gao, Linfeng; Emperle, Max; Guo, Yiran; Grimm, Sara A.; Ren, Wendan; Adam, Sabrina et al. (2020): Comprehensive structure-function characterization of DNMT3B and DNMT3A reveals distinctive de novo DNA methylation mechanisms. In *Nature communications* 11 (1), p. 3355.

García-Cao, Marta; O'Sullivan, Roderick; Peters, Antoine H. F. M.; Jenuwein, Thomas; Blasco, María A. (2004): Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. In *Nature genetics* 36 (1), pp. 94–99.

Gaston, K.; Fried, M. (1995): CpG methylation has differential effects on the binding of YY1 and ETS proteins to the bi-directional promoter of the Surf-1 and Surf-2 genes. In *Nucleic acids research* 23 (6), pp. 901–909.

Gentles, A. J.; Karlin, S. (2001): Genome-scale compositional comparisons in eukaryotes. In *Genome research* 11 (4), pp. 540–546.

Gersbach, Charles A.; Gaj, Thomas; Barbas, Carlos F. (2014): Synthetic zinc finger proteins: the advent of targeted gene regulation and genome modification technologies. In *Accounts of chemical research* 47 (8), pp. 2309–2318.

Gibson, Daniel G.; Young, Lei; Chuang, Ray-Yuan; Venter, J. Craig; Hutchison, Clyde A.; Smith, Hamilton O. (2009): Enzymatic assembly of DNA molecules up to several hundred kilobases. In *Nature methods* 6 (5), pp. 343–345.

Gicquel, Christine; Rossignol, Sylvie; Cabrol, Sylvie; Houang, Muriel; Steunou, Virginie; Barbu, Véronique et al. (2005): Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. In *Nature genetics* 37 (9), pp. 1003–1007.

Ginno, Paul Adrian; Gaidatzis, Dimos; Feldmann, Angelika; Hoerner, Leslie; Imanci, Dilek; Burger, Lukas et al. (2020): A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. In *Nature communications* 11 (1), p. 2680.

Globisch, Daniel; Münzel, Martin; Müller, Markus; Michalakis, Stylianos; Wagner, Mirko; Koch, Susanne et al. (2010): Tissue distribution of 5-hydroxymethylcytosine and search for active de-methylation intermediates. In *PLoS one* 5 (12), e15367.

Glozak, M. A.; Seto, E. (2007): Histone deacetylases and cancer. In *Oncogene* 26 (37), pp. 5420–5432.

Gonzalo, Susana; Jaco, Isabel; Fraga, Mario F.; Chen, Taiping; Li, En; Esteller, Manel; Blasco, María A. (2006): DNA methyltransferases control telomere length and telomere recombination in mammalian cells. In *Nature cell biology* 8 (4), pp. 416–424.

Gordon, J.A.R.; Grandy, R. A.; Lian, J. B.; Stein, J. L.; van Wijnen, A. J.; Stein, G. S. (2013): Chromatin. In: *Brenner's Encyclopedia of Genetics*, vol. 278: Elsevier, pp. 538–541.

Gouil, Quentin; Keniry, Andrew (2019): Latest techniques to study DNA methylation. In *Essays in biochemistry* 63 (6), pp. 639–648.

Gowher, Humaira; Jeltsch, Albert (2018): Mammalian DNA methyltransferases: new discoveries and open questions. In *Biochemical Society transactions* 46 (5), pp. 1191–1202.

Gowher, Humaira; Liebert, Kirsten; Hermann, Andrea; Xu, Guoliang; Jeltsch, Albert (2005): Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. In *The Journal of biological chemistry* 280 (14), pp. 13341–13348.

Gowher, Humaira; Loutchanwoot, Panida; Vorobjeva, Olga; Handa, Vikas; Jurkowska, Renata Z.; Jurkowski, Tomasz P.; Jeltsch, Albert (2006): Mutational analysis of the catalytic domain of the murine Dnmt3a DNA-(cytosine C5)-methyltransferase. In *Journal of molecular biology* 357 (3), pp. 928–941.

Graham, F. L.; Smiley, J.; Russell, W. C.; Nairn, R. (1977): Characteristics of a human cell line transformed by DNA from human adenovirus type 5. In *The Journal of general virology* 36 (1), pp. 59–74.

Greenberg, Maxim V. C.; Bourc'his, Deborah (2019): The diverse roles of DNA methylation in mammalian development and disease. In *Nature reviews. Molecular cell biology* 20 (10), pp. 590–607.

Gujar, Hemant; Weisenberger, Daniel J.; Liang, Gangning (2019): The Roles of Human DNA Methyltransferases and Their Isoforms in Shaping the Epigenome. In *Genes* 10 (2).

Gupta, Rajat M.; Musunuru, Kiran (2014): Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. In *The Journal of clinical investigation* 124 (10), pp. 4154–4161.

Hagarman, James A.; Motley, Michael P.; Kristjansdottir, Katla; Soloway, Paul D. (2013): Coordinate regulation of DNA methylation and H3K27me3 in mouse embryonic stem cells. In *PloS one* 8 (1), e53880.

Hahn, Maria A.; Szabó, Piroska E.; Pfeifer, Gerd P. (2014): 5-Hydroxymethylcytosine: a stable or transient DNA modification? In *Genomics* 104 (5), pp. 314–323.

Hanna, Courtney W.; Demond, Hannah; Kelsey, Gavin (2018): Epigenetic regulation in development: is the mouse a good model for the human? In *Human reproduction update* 24 (5), pp. 556–576.

Hanzawa, Nozomi; Hashimoto, Koshi; Yuan, Xunmei; Kawahori, Kenichi; Tsujimoto, Kazutaka; Hamaguchi, Miho et al. (2020): Targeted DNA demethylation of the Fgf21 promoter by CRISPR/dCas9-mediated epigenome editing. In *Scientific reports* 10 (1), p. 5181.

Hark, A. T.; Schoenherr, C. J.; Katz, D. J.; Ingram, R. S.; Levorse, J. M.; Tilghman, S. M. (2000): CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. In *Nature* 405 (6785), pp. 486–489.

Hashimoto, Hideharu; Vertino, Paula M.; Cheng, Xiaodong (2010): Molecular coupling of DNA methylation and histone methylation. In *Epigenomics* 2 (5), pp. 657–669.

Hashimoto, Hideharu; Wang, Dongxue; Horton, John R.; Zhang, Xing; Corces, Victor G.; Cheng, Xiaodong (2017): Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA. In *Molecular cell* 66 (5), 711-720.e3.

Hata, Kenichiro; Okano, Masaki; Lei, Hong; Li, En (2002): Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. In *Development* (Cambridge, England) 129 (8), pp. 1983–1993.

Hatano, Akiko; Chiba, Hirokazu; Moesa, Harry Amri; Taniguchi, Takeaki; Nagaie, Satoshi; Yamanegi, Koji et al. (2011): CELLPEDIA: a repository for human cell information for cell studies and differentiation analyses. In *Database : the journal of biological databases and curation* 2011, bar046.

Heard, Edith; Disteche, Christine M. (2006): Dosage compensation in mammals: fine-tuning the expression of the X chromosome. In *Genes & development* 20 (14), pp. 1848–1867.

Heintzman, Nathaniel D.; Hon, Gary C.; Hawkins, R. David; Kheradpour, Pouya; Stark, Alexander; Harp, Lindsey F. et al. (2009): Histone modifications at human enhancers reflect global cell-type-specific gene expression. In *Nature* 459 (7243), pp. 108–112.

Heintzman, Nathaniel D.; Stuart, Rhona K.; Hon, Gary; Fu, Yutao; Ching, Christina W.; Hawkins, R. David et al. (2007): Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. In *Nature genetics* 39 (3), pp. 311–318.

Heyn, Patricia; Logan, Clare V.; Fluteau, Adeline; Challis, Rachel C.; Auchynnikava, Tatsiana; Martin, Carol-Anne et al. (2019): Gain-of-function DNMT3A mutations cause microcephalic dwarfism and hypermethylation of Polycomb-regulated regions. In *Nature genetics* 51 (1), pp. 96–105.

Hilton, Isaac B.; D'Ippolito, Anthony M.; Vockley, Christopher M.; Thakore, Pratiksha I.; Crawford, Gregory E.; Reddy, Timothy E.; Gersbach, Charles A. (2015): Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. In *Nature biotechnology* 33 (5), pp. 510–517.

Hofacker, Daniel; Broche, Julian; Laistner, Laura; Adam, Sabrina; Bashtrykov, Pavel; Jeltsch, Albert (2020): Engineering of Effector Domains for Targeted DNA Methylation with Reduced Off-Target Effects. In *International journal of molecular sciences* 21 (2).

Hon, Gary C.; Hawkins, R. David; Caballero, Otavia L.; Lo, Christine; Lister, Ryan; Pelizzola, Mattia et al. (2012): Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. In *Genome research* 22 (2), pp. 246–258.

Hon, Gary C.; Song, Chun-Xiao; Du, Tingting; Jin, Fulai; Selvaraj, Siddarth; Lee, Ah Young et al. (2014): 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. In *Molecular cell* 56 (2), pp. 286–297.

Horii, Takuro; Morita, Sumiyo; Hino, Shinjiro; Kimura, Mika; Hino, Yuko; Kogo, Hiroshi et al. (2020): Successful generation of epigenetic disease model mice by targeted demethylation of the epigenome. In *Genome biology* 21 (1), p. 77.

Hotchkiss, R. D. (1948): The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. In *The Journal of biological chemistry* 175 (1), pp. 315–332.

Howe, Françoise S.; Fischl, Harry; Murray, Struan C.; Mellor, Jane (2017): Is H3K4me3 instructive for transcription activation? In *BioEssays : news and reviews in molecular, cellular and developmental biology* 39 (1), pp. 1–12.

Huang, He; Lin, Shu; Garcia, Benjamin A.; Zhao, Yingming (2015): Quantitative proteomic analysis of histone modifications. In *Chemical reviews* 115 (6), pp. 2376–2418.

Huang, Yung-Hsin; Su, Jianzhong; Lei, Yong; Brunetti, Lorenzo; Gundry, Michael C.; Zhang, Xiaotian et al. (2017): DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A. In *Genome biology* 18 (1), p. 176.

Hyun, Kwangbeom; Jeon, Jongcheol; Park, Kihyun; Kim, Jaehoon (2017): Writing, erasing and reading histone lysine methylations. In *Experimental & molecular medicine* 49 (4), e324.

Ichimura, Takaya; Watanabe, Sugiko; Sakamoto, Yasuo; Aoto, Takahiro; Fujita, Naoyuki; Nakao, Mitsuyoshi (2005): Transcriptional repression and heterochromatin formation by MBD1 and MCAF/AM family proteins. In *The Journal of biological chemistry* 280 (14), pp. 13928–13935.

Iida, Tetsuo; Suetake, Isao; Tajima, Shoji; Morioka, Hiroshi; Ohta, Satoshi; Obuse, Chikashi; Tsunamoto, Toshiki (2002): PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. In *Genes to cells : devoted to molecular & cellular mechanisms* 7 (10), pp. 997–1007.

Ito, Shinsuke; Shen, Li; Dai, Qing; Wu, Susan C.; Collins, Leonard B.; Swenberg, James A. et al. (2011): Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. In *Science (New York, N.Y.)* 333 (6047), pp. 1300–1303.

Iurlaro, Mario; McInroy, Gordon R.; Burgess, Heather E.; Dean, Wendy; Raiber, Eun-Ang; Bachman, Martin et al. (2016): In vivo genome-wide profiling reveals a tissue-specific role for 5-formylcytosine. In *Genome biology* 17 (1), p. 141.

Jacobsen, Linda; Calvin, Susan; Lobenhofer, Edward (2009): Transcriptional effects of transfection: the potential for misinterpretation of gene expression data generated from transiently transfected cells. In *BioTechniques* 47 (1), pp. 617–624.

Jeltsch, Albert (2002): Beyond Watson and Crick: DNA Methylation and Molecular Enzymology of DNA Methyltransferases. In *ChemBioChem* 3 (4), pp. 274–293.

Jeltsch, Albert; Broche, Julian; Bashtrykov, Pavel (2018): Molecular Processes Connecting DNA Methylation Patterns with DNA Methyltransferases and Histone Modifications in Mammalian Genomes. In *Genes* 9 (11).

Jeltsch, Albert; Broche, Julian; Lungu, Cristiana; Bashtrykov, Pavel (2019): Biotechnological Applications of MBD Domain Proteins for DNA Methylation Analysis. In *Journal of molecular biology*.

Jeltsch, Albert; Jurkowska, Renata Z. (2013): Multimerization of the dnmt3a DNA methyltransferase and its functional implications. In *Progress in molecular biology and translational science* 117, pp. 445–464.

Jeltsch, Albert; Jurkowska, Renata Z. (2014): New concepts in DNA methylation. In *Trends in biochemical sciences* 39 (7), pp. 310–318.

Jeltsch, Albert; Jurkowska, Renata Z. (2016): Allosteric control of mammalian DNA methyltransferases - a new regulatory paradigm. In *Nucleic acids research* 44 (18), pp. 8556–8575.

Jeltsch, Albert; Lanio, Thomas (2002): Site-directed mutagenesis by polymerase chain reaction. In *Methods in molecular biology* (Clifton, N.J.) 182, pp. 85–94.

Jenuwein, T.; Allis, C. D. (2001): Translating the histone code. In *Science* (New York, N.Y.) 293 (5532), pp. 1074–1080.

Jeong, Hyeonsoo; Mendizabal, Isabel; Berto, Stefano; Chatterjee, Paramita; Layman, Thomas; Usui, Noriyoshi et al. (2021): Evolution of DNA methylation in the human brain. In *Nature communications* 12 (1), p. 2021.

Jeong, Mira; Goodell, Margaret A. (2014): New answers to old questions from genome-wide maps of DNA methylation in hematopoietic cells. In *Experimental hematology* 42 (8), pp. 609–617.

Jeong, Mira; Sun, Deqiang; Luo, Min; Huang, Yun; Challen, Grant A.; Rodriguez, Benjamin et al. (2014): Large conserved domains of low DNA methylation maintained by Dnmt3a. In *Nature genetics* 46 (1), pp. 17–23.

Jeong, Shinwu; Liang, Gangning; Sharma, Shikhar; Lin, Joy C.; Choi, Si Ho; Han, Han et al. (2009): Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. In *Molecular and cellular biology* 29 (19), pp. 5366–5376.

Jeziorska, Danuta M.; Murray, Robert J. S.; Gobbi, Marco de; Gaentzsch, Ricarda; Garrick, David; Ayyub, Helena et al. (2017): DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. In *Proceedings of the National Academy of Sciences of the United States of America* 114 (36), E7526-E7535.

Jia, Da; Jurkowska, Renata Z.; Zhang, Xing; Jeltsch, Albert; Cheng, Xiaodong (2007): Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. In *Nature* 449 (7159), pp. 248–251.

Jia, Yuanhui; Li, Pishun; Fang, Lan; Zhu, Haijun; Xu, Liangliang; Cheng, Hao et al. (2016): Negative regulation of DNMT3A de novo DNA methylation by frequently overexpressed UHRF family proteins as a mechanism for widespread DNA hypomethylation in cancer. In *Cell discovery* 2, p. 16007.

Jin, Zelin; Liu, Yun (2018): DNA methylation in human diseases. In *Genes & diseases* 5 (1), pp. 1–8.

Jinek, Martin; Chylinski, Krzysztof; Fonfara, Ines; Hauer, Michael; Doudna, Jennifer A.; Charpentier, Emmanuelle (2012): A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. In *Science* (New York, N.Y.) 337 (6096), pp. 816–821.

Johnson, Treat B.; Coghill, Robert D. (1925): RESEARCHES ON PYRIMIDINES. C111. THE DISCOVERY OF 5-METHYL-CYTOSINE IN TUBERCULINIC ACID, THE NUCLEIC ACID OF THE TUBERCLE BACILLUS 1. In *J. Am. Chem. Soc.* 47 (11), pp. 2838–2844.

Jones, Peter A. (2012): Functions of DNA methylation: islands, start sites, gene bodies and beyond. In *Nature reviews. Genetics* 13 (7), pp. 484–492.

Jones, Peter A.; Liang, Gangning (2009): Rethinking how DNA methylation patterns are maintained. In *Nature reviews. Genetics* 10 (11), pp. 805–811.

Jurkowska, Renata Z.; Anspach, Nils; Urbanke, Claus; Jia, Da; Reinhardt, Richard; Nellen, Wolfgang et al. (2008): Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. In *Nucleic acids research* 36 (21), pp. 6656–6663.

Jurkowska, Renata Z.; Rajavelu, Arumugam; Anspach, Nils; Urbanke, Claus; Jankevicius, Gytis; Ragozin, Sergey et al. (2011): Oligomerization and binding of the Dnmt3a DNA methyltransferase to parallel DNA molecules: heterochromatic localization and role of Dnmt3L. In *The Journal of biological chemistry* 286 (27), pp. 24200–24207.

Kareta, Michael S.; Botello, Zaida M.; Ennis, Joshua J.; Chou, Christina; Chédin, Frédéric (2006): Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. In *The Journal of biological chemistry* 281 (36), pp. 25893–25902.

Karlič, Rosa; Chung, Ho-Ryun; Lasserre, Julia; Vlahovicek, Kristian; Vingron, Martin (2010): His-tone modification levels are predictive for gene expression. In *Proceedings of the National Academy of Sciences of the United States of America* 107 (7), pp. 2926–2931.

Karvelis, Tautvydas; Gasiunas, Giedrius; Young, Joshua; Bigelyte, Greta; Silanskas, Arunas; Cigan, Mark; Siksnys, Virginijus (2015): Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. In *Genome biology* 16, p. 253.

Kasinath, Vignesh; Beck, Curtis; Sauer, Paul; Poepfel, Simon; Kosmatka, Jennifer; Faini, Marco et al. (2021): JARID2 and AEBP2 regulate PRC2 in the presence of H2AK119ub1 and other his-tone modifications. In *Science (New York, N.Y.)* 371 (6527).

Kernohan, Kristin D.; Vernimmen, Douglas; Gloor, Gregory B.; Bérubé, Nathalie G. (2014): Analysis of neonatal brain lacking ATRX or MeCP2 reveals changes in nucleosome density, CTCF binding and chromatin looping. In *Nucleic acids research* 42 (13), pp. 8356–8368.

Kim, Moon-Soo; Kini, Anu Ganesh (2017): Engineering and Application of Zinc Finger Proteins and TALEs for Biomedical Research. In *Molecules and cells* 40 (8), pp. 533–541.

Kim, Tae Hoon; Abdullaev, Ziedulla K.; Smith, Andrew D.; Ching, Keith A.; Loukinov, Dmitri I.; Green, Roland D. et al. (2007): Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. In *Cell* 128 (6), pp. 1231–1245.

King, Andrew D.; Huang, Kevin; Rubbi, Liudmilla; Liu, Shuo; Wang, Cun-Yu; Wang, Yinsheng et al. (2016): Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells. In *Cell reports* 17 (1), pp. 289–302.

Knutson, Sarah K.; Kawano, Satoshi; Minoshima, Yukinori; Warholic, Natalie M.; Huang, Kuan-Chun; Xiao, Yonghong et al. (2014): Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma. In *Molecular cancer therapeutics* 13 (4), pp. 842–854.

Kong, Lingchun; Tan, Li; Lv, Ruitu; Shi, Zhennan; Xiong, Lijun; Wu, Feizhen et al. (2016): A primary role of TET proteins in establishment and maintenance of De Novo bivalency at CpG islands. In *Nucleic acids research* 44 (18), pp. 8682–8692.

Kouzarides, Tony (2007): Chromatin modifications and their function. In *Cell* 128 (4), pp. 693–705.

Kozlenkov, Alexey; Wang, Minghui; Roussos, Panos; Rudchenko, Sergei; Barbu, Mihaela; Bibikova, Marina et al. (2016): Substantial DNA methylation differences between two major neuronal subtypes in human brain. In *Nucleic acids research* 44 (6), pp. 2593–2612.

Kumar, S.; Cheng, X.; Klimasauskas, S.; Mi, S.; Posfai, J.; Roberts, R. J.; Wilson, G. G. (1994): The DNA (cytosine-5) methyltransferases. In *Nucleic acids research* 22 (1), pp. 1–10.

Kungulovski, Goran; Jeltsch, Albert (2016): Epigenome Editing: State of the Art, Concepts, and Perspectives. In *Trends in genetics : TIG* 32 (2), pp. 101–113.

Kungulovski, Goran; Nunna, Suneetha; Thomas, Maria; Zanger, Ulrich M.; Reinhardt, Richard; Jeltsch, Albert (2015): Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. In *Epigenetics & chromatin* 8, p. 12.

Kurukuti, Sreenivasulu; Tiwari, Vijay Kumar; Tavoosidana, Gholamreza; Pugacheva, Elena; Murrell, Adele; Zhao, Zhihu et al. (2006): CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. In *Proceedings of the National Academy of Sciences of the United States of America* 103 (28), pp. 10684–10689.

Kwon, Deborah Y.; Zhao, Ying-Tao; Lamonica, Janine M.; Zhou, Zhaolan (2017): Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. In *Nature communications* 8, p. 15315.

Laget, Sophie; Joulie, Michael; Le Masson, Florent; Sasai, Nobuhiro; Christians, Elisabeth; Prahdan, Sriharsa et al. (2010): The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA. In *PloS one* 5 (8), e11982.

Larson, Adam G.; Narlikar, Geeta J. (2018): The Role of Phase Separation in Heterochromatin Formation, Function, and Regulation. In *Biochemistry* 57 (17), pp. 2540–2548.

Lauberth, Shannon M.; Nakayama, Takahiro; Wu, Xiaolin; Ferris, Andrea L.; Tang, Zhanyun; Hughes, Stephen H.; Roeder, Robert G. (2013): H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. In *Cell* 152 (5), pp. 1021–1036.

Laufer, Benjamin I.; Singh, Shiva M. (2015): Strategies for precision modulation of gene expression by epigenome editing: an overview. In *Epigenetics & chromatin* 8, p. 34.

Lebar, Tina; Lainšček, Duško; Merljak, Estera; Aupič, Jana; Jerala, Roman (2020): A tunable orthogonal coiled-coil interaction toolbox for engineering mammalian cells. In *Nature chemical biology* 16 (5), pp. 513–519.

Lee, Gun E.; Kim, Joo Hee; Taylor, Michael; Muller, Mark T. (2010): DNA methyltransferase 1-associated protein (DNMT1) is a co-repressor that stimulates DNA methylation globally and locally at sites of double strand break repair. In *The Journal of biological chemistry* 285 (48), pp. 37630–37640.

Lee, Heather J.; Hore, Timothy A.; Reik, Wolf (2014): Reprogramming the methylome: erasing memory and creating diversity. In *Cell stem cell* 14 (6), pp. 710–719.

Lehnertz, Bernhard; Ueda, Yoshihide; Derijck, Alwin A.H.A.; Braunschweig, Ulrich; Perez-Burgos, Laura; Kubicek, Stefan et al. (2003): Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. In *Current Biology* 13 (14), pp. 1192–1200.

Lev Maor, Galit; Yearim, Ahuvi; Ast, Gil (2015): The alternative role of DNA methylation in splicing regulation. In *Trends in genetics* : TIG 31 (5), pp. 274–280.

Li, En; Bestor, Timothy H.; Jaenisch, Rudolf (1992): Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. In *Cell* 69 (6), pp. 915–926.

Li, Haojie; Liefke, Robert; Jiang, Junyi; Kurland, Jesse Vigoda; Tian, Wei; Deng, Pujuan et al. (2017): Polycomb-like proteins link the PRC2 complex to CpG islands. In *Nature* 549 (7671), pp. 287–291.

Li, Hongwei; Rauch, Tibor; Chen, Zhao-Xia; Szabó, Piroska E.; Riggs, Arthur D.; Pfeifer, Gerd P. (2006): The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A inter-act directly and localize to promoters silenced in cancer cells. In *The Journal of biological chemistry* 281 (28), pp. 19489–19500.

Li, Jie; Ahn, Jeong Hyun; Wang, Gang Greg (2019): Understanding histone H3 lysine 36 methylation and its deregulation in disease. In *Cellular and molecular life sciences* : CMLS 76 (15), pp. 2899–2916.

Li, Jingting; Ohliger, James; Pei, Ming (2014): Significance of epigenetic landscape in cartilage regeneration from the cartilage development and pathology perspective. In *Stem cells and development* 23 (11), pp. 1178–1194.

Li, Tao; Wang, Linsheng; Du, Yongming; Xie, Si; Yang, Xi; Lian, Fuming et al. (2018): Structural and mechanistic insights into UHRF1-mediated DNMT1 activation in the maintenance DNA methylation. In *Nucleic acids research* 46 (6), pp. 3218–3231.

Li, Yingrui; Zhu, Jingde; Tian, Geng; Li, Ning; Li, Qibin; Ye, Mingzhi et al. (2010): The DNA methylome of human peripheral blood mononuclear cells. In *PLoS biology* 8 (11), e1000533.

Lin, Chien-Chu; Chen, Yi-Ping; Yang, Wei-Zen; Shen, James C. K.; Yuan, Hanna S. (2020): Structural insights into CpG-specific DNA methylation by human DNA methyltransferase 3B. In *Nucleic acids research* 48 (7), pp. 3949–3961.

Lin, Yao-Cheng; Boone, Morgane; Meuris, Leander; Lemmens, Irma; van Roy, Nadine; Soete, Arne et al. (2014): Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. In *Nature communications* 5, p. 4767.

Lisanti, Sofia; Omar, Wan A. W.; Tomaszewski, Bartłomiej; Prins, Sofie de; Jacobs, Griet; Koppen, Gudrun et al. (2013): Comparison of methods for quantification of global DNA methylation in human cells and tissues. In *PLoS one* 8 (11), e79044.

Lister, Ryan; Pelizzola, Mattia; Downen, Robert H.; Hawkins, R. David; Hon, Gary; Tonti-Filipini, Julian et al. (2009): Human DNA methylomes at base resolution show widespread epigenomic differences. In *Nature* 462 (7271), pp. 315–322.

Liu, P. Q.; Rebar, E. J.; Zhang, L.; Liu, Q.; Jamieson, A. C.; Liang, Y. et al. (2001): Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. In *The Journal of biological chemistry* 276 (14), pp. 11323–11334.

Liu, X. Shawn; Wu, Hao; Ji, Xiong; Stelzer, Yonatan; Wu, Xuebing; Czauderna, Szymon et al. (2016): Editing DNA Methylation in the Mammalian Genome. In *Cell* 167 (1), 233-247.e17.

- Liu, Zhixi; Liao, Zhi; Chen, Yan; Han, Lizhu; Yin, Qinan; Xiao, Hongtao (2020): Application of Various Delivery Methods for CRISPR/dCas9. In *Molecular biotechnology* 62 (8), pp. 355–363.
- Long, Hannah K.; King, Hamish W.; Patient, Roger K.; Odom, Duncan T.; Klose, Robert J. (2016): Protection of CpG islands from DNA methylation is DNA-encoded and evolutionarily conserved. In *Nucleic acids research* 44 (14), pp. 6693–6706.
- Lorsbach, R. B.; Moore, J.; Mathew, S.; Raimondi, S. C.; Mukatira, S. T.; Downing, J. R. (2003): TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). In *Leukemia* 17 (3), pp. 637–641.
- Luco, Reini F.; Pan, Qun; Tominaga, Kaoru; Blencowe, Benjamin J.; Pereira-Smith, Olivia M.; Misteli, Tom (2010): Regulation of alternative splicing by histone modifications. In *Science (New York, N.Y.)* 327 (5968), pp. 996–1000.
- Lunyak, Victoria V.; Burgess, Robert; Prefontaine, Gratien G.; Nelson, Charles; Sze, Sing-Hoi; Chenoweth, Josh et al. (2002): Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. In *Science (New York, N.Y.)* 298 (5599), pp. 1747–1752.
- Luo, Guan-Zheng; Blanco, Mario Andres; Greer, Eric Lieberman; He, Chuan; Shi, Yang (2015): DNA N(6)-methyladenine: a new epigenetic mark in eukaryotes? In *Nature reviews. Molecular cell biology* 16 (12), pp. 705–710.
- Luo, Minkui (2018): Chemical and Biochemical Perspectives of Protein Lysine Methylation. In *Chemical reviews* 118 (14), pp. 6656–6705.
- Lynch, Magnus D.; Smith, Andrew J. H.; Gobbi, Marco de; Flenley, Maria; Hughes, Jim R.; Ver-nimmen, Douglas et al. (2012): An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. In *The EMBO journal* 31 (2), pp. 317–329.
- Lyst, Matthew J.; Ekiert, Robert; Ebert, Daniel H.; Merusi, Cara; Nowak, Jakub; Selfridge, Jim et al. (2013): Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. In *Nature neuroscience* 16 (7), pp. 898–902.
- Machida, Shinichi; Takizawa, Yoshimasa; Ishimaru, Masakazu; Sugita, Yukihiro; Sekine, Satoshi; Nakayama, Jun-Ichi et al. (2018): Structural Basis of Heterochromatin Formation by Human HP1. In *Molecular cell* 69 (3), 385-397.e8.
- Maeder, Morgan L.; Angstman, James F.; Richardson, Marcy E.; Linder, Samantha J.; Cascio, Vincent M.; Tsai, Shengdar Q. et al. (2013): Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. In *Nature biotechnology* 31 (12), pp. 1137–1142.
- Mahé, Elise A.; Madigou, Thierry; Sérandour, Aurélien A.; Bizot, Maud; Avner, Stéphane; Chalmel, Frédéric et al. (2017): Cytosine modifications modulate the chromatin architecture of transcriptional enhancers. In *Genome research* 27 (6), pp. 947–958.
- Mak, Amanda Nga-Sze; Bradley, Philip; Bogdanove, Adam J.; Stoddard, Barry L. (2013): TAL effectors: function, structure, engineering and applications. In *Current opinion in structural biology* 23 (1), pp. 93–99.

Manzo, Massimiliano; Wirz, Joël; Ambrosi, Christina; Villaseñor, Rodrigo; Roschitzki, Bernd; Baubec, Tuncay (2017): Isoform-specific localization of DNMT3A regulates DNA methylation fidelity at bivalent CpG islands. In *The EMBO journal* 36 (23), pp. 3421–3434.

Margueron, Raphaël; Reinberg, Danny (2011): The Polycomb complex PRC2 and its mark in life. In *Nature* 469 (7330), pp. 343–349.

Marholz, Laura J.; Wang, Wei; Zheng, Yu; Wang, Xiang (2016): A Fluorescence Polarization Bio-physical Assay for the Naegleria DNA Hydroxylase Tet1. In *ACS medicinal chemistry letters* 7 (2), pp. 167–171.

Markusic, David; Oude-Elferink, Ronald; Das, Atze T.; Berkhout, Ben; Seppen, Jurgen (2005): Comparison of single regulated lentiviral vectors with rtTA expression driven by an autoregulatory loop or a constitutive promoter. In *Nucleic acids research* 33 (6), e63.

Martins, Nuno M. C.; Cisneros-Soberanis, Fernanda; Pesenti, Elisa; Kochanova, Natalia Y.; Shang, Wei-Hao; Hori, Tetsuya et al. (2020): H3K9me3 maintenance on a human artificial chromosome is required for segregation but not centromere epigenetic memory. In *Journal of cell science* 133 (14).

Marzese, Diego M.; Hoon, Dave Sb (2015): Emerging technologies for studying DNA methylation for the molecular diagnosis of cancer. In *Expert review of molecular diagnostics* 15 (5), pp. 647–664.

Meissner, Alexander; Mikkelsen, Tarjei S.; Gu, Hongchang; Wernig, Marius; Hanna, Jacob; Sivachenko, Andrey et al. (2008): Genome-scale DNA methylation maps of pluripotent and differentiated cells. In *Nature* 454 (7205), pp. 766–770.

Melamed, Philippa; Yosefzon, Yahav; David, Cfir; Tsukerman, Anna; Pnueli, Lilach (2018): Tet Enzymes, Variants, and Differential Effects on Function. In *Frontiers in cell and developmental biology* 6, p. 22.

Mellén, Marian; Ayata, Pinar; Dewell, Scott; Kriaucionis, Skirmantas; Heintz, Nathaniel (2012): MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. In *Cell* 151 (7), pp. 1417–1430.

Mikkelsen, Tarjei S.; Ku, Manching; Jaffe, David B.; Issac, Biju; Lieberman, Erez; Giannoukos, Georgia et al. (2007): Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. In *Nature* 448 (7153), pp. 553–560.

Monaghan, Laura; Massett, Matthew E.; Bunschoten, Roderick P.; Hoose, Alex; Pirvan, Petrisor-Alin; Liskamp, Robert M. J. et al. (2019): The Emerging Role of H3K9me3 as a Potential Therapeutic Target in Acute Myeloid Leukemia. In *Frontiers in oncology* 9, p. 705.

Morgan, Richard A.; Gray, David; Lomova, Anastasia; Kohn, Donald B. (2017): Hematopoietic Stem Cell Gene Therapy: Progress and Lessons Learned. In *Cell stem cell* 21 (5), pp. 574–590.

Morita, S.; Kojima, T.; Kitamura, T. (2000): Plat-E: an efficient and stable system for transient packaging of retroviruses. In *Gene therapy* 7 (12), pp. 1063–1066.

Morita, Sumiyo; Noguchi, Hirofumi; Horii, Takuro; Nakabayashi, Kazuhiko; Kimura, Mika; Okamura, Kohji et al. (2016): Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. In *Nature biotechnology* 34 (10), pp. 1060–1065.

Mossman, David; Scott, Rodney J. (2006): Epimutations, inheritance and causes of aberrant DNA methylation in cancer. In *Hereditary cancer in clinical practice* 4 (2), pp. 75–80.

Moulton, T.; Crenshaw, T.; Hao, Y.; Moosikasuwan, J.; Lin, N.; Dembitzer, F. et al. (1994): Epigenetic lesions at the H19 locus in Wilms' tumour patients. In *Nature genetics* 7 (3), pp. 440–447.

Mulholland, Christopher B.; Traube, Franziska R.; Ugur, Enes; Parsa, Edris; Eckl, Eva-Maria; Schönung, Maximilian et al. (2020): Distinct and stage-specific contributions of TET1 and TET2 to stepwise cytosine oxidation in the transition from naive to primed pluripotency. In *Scientific reports* 10 (1), p. 12066.

Murrell, Adele; Heeson, Sarah; Cooper, Wendy N.; Douglas, Eleanor; Apostolidou, Sophia; Moore, Gudrun E. et al. (2004): An association between variants in the IGF2 gene and Beckwith-Wiedemann syndrome: interaction between genotype and epigenotype. In *Human molecular genetics* 13 (2), pp. 247–255.

Nativio, Raffaella; Sparago, Angela; Ito, Yoko; Weksberg, Rosanna; Riccio, Andrea; Murrell, Adele (2011): Disruption of genomic neighbourhood at the imprinted IGF2-H19 locus in Beckwith-Wiedemann syndrome and Silver-Russell syndrome. In *Human molecular genetics* 20 (7), pp. 1363–1374.

Nativio, Raffaella; Wendt, Kerstin S.; Ito, Yoko; Huddleston, Joanna E.; Uribe-Lewis, Santiago; Woodfine, Kathryn et al. (2009): Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. In *PLoS genetics* 5 (11), e1000739.

Nayak, Vinod; Xu, Chao; Min, Jinrong (2011): Composition, recruitment and regulation of the PRC2 complex. In *Nucleus (Austin, Tex.)* 2 (4), pp. 277–282.

Nelson, Joel D.; Denisenko, Oleg; Bomsztyk, Karol (2006): Protocol for the fast chromatin immunoprecipitation (ChIP) method. In *Nature protocols* 1 (1), pp. 179–185.

Neri, Francesco; Incarnato, Danny; Krepelova, Anna; Rapelli, Stefania; Anselmi, Francesca; Par-lato, Caterina et al. (2015): Single-Base Resolution Analysis of 5-Formyl and 5-Carboxyl Cytosine Reveals Promoter DNA Methylation Dynamics. In *Cell reports* 10 (5), pp. 674–683.

Neri, Francesco; Incarnato, Danny; Krepelova, Anna; Rapelli, Stefania; Pagnani, Andrea; Zecchina, Riccardo et al. (2013): Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. In *Genome biology* 14 (8), R91.

Netchine, Irène; Rossignol, Sylvie; Dufourg, Marie-Noëlle; Azzi, Salah; Rousseau, Alexandra; Perin, Laurence et al. (2007): 11p15 imprinting center region 1 loss of methylation is a common and specific cause of typical Russell-Silver syndrome: clinical scoring system and epigenetic-phenotypic correlations. In *The Journal of clinical endocrinology and metabolism* 92 (8), pp. 3148–3154.

Nguyen, Michelle L. T.; Jones, Sarah A.; Prier, Julia E.; Russ, Brendan E. (2015): Transcriptional Enhancers in the Regulation of T Cell Differentiation. In *Frontiers in immunology* 6, p. 462.

Noh, Kyung-Min; Wang, Haibo; Kim, Hyunjae R.; Wenderski, Wendy; Fang, Fang; Li, Charles H. et al. (2015): Engineering of a Histone-Recognition Domain in Dnmt3a Alters the Epigenetic Landscape and Phenotypic Features of Mouse ESCs. In *Molecular cell* 59 (1), pp. 89–103.

- Nordin, M.; Bergman, D.; Halje, M.; Engström, W.; Ward, A. (2014): Epigenetic regulation of the Igf2/H19 gene cluster. In *Cell proliferation* 47 (3), pp. 189–199.
- Norvil, Allison B.; Petell, Christopher J.; Alabdi, Lama; Wu, Lan Chen; Rossie, Sandra; Gowher, Humaira (2018): Dnmt3b Methylates DNA by a Noncooperative Mechanism, and Its Activity Is Unaffected by Manipulations at the Predicted Dimer Interface. In *Biochemistry* 57 (29), pp. 4312–4324.
- Nowialis, Pawel; Lopusna, Katarina; Opavska, Jana; Haney, Staci L.; Abraham, Ajay; Sheng, Peike et al. (2019): Catalytically inactive Dnmt3b rescues mouse embryonic development by accessory and repressive functions. In *Nature communications* 10 (1), p. 4374.
- Nuñez, James K.; Chen, Jin; Pommier, Greg C.; Cogan, J. Zachery; Replogle, Joseph M.; Adri-aens, Carmen et al. (2021): Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. In *Cell* 184 (9), 2503-2519.e17.
- O'Geen, Henriette; Bates, Sofie L.; Carter, Sakereh S.; Nisson, Karly A.; Halmai, Julian; Fink, Kyle D. et al. (2019): Ezh2-dCas9 and KRAB-dCas9 enable engineering of epigenetic memory in a context-dependent manner. In *Epigenetics & chromatin* 12 (1), p. 26.
- Okano, Masaki; Bell, Daphne W.; Haber, Daniel A.; Li, En (1999): DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. In *Cell* 99 (3), pp. 247–257.
- O'Neill, L. (2003): Immunoprecipitation of native chromatin: NChIP. In *Methods* 31 (1), pp. 76–82.
- Orlando, Valerio (2000): Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. In *Trends in biochemical sciences* 25 (3), pp. 99–104.
- Park, Elizabeth; Kim, Yunha; Ryu, Hyun; Kowall, Neil W.; Lee, Junghee; Ryu, Hoon (2014): Epi-genetic mechanisms of Rubinstein-Taybi syndrome. In *Neuromolecular medicine* 16 (1), pp. 16–24.
- Pastor, William A.; Pape, Utz J.; Huang, Yun; Henderson, Hope R.; Lister, Ryan; Ko, Myunggon et al. (2011): Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. In *Nature* 473 (7347), pp. 394–397.
- Patnaik, Debasis; Estève, Pierre-Olivier; Pradhan, Sriharsa (2018): Targeting the SET and RING-associated (SRA) domain of ubiquitin-like, PHD and ring finger-containing 1 (UHRF1) for anti-cancer drug development. In *Oncotarget* 9 (40), pp. 26243–26258.
- Peng, Lina; Li, Yan; Xi, Yanping; Li, Wei; Li, Jin; Lv, Ruitu et al. (2016): MBD3L2 promotes Tet2 enzymatic activity for mediating 5-methylcytosine oxidation. In *Journal of cell science* 129 (5), pp. 1059–1071.
- Penn, N. W.; Suwalski, R.; O'Riley, C.; Bojanowski, K.; Yura, R. (1972): The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. In *The Biochemical journal* 126 (4), pp. 781–790.
- Pfaffl, M. W. (2001): A new mathematical model for relative quantification in real-time RT-PCR. In *Nucleic acids research* 29 (9), e45.

Pfister, Sophia X.; Ahrabi, Sara; Zalmas, Lykourgos-Panagiotis; Sarkar, Sovan; Aymard, François; Bachrati, Csanád Z. et al. (2014): SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. In *Cell reports* 7 (6), pp. 2006–2018.

Pflueger, Christian; Tan, Dennis; Swain, Tessa; Nguyen, Trung; Pflueger, Jahnvi; Nefzger, Christian et al. (2018): A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. In *Genome research* 28 (8), pp. 1193–1206.

Pinheiro, Inês; Margueron, Raphaël; Shukeir, Nicholas; Eisold, Michael; Fritsch, Christoph; Richter, Florian M. et al. (2012): Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. In *Cell* 150 (5), pp. 948–960.

Podobinska, Martyna; Szablowska-Gadomska, Ilona; Augustyniak, Justyna; Sandvig, Ioanna; Sandvig, Axel; Buzanska, Leonora (2017): Epigenetic Modulation of Stem Cells in Neurodevelopment: The Role of Methylation and Acetylation. In *Frontiers in cellular neuroscience* 11, p. 23.

Ponchel, Frederique; Toomes, Carmel; Bransfield, Kieran; Leong, Fong T.; Douglas, Susan H.; Field, Sarah L. et al. (2003): Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. In *BMC biotechnology* 3, p. 18.

Portela, Anna; Esteller, Manel (2010): Epigenetic modifications and human disease. In *Nature biotechnology* 28 (10), pp. 1057–1068.

Pradhan, Mihika; Estève, Pierre-Olivier; Chin, Hang Gyeong; Samaranayke, Mala; Kim, Gun-Do; Pradhan, Sriharsa (2008): CXXC domain of human DNMT1 is essential for enzymatic activity. In *Biochemistry* 47 (38), pp. 10000–10009.

Qi, Lei S.; Larson, Matthew H.; Gilbert, Luke A.; Doudna, Jennifer A.; Weissman, Jonathan S.; Arkin, Adam P.; Lim, Wendell A. (2013): Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. In *Cell* 152 (5), pp. 1173–1183.

Qin, Jane Yuxia; Zhang, Li; Clift, Kayla L.; Hular, Imge; Xiang, Andy Peng; Ren, Bing-Zhong; Lahn, Bruce T. (2010): Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. In *PloS one* 5 (5), e10611.

Qin, Su; Min, Jinrong (2014): Structure and function of the nucleosome-binding PWWP domain. In *Trends in biochemical sciences* 39 (11), pp. 536–547.

Raiber, Eun-Ang; Beraldi, Dario; Ficiz, Gabriella; Burgess, Heather E.; Branco, Miguel R.; Murat, Pierre et al. (2012): Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. In *Genome biology* 13 (8), R69.

Raisner, Ryan; Kharbanda, Samir; Jin, Lingyan; Jeng, Edwin; Chan, Emily; Merchant, Mark et al. (2018): Enhancer Activity Requires CBP/P300 Bromodomain-Dependent Histone H3K27 Acetylation. In *Cell reports* 24 (7), pp. 1722–1729.

Rajavelu, Arumugam; Jurkowska, Renata Z.; Fritz, Jürgen; Jeltsch, Albert (2012): Function and disruption of DNA methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation. In *Nucleic acids research* 40 (2), pp. 569–580.

Rajavelu, Arumugam; Lungu, Cristiana; Emperle, Max; Dukatz, Michael; Bröhm, Alexander; Broche, Julian et al. (2018): Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. In *Nucleic acids research* 46 (17), pp. 9044–9056.

Rangasamy, Sampathkumar; D'Mello, Santosh R.; Narayanan, Vinodh (2013): Epigenetics, autism spectrum, and neurodevelopmental disorders. In *Neurotherapeutics: the journal of the American Society for Experimental NeuroTherapeutics* 10 (4), pp. 742–756.

Rathert, Philipp; Roth, Mareike; Neumann, Tobias; Muerdter, Felix; Roe, Jae-Seok; Muhar, Matthias et al. (2015): Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. In *Nature* 525 (7570), pp. 543–547.

Ravichandran, M.; Jurkowska, R. Z.; Jurkowski, T. P. (2018): Target specificity of mammalian DNA methylation and demethylation machinery. In *Organic & biomolecular chemistry* 16 (9), pp. 1419–1435.

Rawłuszko-Wieczorek, Agnieszka Anna; Siera, Agnieszka; Horbacka, Karolina; Horst, Nikodem; Krokowicz, Piotr; Jagodziński, Paweł Piotr (2015): Clinical significance of DNA methylation mRNA levels of TET family members in colorectal cancer. In *Journal of cancer research and clinical oncology* 141 (8), pp. 1379–1392.

Ren, Wendan; Fan, Huitao; Grimm, Sara A.; Kim, Jae Jin; Li, Linhui; Guo, Yiran et al. (2021): DNMT1 reads heterochromatic H4K20me3 to reinforce LINE-1 DNA methylation. In *Nature communications* 12 (1), p. 2490.

Rice, Judd C.; Briggs, Scott D.; Ueberheide, Beatrix; Barber, Cynthia M.; Shabanowitz, Jeffrey; Hunt, Donald F. et al. (2003): Histone Methyltransferases Direct Different Degrees of Methylation to Define Distinct Chromatin Domains. In *Molecular cell* 12 (6), pp. 1591–1598.

Rivenbark, Ashley G.; Stolzenburg, Sabine; Beltran, Adriana S.; Yuan, Xinni; Rots, Marianne G.; Strahl, Brian D.; Blancafort, Pilar (2012): Epigenetic reprogramming of cancer cells via targeted DNA methylation. In *Epigenetics* 7 (4), pp. 350–360.

Rozenberg, Julian M.; Shlyakhtenko, Andrey; Glass, Kimberly; Rishi, Vikas; Myakishev, Maxim V.; FitzGerald, Peter C.; Vinson, Charles (2008): All and only CpG containing sequences are enriched in promoters abundantly bound by RNA polymerase II in multiple tissues. In *BMC genomics* 9, p. 67.

Sado, T.; Fenner, M. H.; Tan, S. S.; Tam, P.; Shioda, T.; Li, E. (2000): X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. In *Developmental biology* 225 (2), pp. 294–303.

Saksouk, Nehmé; Simboeck, Elisabeth; Déjardin, Jérôme (2015): Constitutive heterochromatin formation and transcription in mammals. In *Epigenetics & chromatin* 8, p. 3.

Sanchez-Delgado, Marta; Court, Franck; Vidal, Enrique; Medrano, Jose; Monteagudo-Sánchez, Ana; Martin-Trujillo, Alex et al. (2016): Human Oocyte-Derived Methylation Differences Persist in the Placenta Revealing Widespread Transient Imprinting. In *PLoS genetics* 12 (11), e1006427.

Sankaran, Saumya M.; Wilkinson, Alex W.; Elias, Joshua E.; Gozani, Or (2016): A PWWP Domain of Histone-Lysine N-Methyltransferase NSD2 Binds to Dimethylated Lys-36 of Histone H3 and Regulates NSD2 Function at Chromatin. In *The Journal of biological chemistry* 291 (16), pp. 8465–8474.

Santa, Francesca de; Barozzi, Iros; Mietton, Flore; Ghisletti, Serena; Polletti, Sara; Tusi, Betsa-beh Khoramian et al. (2010): A large fraction of extragenic RNA pol II transcription sites overlap enhancers. In *PLoS biology* 8 (5), e1000384.

Saunderson, Emily A.; Stepper, Peter; Gomm, Jennifer J.; Hoa, Lily; Morgan, Adrienne; Allen, Michael D. et al. (2017): Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. In *Nature communications* 8 (1), p. 1450.

Saxonov, Serge; Berg, Paul; Brutlag, Douglas L. (2006): A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. In *Proceedings of the National Academy of Sciences of the United States of America* 103 (5), pp. 1412–1417.

Schägger, Hermann (2006): Tricine-SDS-PAGE. In *Nature protocols* 1 (1), pp. 16–22.

Schambach, Axel; Galla, Melanie; Modlich, Ute; Will, Elke; Chandra, Saurabh; Reeves, Lilith et al. (2006): Lentiviral vectors pseudotyped with murine ecotropic envelope: increased biosafety and convenience in preclinical research. In *Experimental hematology* 34 (5), pp. 588–592.

Schlesinger, Yeshayahu; Straussman, Ravid; Keshet, Ilana; Farkash, Shlomit; Hecht, Merav; Zimmerman, Joseph et al. (2007): Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. In *Nature genetics* 39 (2), pp. 232–236.

Schuermann, David; Weber, Alain R.; Schär, Primo (2016): Active DNA demethylation by DNA repair: Facts and uncertainties. In *DNA repair* 44, pp. 92–102.

Schulz, Wolfgang A.; Lang, Alexander; Koch, Julian; Greife, Annemarie (2019): The histone de-methylase UTX/KDM6A in cancer: Progress and puzzles. In *International journal of cancer* 145 (3), pp. 614–620.

Sendžikaitė, Gintarė; Hanna, Courtney W.; Stewart-Morgan, Kathleen R.; Ivanova, Elena; Kelsey, Gavin (2019): A DNMT3A PWWP mutation leads to methylation of bivalent chromatin and growth retardation in mice. In *Nature communications* 10 (1), p. 1884.

Shao, Shipeng; Chang, Lei; Sun, Yuao; Hou, Yingping; Fan, Xiaoying; Sun, Yujie (2018): Multiplexed sgRNA Expression Allows Versatile Single Nonrepetitive DNA Labeling and Endogenous Gene Regulation. In *ACS synthetic biology* 7 (1), pp. 176–186.

Sharif, Jafar; Muto, Masahiro; Takebayashi, Shin-ichiro; Suetake, Isao; Iwamatsu, Akihiro; Endo, Takaho A. et al. (2007): The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. In *Nature* 450 (7171), pp. 908–912.

Sharifi-Zarchi, Ali; Gerovska, Daniela; Adachi, Kenjiro; Totonchi, Mehdi; Pezeshk, Hamid; Taft, Ryan J. et al. (2017): DNA methylation regulates discrimination of enhancers from promoters through a H3K4me1-H3K4me3 seesaw mechanism. In *BMC genomics* 18 (1), p. 964.

She, Xinwei; Rohl, Carol A.; Castle, John C.; Kulkarni, Amit V.; Johnson, Jason M.; Chen, Rong-hua (2009): Definition, conservation and epigenetics of housekeeping and tissue-enriched genes. In *BMC genomics* 10, p. 269.

- Shibata, Mikihiro; Nishimasu, Hiroshi; Kodera, Noriyuki; Hirano, Seiichi; Ando, Toshio; Uchihashi, Takayuki; Nureki, Osamu (2017): Real-space and real-time dynamics of CRISPR-Cas9 visualized by high-speed atomic force microscopy. In *Nature communications* 8 (1), p. 1430.
- Shilatifard, Ali (2012): The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. In *Annual review of biochemistry* 81, pp. 65–95.
- Shimbo, Takashi; Du, Ying; Grimm, Sara A.; Dhasarathy, Archana; Mav, Deepak; Shah, Ruchir R. et al. (2013): MBD3 localizes at promoters, gene bodies and enhancers of active genes. In *PLoS genetics* 9 (12), e1004028.
- Siddique, Abu Nasar; Nunna, Suneetha; Rajavelu, Arumugam; Zhang, Yingying; Jurkowska, Renata Z.; Reinhardt, Richard et al. (2013): Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. In *Journal of molecular biology* 425 (3), pp. 479–491.
- Simmen, Martin W. (2008): Genome-scale relationships between cytosine methylation and dinucleotide abundances in animals. In *Genomics* 92 (1), pp. 33–40.
- Smiraglia, D. J.; Rush, L. J.; Frühwald, M. C.; Dai, Z.; Held, W. A.; Costello, J. F. et al. (2001): Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. In *Human molecular genetics* 10 (13), pp. 1413–1419.
- Smith, Zachary D.; Chan, Michelle M.; Humm, Kathryn C.; Karnik, Rahul; Mekhoubad, Shila; Regev, Aviv et al. (2014): DNA methylation dynamics of the human preimplantation embryo. In *Nature* 511 (7511), pp. 611–615.
- Spruijt, Cornelia G.; Gnerlich, Felix; Smits, Arne H.; Pfaffeneder, Toni; Jansen, Pascal W. T. C.; Bauer, Christina et al. (2013): Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. In *Cell* 152 (5), pp. 1146–1159.
- Stasevich, Timothy J.; Hayashi-Takanaka, Yoko; Sato, Yuko; Maehara, Kazumitsu; Ohkawa, Yasuyuki; Sakata-Sogawa, Kumiko et al. (2014): Regulation of RNA polymerase II activation by histone acetylation in single living cells. In *Nature* 516 (7530), pp. 272–275.
- Steenman, M. J.; Rainier, S.; Dobry, C. J.; Grundy, P.; Horon, I. L.; Feinberg, A. P. (1994): Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. In *Nature genetics* 7 (3), pp. 433–439.
- Stepanenko, A. A.; Dmitrenko, V. V. (2015): HEK293 in cell biology and cancer research: phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. In *Gene* 569 (2), pp. 182–190.
- Stepanenko, Aleksei A.; Heng, Henry H. (2017): Transient and stable vector transfection: Pitfalls, off-target effects, artifacts. In *Mutation research* 773, pp. 91–103.
- Stepper, Peter; Kungulovski, Goran; Jurkowska, Renata Z.; Chandra, Tamir; Krueger, Felix; Reinhardt, Richard et al. (2017): Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. In *Nucleic acids research* 45 (4), pp. 1703–1713.
- Stolzenburg, S.; Beltran, A. S.; Swift-Scanlan, T.; Rivenbark, A. G.; Rashwan, R.; Blancafort, P. (2015): Stable oncogenic silencing in vivo by programmable and targeted de novo DNA methylation in breast cancer. In *Oncogene* 34 (43), pp. 5427–5435.

Strahl, B. D.; Allis, C. D. (2000): The language of covalent histone modifications. In *Nature* 403 (6765), pp. 41–45.

Strahl, B. D.; Ohba, R.; Cook, R. G.; Allis, C. D. (1999): Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. In *Proceedings of the National Academy of Sciences of the United States of America* 96 (26), pp. 14967–14972.

Su, Jianzhong; Shao, Xiujuan; Liu, Hongbo; Liu, Shengqiang; Wu, Qiong; Zhang, Yan (2012): Genome-wide dynamic changes of DNA methylation of repetitive elements in human embryonic stem cells and fetal fibroblasts. In *Genomics* 99 (1), pp. 10–17.

Sudhamalla, Babu; Dey, Debasis; Breski, Megan; Islam, Kabirul (2017): A rapid mass spectrometric method for the measurement of catalytic activity of ten-eleven translocation enzymes. In *Analytical biochemistry* 534, pp. 28–35.

Sudhamalla, Babu; Wang, Sinan; Snyder, Valerie; Kavooosi, Sam; Arora, Simran; Islam, Kabirul (2018): Complementary Steric Engineering at the Protein-Ligand Interface for Analogue-Sensitive TET Oxygenases. In *Journal of the American Chemical Society* 140 (32), pp. 10263–10269.

Suetake, Isao; Mishima, Yuichi; Kimura, Hironobu; Lee, Young-Ho; Goto, Yuji; Takeshima, Hideyuki et al. (2011): Characterization of DNA-binding activity in the N-terminal domain of the DNA methyltransferase Dnmt3a. In *The Biochemical journal* 437 (1), pp. 141–148.

Suzuki, M.; Yamada, T.; Kihara-Negishi, F.; Sakurai, T.; Hara, E.; Tenen, D. G. et al. (2006): Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. In *Oncogene* 25 (17), pp. 2477–2488.

Suzuki, Mitsuhiro; Yamada, Toshiyuki; Kihara-Negishi, Fumiko; Sakurai, Takuya; Oikawa, Tsuneyuki (2003): Direct association between PU.1 and MeCP2 that recruits mSin3A-HDAC complex for PU.1-mediated transcriptional repression. In *Oncogene* 22 (54), pp. 8688–8698.

Szerlong, Heather J.; Hansen, Jeffrey C. (2011): Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. In *Biochemistry and cell biology = Biochimie et biologie cellulaire* 89 (1), pp. 24–34.

Tahiliani, Mamta; Koh, Kian Peng; Shen, Yinghua; Pastor, William A.; Bandukwala, Hozefa; Brudno, Yevgeny et al. (2009): Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. In *Science (New York, N.Y.)* 324 (5929), pp. 930–935.

Takai, Daiya; Jones, Peter A. (2002): Comprehensive analysis of CpG islands in human chromosomes 21 and 22. In *Proceedings of the National Academy of Sciences of the United States of America* 99 (6), pp. 3740–3745.

Takeshima, Hideyuki; Wakabayashi, Mika; Hattori, Naoko; Yamashita, Satoshi; Ushijima, Toshi-kazu (2015): Identification of coexistence of DNA methylation and H3K27me3 specifically in cancer cells as a promising target for epigenetic therapy. In *Carcinogenesis* 36 (2), pp. 192–201.

Tanenbaum, Marvin E.; Gilbert, Luke A.; Qi, Lei S.; Weissman, Jonathan S.; Vale, Ronald D. (2014): A protein-tagging system for signal amplification in gene expression and fluorescence imaging. In *Cell* 159 (3), pp. 635–646.

- Taverna, Sean D.; Li, Haitao; Ruthenburg, Alexander J.; Allis, C. David; Patel, Dinshaw J. (2007): How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. In *Nature structural & molecular biology* 14 (11), pp. 1025–1040.
- Thurman, Robert E.; Rynes, Eric; Humbert, Richard; Vierstra, Jeff; Maurano, Matthew T.; Haugen, Eric et al. (2012): The accessible chromatin landscape of the human genome. In *Nature* 489 (7414), pp. 75–82.
- Toyota, M.; Ahuja, N.; Ohe-Toyota, M.; Herman, J. G.; Baylin, S. B.; Issa, J. P. (1999): CpG island methylator phenotype in colorectal cancer. In *Proceedings of the National Academy of Sciences of the United States of America* 96 (15), pp. 8681–8686.
- Trojer, Patrick; Reinberg, Danny (2007): Facultative heterochromatin: is there a distinctive molecular signature? In *Molecular cell* 28 (1), pp. 1–13.
- Tsagaratou, Ageliki; Äijö, Tarmo; Lio, Chan-Wang J.; Yue, Xiaojing; Huang, Yun; Jacobsen, Steven E. et al. (2014): Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. In *Proceedings of the National Academy of Sciences of the United States of America* 111 (32), E3306-15.
- Tucci, Valter; Isles, Anthony R.; Kelsey, Gavin; Ferguson-Smith, Anne C. (2019): Genomic Imprinting and Physiological Processes in Mammals. In *Cell* 176 (5), pp. 952–965.
- Ueda, Minoru; Seki, Motoaki (2020): Histone Modifications Form Epigenetic Regulatory Networks to Regulate Abiotic Stress Response. In *Plant physiology* 182 (1), pp. 15–26.
- Uyehara, Christopher M.; Nystrom, Spencer L.; Niederhuber, Matthew J.; Leatham-Jensen, Mary; Ma, Yiqin; Buttitta, Laura A.; McKay, Daniel J. (2017): Hormone-dependent control of developmental timing through regulation of chromatin accessibility. In *Genes & development* 31 (9), pp. 862–875.
- van den Ent, Fusinita; Löwe, Jan (2006): RF cloning: a restriction-free method for inserting target genes into plasmids. In *Journal of biochemical and biophysical methods* 67 (1), pp. 67–74.
- Veland, Nicolas; Chen, Taiping (2017): Mechanisms of DNA Methylation and Demethylation During Mammalian Development. In : *Handbook of Epigenetics*, vol. 9: Elsevier, pp. 11–24.
- Verdone, Loredana; Agricola, Eleonora; Caserta, Micaela; Di Mauro, Ernesto (2006): Histone acetylation in gene regulation. In *Briefings in functional genomics & proteomics* 5 (3), pp. 209–221.
- Viré, Emmanuelle; Brenner, Carmen; Deplus, Rachel; Blanchon, Loïc; Fraga, Mario; Didelot, Céline et al. (2006): The Polycomb group protein EZH2 directly controls DNA methylation. In *Nature* 439 (7078), pp. 871–874.
- Vojta, Aleksandar; Dobrinić, Paula; Tadić, Vanja; Bočkor, Luka; Korać, Petra; Julg, Boris et al. (2016): Repurposing the CRISPR-Cas9 system for targeted DNA methylation. In *Nucleic acids research* 44 (12), pp. 5615–5628.
- Waddington, C. H. (1942): CANALIZATION OF DEVELOPMENT AND THE INHERITANCE OF ACQUIRED CHARACTERS. In *Nature* 150 (3811), pp. 563–565.

- Waddington, C. H. (2012): The epigenotype. 1942. In *International journal of epidemiology* 41 (1), pp. 10–13.
- Walsh, C. P.; Chaillet, J. R.; Bestor, T. H. (1998): Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. In *Nature genetics* 20 (2), pp. 116–117.
- Wang, Hao; Maurano, Matthew T.; Qu, Hongzhu; Varley, Katherine E.; Gertz, Jason; Pauli, Flor-encia et al. (2012): Widespread plasticity in CTCF occupancy linked to DNA methylation. In *Ge-nome research* 22 (9), pp. 1680–1688.
- Wang, Xueyin; Paucek, Richard D.; Gooding, Anne R.; Brown, Zachary Z.; Ge, Eva J.; Muir, Tom W.; Cech, Thomas R. (2017): Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. In *Nature structural & molecular biology* 24 (12), pp. 1028–1038.
- Wang, Zhibin; Zang, Chongzhi; Rosenfeld, Jeffrey A.; Schones, Dustin E.; Barski, Artem; Cuddapah, Suresh et al. (2008): Combinatorial patterns of histone acetylations and methylations in the human genome. In *Nature genetics* 40 (7), pp. 897–903.
- Weber, Michael; Hellmann, Ines; Stadler, Michael B.; Ramos, Liliana; Pääbo, Svante; Rebhan, Michael; Schübeler, Dirk (2007): Distribution, silencing potential and evolutionary impact of pro-moter DNA methylation in the human genome. In *Nature genetics* 39 (4), pp. 457–466.
- Weinberg, Daniel N.; Papillon-Cavanagh, Simon; Chen, Haifen; Yue, Yuan; Chen, Xiao; Rajagopalan, Kartik N. et al. (2019): The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. In *Nature* 573 (7773), pp. 281–286.
- Weinberg, Daniel N.; Rosenbaum, Phillip; Chen, Xiao; Barrows, Douglas; Horth, Cynthia; Marun-de, Matthew R. et al. (2021): Two competing mechanisms of DNMT3A recruitment regulate the dynamics of de novo DNA methylation at PRC1-targeted CpG islands. In *Nature genetics* 53 (6), pp. 794–800.
- Weinert, Brian T.; Narita, Takeo; Satpathy, Shankha; Srinivasan, Balaji; Hansen, Bogi K.; Schölz, Christian et al. (2018): Time-Resolved Analysis Reveals Rapid Dynamics and Broad Scope of the CBP/p300 Acetylome. In *Cell* 174 (1), 231-244.e12.
- Weinhold, Bob (2006): Epigenetics: the science of change. In *Environmental health perspectives* 114 (3), A160-7.
- Whyte, Warren A.; Bilodeau, Steve; Orlando, David A.; Hoke, Heather A.; Frampton, Garrett M.; Foster, Charles T. et al. (2012): Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. In *Nature* 482 (7384), pp. 221–225.
- Widschwendter, Martin; Fiegl, Heidi; Egle, Daniel; Mueller-Holzner, Elisabeth; Spizzo, Gilbert; Marth, Christian et al. (2007): Epigenetic stem cell signature in cancer. In *Nature genetics* 39 (2), pp. 157–158.
- Wilhelm, Brian T.; Marguerat, Samuel; Aligianni, Sofia; Codlin, Sandra; Watt, Stephen; Bähler, Jürg (2011): Differential patterns of intronic and exonic DNA regions with respect to RNA polymerase II occupancy, nucleosome density and H3K36me3 marking in fission yeast. In *Genome biology* 12 (8), R82.
- Williams, Kristine; Christensen, Jesper; Pedersen, Marianne Terndrup; Johansen, Jens V.; Cloos, Paul A. C.; Rappsilber, Juri; Helin, Kristian (2011): TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. In *Nature* 473 (7347), pp. 343–348.

- Wolfe, S. A.; Neklodova, L.; Pabo, C. O. (2000): DNA recognition by Cys2His2 zinc finger proteins. In *Annual review of biophysics and biomolecular structure* 29, pp. 183–212.
- Wu, Ct; Morris, J. R. (2001): Genes, genetics, and epigenetics: a correspondence. In *Science* (New York, N.Y.) 293 (5532), pp. 1103–1105.
- Wu, Rong; Terry, Anna V.; Singh, Prim B.; Gilbert, David M. (2005): Differential subnuclear localization and replication timing of histone H3 lysine 9 methylation states. In *Molecular biology of the cell* 16 (6), pp. 2872–2881.
- Wu, Tao P.; Wang, Tao; Seetin, Matthew G.; Lai, Yongquan; Zhu, Shijia; Lin, Kaixuan et al. (2016): DNA methylation on N(6)-adenine in mammalian embryonic stem cells. In *Nature* 532 (7599), pp. 329–333.
- Xiao, Chuan-Le; Zhu, Song; He, Minghui; Chen, De; Zhang, Qian; Chen, Ying et al. (2018): N6-Methyladenine DNA Modification in the Human Genome. In *Molecular cell* 71 (2), 306–318.e7.
- Xie, Shaoping; Wang, Zhenjuan; Okano, Masaki; Nogami, Masahiro; Li, Yuan; He, Wei-Wu et al. (1999): Cloning, expression and chromosome locations of the human DNMT3 gene family. In *Gene* 236 (1), pp. 87–95.
- Xie, Wei; Schultz, Matthew D.; Lister, Ryan; Hou, Zhonggang; Rajagopal, Nisha; Ray, Pradipta et al. (2013): Epigenomic analysis of multilineage differentiation of human embryonic stem cells. In *Cell* 153 (5), pp. 1134–1148.
- Xiong, Tina; Rohm, Dahlia; Workman, Rachael E.; Roundtree, Lauren; Novina, Carl D.; Timp, Winston; Ostermeier, Marc (2018): Protein engineering strategies for improving the selective methylation of target CpG sites by a dCas9-directed cytosine methyltransferase in bacteria. In *PloS one* 13 (12), e0209408.
- Yildirim, Ozlem; Li, Ruowang; Hung, Jui-Hung; Chen, Poshen B.; Dong, Xianjun; Ee, Ly-Sha et al. (2011): Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. In *Cell* 147 (7), pp. 1498–1510.
- Yin, Yimeng; Morgunova, Ekaterina; Jolma, Arttu; Kaasinen, Eevi; Sahu, Biswajyoti; Khund-Sayeed, Syed et al. (2017): Impact of cytosine methylation on DNA binding specificities of human transcription factors. In *Science* (New York, N.Y.) 356 (6337).
- Zemach, Assaf; Zilberman, Daniel (2010): Evolution of eukaryotic DNA methylation and the pursuit of safer sex. In *Current biology : CB* 20 (17), R780-5.
- Zeng, Yang; Chen, Taiping (2019): DNA Methylation Reprogramming during Mammalian Development. In *Genes* 10 (4).
- Zhang, Lijian; Li, Zhaojin; Chen, Guangzhao; Huang, Qionglin; Zhang, Junjie; Wen, Juan et al. (2016): Validation and quantification of genomic 5-carboxylcytosine (5caC) in mouse brain tissue by liquid chromatography-tandem mass spectrometry. In *Anal. Methods* 8 (29), pp. 5812–5817.
- Zhang, Tianyi; Cooper, Sarah; Brockdorff, Neil (2015): The interplay of histone modifications - writers that read. In *EMBO reports* 16 (11), pp. 1467–1481.

Zhang, Xiaotian; Jeong, Mira; Huang, Xingfan; Wang, Xue Qing; Wang, Xinyu; Zhou, Wanding et al. (2020): Large DNA Methylation Nadirs Anchor Chromatin Loops Maintaining Hematopoietic Stem Cell Identity. In *Molecular cell* 78 (3), 506-521.e6.

Zhang, Yingying; Rohde, Christian; Tierling, Sascha; Jurkowski, Tomasz P.; Bock, Christoph; Santacruz, Diana et al. (2009): DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. In *PLoS genetics* 5 (3), e1000438.

Zhang, Zhi-Min; Lu, Rui; Wang, Pengcheng; Yu, Yang; Chen, Dongliang; Gao, Linfeng et al. (2018): Structural basis for DNMT3A-mediated de novo DNA methylation. In *Nature* 554 (7692), pp. 387–391.

Zheng, Xiaoguo; Li, Zhenhua; Wang, Guishuan; Li, Zhengzheng; Liang, Ajuan; Wang, Hanshu et al. (2017): Overexpression of Human-Derived DNMT3A Induced Intergenerational Inheritance of Active DNA Methylation Changes in Rat Sperm. In *Frontiers in genetics* 8, p. 207.

Zhu, Yueyan; Sun, Daijing; Jakovcevski, Mira; Jiang, Yan (2020): Epigenetic mechanism of SETDB1 in brain: implications for neuropsychiatric disorders. In *Translational psychiatry* 10 (1), p. 115.

Author's contributions

Research articles:

- Lungu, C., Pinter, S., Broche, J., Rathert, P. & Albert Jeltsch (2017). Modular fluorescence complementation sensors for live cell detection of epigenetic signals at endogenous genomic sites. *Nature Communications*, 8(1), 649.

C.L. and A.J. designed the research. C.L. and S.P. performed the experimental work. P.R. provided advice and materials for the generation of the dox-inducible cell lines. J.B. performed the bioinformatic search for TALE binding sites. C.L. and A.J. wrote the manuscript draft. All authors were involved in data analysis and finalizing of the manuscript.

- Rajavelu, A., Lungu, C., Emperle, M., Dukatz, M., Bröhm, A., **Broche, J.**, Hanelt, I., Parsa, E., Schiffers, S., Karnik R., Meissner, A., Carell, T., Rathert, P., Jurkowska, R. Z. & Jeltsch, A. (2018). Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. *Nucleic Acids Research*, 6(17), 9044-9056.

A.J., A.R., C.L. and R.Z.J. devised the project and analyzed the data. C.L. and A.R. conducted the biochemical assays with contributions from I.H., M.D., A.B. and M.E. C.L. and A.R. performed the fluorescence microscopy experiments. C.L. and M.E. performed the cell culture and biochemical work for the data shown in Figure 55 with contribution from P.R. and J.B. S.S., E.P. and T.C. performed the LC-ESI-MS/MS. All authors contributed to data interpretation and discussion, read and approved the final manuscript.

- Hofacker, D.*, Broche, J.*, Laistner L., Adam, S., Bashtrykov, P. & Jeltsch, A. (2020). Engineering of Effector Domains for Targeted DNA Methylation with Reduced Off-Target Effects. *International Journal of Molecular Sciences*, 21(2), 502.

*co-shared first authors

Conceptualization, P.B. and A.J.; development of methodology, J.B. and P.B.; experimental studies and targeted bis-seq studies, D.H., with help by L.L., S.A., and P.B.; MBD-seq studies, J.B.; original draft preparation, P.B. and A.J.; review and editing of the manuscript, all authors; visualization, D.H., P.B., and J.B.; supervision, P.B. and A.J.; funding acquisition, A.J. All authors have read and agreed to the published version of the manuscript.

- Broche, J., Kungulovski, G., Bashtrykov, P., Rathert P. & Jeltsch, A. (2021). Genome-wide investigation of the dynamic changes of epigenome modifications after global DNA methylation editing. *Nucleic Acids Research*, 49(1), 158-176.

J.B. conducted all experiments with support from G.K., P.B. and P.R. A.J. and G.K. conceived and devised the study. A.J., G.K., P.R. and P.B. supervised the work. J.B. and A.J. prepared the manuscript draft and figures. All authors contributed to data interpretation and editing of the manuscript. The final manuscript was approved by all authors.

Reviews:

- Jeltsch, A., Broche, J. & Pavel Bashtrykov (2018). Molecular Processes Connecting DNA Methylation Patterns with DNA Methyltransferases and Histone Modifications in Mammalian Genomes. *Genes (Basel)*, 9(11), 566.

JB contributed to the writing of the manuscript.

- Jeltsch, A., Broche, J., Lungu, C. & Pavel Bashtrykov (2019). Biotechnological Applications of MBD Domain Proteins for DNA Methylation Analysis. *Journal of Molecular Biology*, 432(6), 1816-1823.

JB contributed to the writing of the manuscript.

Appendix

(not included in the published version)

Appendix 1:

Broche, J., Kungulovski, G., Bashtrykov, P., Rathert P. & Jeltsch, A. (2021). Genome-wide investigation of the dynamic changes of epigenome modifications after global DNA methylation editing. *Nucleic Acids Research*, 49(1), 158-176.

Appendix 2:

Hofacker, D.*, Broche, J.*, Laistner L., Adam, S., Bashtrykov, P. & Jeltsch, A. (2020). Engineering of Effector Domains for Targeted DNA Methylation with Reduced Off-Target Effects. *International Journal of Molecular Sciences*, 21(2), 502.

*co-shared first authors