CRISPR-Cas9 fusions for synthetic epigenetics

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

Peter Stepper

aus Filderstadt

Hauptberichter: Jun. Prof. Dr. Tomasz Jurkowski

Mitberichter: Prof. Dr. Jörn Lausen

Mitberichterin: Dr. Melita Vidaković

Tag der mündlichen Prüfung: 10.08.2020

Institut für Biochemie und Technische Biochemie
der Universität Stuttgart

2020

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbst verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Wien, August 2020

Peter Stepper

Contents

Acknowledgments

This thesis would not have been possible without the great work and support of many people.

First of all, I want to thank my supervisor Dr. Tomasz Jurkowski for his enthusiasm, endless terrific ideas, guidance and support. He gave me the chance to work on the exciting new topic of CRISPR-Cas9 based epigenome editing back when nobody even knew whether it would work.

I'm thankful to Dr. Melita Vidaković and Prof. Dr. Jörn Lausen for taking the time to examine my thesis and be co-referees in my committee; and Prof. Dr. Markus Morrison for chairing my committee. Furthermore, I'd like to thank Prof. Dr. Olayioye, Prof. Dr. Wege and Prof. Dr. Kontermann for agreeing to review my thesis.

I'm grateful to the groups of Dr. Gabriella Ficz, Dr. Melita Vidaković and Dr. Tim Hore for the pleasure of collaborating with them on exciting projects and their outstanding work.

My sincere gratitude goes to all my colleagues who trained me, helped me, and made life in the lab fun most of the time. I'd like to especially thank Miruna and Phil for everything they helped me with and the great time we shared, and of course Anja, Micha, Rebekka, Johannes, Denis, Nicole, Goran and Sara and everyone else. Furthermore, I am thankful to Dr. Renata Jurkowska for her infectious enthusiasm and helpful discussions.

I want to thank all my students for the privilege of mentoring them, for everything they taught me and all their help with my projects. Liv, Roman, Sascha, Jule and Vivien, I really enjoyed working and spending time with you! The same goes for all the interns and guests who were working in our lab.

Furthermore, I'd like to thank Dr. Stephan Eisler and Melanie Noack for their FACS expertise and help.

Last but definitely not least, I'll be forever grateful to my family and especially my parents, who always supported me in every way and played a large part in shaping who I am today. The same goes for my non-scientist friends, also for taking my mind off science every now and then and enjoying the other great things in life with me.

List of abbreviations

5caC 5-carboxylcytosine

5fC 5-formylcytosine

5hmC 5-hydroxymethylcytosine

5mC 5-methylcytosine

ACTB actin beta

ADD ATRX-DNMT3-DNMT3L

Arx aristaless- related homeobox

AS Angelman syndrome

ASCL1 achaete-scute family bHLH transcription factor 1

BER base excision repair

BOMB Bio-On-Magnetic-Beads

bp base pair

BRCA breast cancer

BRCA1 BRCA1 DNA repair associated

C2H2 Cys2-His2

CDKN2A cyclin dependent kinase inhibitor 2A

cDNA complementary DNA

CGI CpG island

ChIP chromatin immunoprecipitation

CpG Cytosin phosphate Guanine

Cq Quantification cycle, the cycle at which the fluorescence threshold is crossed in qPCR

CRISPR clustered regularly interspaced short palindromic repeats

crRNA CRISPR RNA

CXCR4 C-X-C motif chemokine receptor 4

CXXC4 CXXC Finger Protein 4

dCas9 dead Cas9

DMEM Dulbecco's Modified Eagle's Medium

DMOG Dimethyloxalylglycine

DMSO Dimethyl sulfoxide

DNMT DNA Methyltransferase

DOT1 disruptor of telomeric silencing 1

DOT1L DOT1 like histone lysine methyltransferase

DPBS Dulbecco's Phosphate Buffered Saline

dpt days past transfection

DSB double-strand break

ED effector domain

EDTA ethylenediaminetetraacetic acid

EPCAM epithelial cell adhesion molecule

ESC embryonic stem cell

EZH2 enhancer of zeste 2 polycomb repressive complex 2 subunit

FACS fluorescence-activated cell sorting

FC flow cytometry

FCS fetal calf serum

GITC guanidine thiocyanate

HDAC histone deacetylase

HEK293 human embryonic kidney 293

HKMT histone lysine methyltransferase

IL6ST interleukin 6 signal transducer

kb kilo base, thousand DNA bases

KDM2B lysine demethylase 2B

KMT2A lysine methyltransferase 2A

KO knockout

KRAB Krüppel associated box

LIN28A lin-28 homolog A

MAA methacrylic acid

MACS Magnetic-Activated Cell Sorting

MAGEB2 MAGE (melanoma antigen) family member B2

MECP2 methyl-CpG binding protein 2

miRNA micro RNA

MLL myeloid/lymphoid or mixed-lineage leukemia

MNP magnetic nano-particle

NaOH sodium hydroxide

PAM protospacer adjacent motif

PCNA proliferating cell nuclear antigen

PHD plant homeodomain

pp percentage point

PRC2 polycomb repressive complex 2

PTM post-translational modification

PWS Prader–Willi syndrome

PWWP Pro-Trp-Trp-Pro

RB retinoblastoma

RB1 RB transcriptional corepressor 1

REEP5 receptor accessory protein 5

RING really interesting new gene

RT-qPCR reverse-transcription quantitative PCR

RTT Rett syndrome

SAM S-adenosyl L-methionine

SFM scanning force microscopy

sgRNA single guide RNA

shRNA short hairpin RNA

siRNAs small interfering RNA

SIRT1 sirtuin 1

SPRI solid-phase reversible immobilization

SRA SET and RING-associated

TAD topologically associating domain

TALE Transcription-Activator-Like Effector

TDG thymine DNA glycosylase

TEOS tetraethylorthosilicate

TET ten-eleven translocation

TFRC transferrin receptor protein

tracrRNA trans-activating RNA

UBL ubiquitin-like

UHRF1 ubiquitin like with PHD and ring finger domains 1. Also known as Np95

VP64 four copies of viral protein 16 from herpes simplex

VPR VP64-p65-Rta

WDR5 WD repeat domain 5

WT wild-type

αKG α-ketoglutarate

Zusammenfassung

Obwohl fast alle Zellen im menschlichen Körper dieselbe DNA-Sequenz besitzen, variieren die Phänotypen und Funktionen der über hundert verschiedenen Zelltypen beträchtlich. Für diese Unterschiede sind spezifische Genaktivitätsmuster verantwortlich, die für jeden Zelltyp einzigartig sind. Diese Transkriptionsmuster werden durch epigenetische Mechanismen reguliert. Nicht-kodierende RNAs, post-translationale Histonmodifikationen und chemische Markierungen an den DNA-Basen bestimmen die Genexpressionsstärken und die Zellidentität. Allerdings sind trotz des enormen Fortschritts bei der Beschreibung der Korrelation zwischen den spezifischen Markierungen und der Genexpression viele Fragen zum epigenetischen Netzwerk nach wie vor offen. Eine vielversprechende Methode, um den epigenetischen Code zu verstehen, ist die gezielte Störung des epigenetischen Netzwerks und die Analyse dessen Antwort. Diese Methode der Epigenomeditierung verwendet eine programmierbare DNA-Bindedomäne und epigenetische Effektordomänen, um an spezifischen Stellen im Genom den lokalen Chromatinzustand und die Genexpression zu verändern. Das revolutionäre CRISPR-Cas9 System, mit einem katalytisch inaktiven dCas9-Protein als genaues, schnelles und praktikables Werkzeug für die gezielte DNA-Bindung, hat dem noch jungen Feld der Epigenomeditierung enorme Sprünge ermöglicht.

In dieser Arbeit nutzte ich dCas9 mit mehreren Epigenommodifizierern zur gezielten DNA-Methylierung und Genrepression, sowie zur gezielten DNA-Demethylierung und Aktivierung der Genexpression in menschlichen Zelllinien. Ich zeigte, dass die gezielte Rekrutierung des Fusionsproteins Dnmt3a3L zu DNA-Methylierung führte, die den gesamten Zielpromoter bedeckte, und dass diese starke Ausbreitung durch die Multimerisierung der DNA-Methyltransferasen entlang der DNA verursacht wurde. Allerdings ging die eingebrachte DNA-Methylierung innerhalb weniger Tage vollständig zurück, wie es auch in anderen Studien beobachtet wurde. Ich lege dar, dass ein vollständigeres Umschalten des epigenetischen Zustands der Zielgene durch die kombinierten Effekte von Dnmt3a3L und KRAB als Fusionsprotein die Stärke und Stabilität der repremierten Genexpression erhöhte. Diese Effekte wurden durch die zusätzliche Behandlung der editierten Zellen mit einem MLL1-Inhibitor verstärkt. Allerdings wurde die Verlustrate der DNA-Methylierung nicht reduziert, wenn die TET-Enzyme durch einen Dioxygenaseinhibitor inhibiert wurden. Das deutet darauf hin, dass die passive Demethylierung durch die beeinträchtigte Aufrechterhaltung der DNA-Methylierung die Ursache war, und nicht die aktive Oxidierung durch TET-Enzyme.

Außerdem konnte ich gezielte, stabile Genaktivierung für mindestens 46 Tage durch DNA-Demethylierung mit veränderten, hyperaktiven katalytischen Domänen des Tet3-Enzyms

aus Mäusen erreichen. Die Aktivierung der Zielgene konnte durch eine Fusion von Tet3 an den Transkriptionsaktivierer VPR synergistisch verstärkt werden. Interessanterweise unterschieden sich die Effekte zwischen den Zielgenen, was die Bedeutung des individiuellen epigenetischen Status des Zielgens hervorhebt.

Ich wendete diese Erkenntnisse in Kooperationen an, um Folgendes zu zeigen: Erstens, dass die Hypermethylierung des Tumorsupressors p16 ausreichte, um den Eintritt in die Seneszenz bei gesunden Brustgewebezellen zu verhindern, was einen ersten Schritt der Tumorinitiierung darstellt. Wir erreichten zweitens die Induktion der Insulinproduktion in pankreatischen α-Zellen aus Mäusen durch die gezielte DNA-Methylierung des Promoters des Hauptreglers *Arx*.

Außerdem haben wir kostengünstige Methoden für die auf magnetischen Kugeln basierenden Hochdurchsatzisolierung von RNA und DNA aus kultivierten Säugetier- und Bakterienzellen, sowie für die Bisulfitkonvertierung entwickelt. Die Methoden zur Isolierung und Bearbeitung von Nukleinsäuren, die Protokolle für die Herstellung der magnetischen Kugeln und die selbst hergestellten magnetischen Probenhalterungen ergeben zusammen die quelloffene Plattform Bio-On-Magnetic-Beads (BOMB).

Abstract

Although almost all cells in a human body contain the same DNA sequence, the phenotypes and functions of the hundreds of different cell types vary considerably. Responsible for these differences are specific patterns of gene activity, which differ for each cell type. These transcriptional patterns are regulated by epigenetic mechanisms. Non-coding RNAs, post-translational modifications of histones and chemical marks on the DNA bases determine gene expression strength and cell identity. However, despite the enormous progress describing the correlation between specific marks and gene expression, many questions about the epigenetic network remain unanswered. One promising method to discern the epigenetic code is the controlled perturbation of the epigenetic network and the analysis of its response. This method, epigenome editing, relies on a programmable DNA binding domain and epigenetic effector domains to change the local chromatin state and gene expression at specific loci. The revolutionary CRISPR-Cas9 system with a catalytically dead dCas9 as a precise, fast and convenient targeting device has led to enormous leaps for the young field of epigenome editing.

In this work I used dCas9 with a toolbox of epigenome modifiers for both targeted DNA methylation and gene repression as well as targeted DNA demethylation and activation of gene expression in human cell lines. I showed that targeting the single-chain fusion protein Dnmt3a3L lead to DNA methylation spreading across the entire targeted promoter region and that the widespread methylation was caused by the multimerization of DNA methyltransferases along the DNA. However, the introduced DNA methylation was lost quickly after a few days, as observed in other studies as well. I demonstrated that a more complete switch of the epigenetic state of the target genes by combining the effects of Dnmt3a3L and KRAB as a single fusion protein increased the strength and stability of the repressed expression state. These effects were enhanced by additionally treating the edited cells with an MLL1 inhibitor. However, inhibiting TET enzymes with a dioxygenase inhibitor did not reduce the rate of loss of targeted DNA methylation. This hints at passive DNA demethylation through cell division and impaired DNA methylation maintenance as the cause, rather than active oxidation via TET.

Furthermore, I was able to demonstrate stable gene activation via targeted DNA demethylation using engineered, hyperactive catalytic domains of murine Tet3, lasting at least 46 days. Target gene activation could be boosted synergistically with fusions of Tet3 to the transcriptional activator VPR. Interestingly, the effects differed by target gene, highlighting the importance of the individual epigenetic state of the targeted gene.

I applied these findings in collaborations to show, firstly, that hypermethylation of the

tumor suppressor p16 was enough to suppress senescence entry in healthy breast cells as a first step of tumor initiation. Secondly, we achieved the induction of insulin production in murine pancreatic α-cells via targeted DNA methylation of the promoter of the master regulator *Arx*.

Furthermore, we developed cost-effective high-throughput methods for the isolation of RNA and DNA from cultured mammalian and bacterial cells and bisulfite conversion based on home-made functionalized magnetic beads. The nucleic acid isolation and handling methods, the magnetic bead preparation protocols and self-made magnetic racks form the open source platform Bio-On-Magnetic-Beads (BOMB).

1 Introduction

1.1 Epigenetics

The term epigenetics was coined by Conrad Waddington. His metaphor of the epigenetic landscape in figure 1.1A (Waddington, 1957) is still influential today (Allen, 2015). The picture of a marble at the top of a hill with many branching valleys depicts the cell differentiation process during which naive cells gradually change from totipotency to one of the final differentiated cell types at the bottom of the hill. Another possible visualization of this process would be a funnel, restricting the differentiation options of the differentiation cells as it travels downwards from the undifferentiated state to the final cell type (Roy and Kundu, 2014). Waddington's lesser known depiction of the epigenetic landscape in figure 1.1B illustrates how this epigenetic landscape is formed: by the genes acting as pegs with lines molding the mountainside into a particular form to shape the possible differentiation pathways.

However, Waddington published his theories around the time when the structure of DNA was just being resolved. Since then, our understanding of the molecular mechanisms of genetics and epigenetics has improved dramatically (Goldberg, Allis, and Bernstein, 2007). The current definition of an epigenetic trait with much clearer knowledge of DNA and genes was agreed upon in a Cold Spring Harbor meeting as a "stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" (Berger et al.,

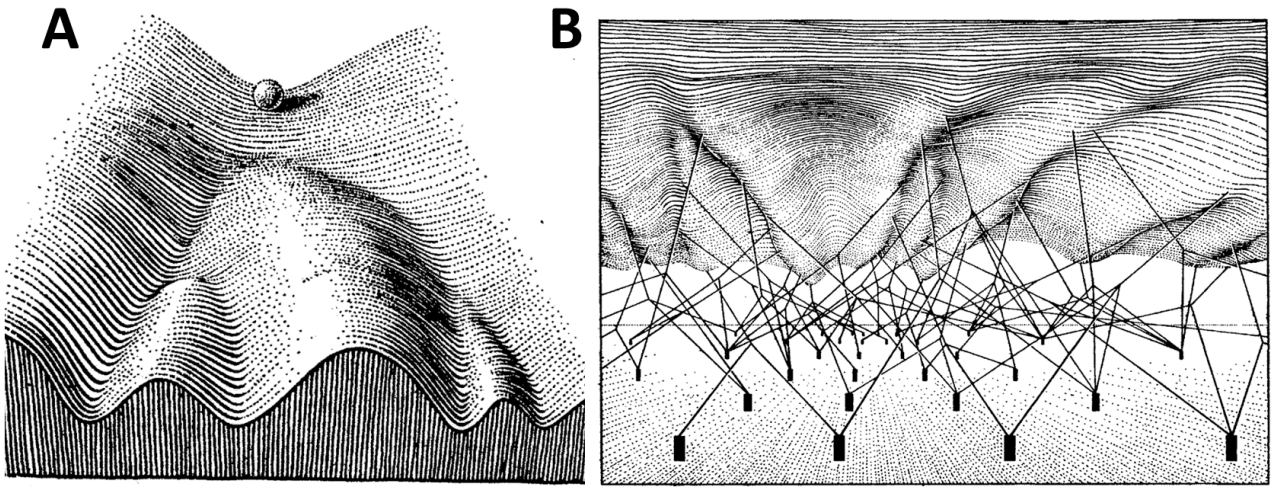

Figure 1.1: Waddington's epigenetic landscape. (A) Cell differentiation as the path of a marble rolling down the forking valleys of a mountainside. **(B)** The complex system of interactions molding the epigenetic landscape, with the genes as pegs pulling the slopes of the epigenetic landscape. Figures taken from Waddington (1957).

2009). It is important to note that heritable in this context refers mainly to daughter cells inheriting the phenotype of the dividing cell, and less to intergenerational inheritance for whole organisms, which is still a hotly debated topic (Horsthemke, 2018). Furthermore, many scholars also consider stable, but not necessarily heritable changes in the transcriptional state to constitute epigenetic traits[1] (Bird, 2007).

These heritable phenotypes without variations of the DNA sequence are possible because of changes in the expression of genes. To come back to Waddington's epigenetic landscape, these changes in gene expression would represent the pegs pulling the hillside into a specific shape, thereby forcing the cell to take a certain pathway for differentiation. The patterns of gene activity during and after cellular differentiation are established and then maintained by epigenetic marks (Atlasi and Stunnenberg, 2017). These can be either (i) chemical modification of the DNA bases or (ii) of the amino acids of the histone proteins around which the DNA is wrapped in eukaryotic cells, or (iii) non-coding RNAs.

In the following sections I will introduce the different epigenetic modifications and mechanisms, focusing on DNA methylation and the regulation of its maintenance. Furthermore, the contribution of epigenetic factors to diseases like cancer and to aging will be discussed. Finally, I'll explain the necessity for the relatively new method of studying epigenetics via the careful perturbation of the epigenetic network in contrast to purely descriptive approaches.

1.1.1 Chromatin

The 2 m long DNA in human cells is packaged into nuclei measuring just a few micrometers. This is possible by wrapping the DNA around proteins called histones, forming the nucleosome. As illustrated in figure 1.2, the nucleosome consists of 146-147 bp of DNA and an octamer of two copies each of the histone core proteins H2A, H2B, H3 and H4 (Luger et al., 1997). This basic unit of the chromatin can then be linked together very tightly, compacting the DNA. At the same time, at least the actively transcribed parts of the genome have to be accessible for the regulatory and transcriptional machinery. Therefore, the DNA is maintained in a very ordered conformation which itself already forms a part of the regulation of gene expression.

At the highest level, DNA can be broadly divided into two different states: accessible euchromatin and densely packed heterochromatin. Generally, heterochromatin contains silenced parts of the genome and is located closer to the laminar regions of the nucleus. On the other hand, actively transcribed genes are part of the euchromatin closer the center of the nucleus. However, these two states be further divided into many different subcategories based on their location, accessibility, the epigenetic marks present at the loci and their transcriptional states (Barski et al., 2007).

Another critical component is the three-dimensional conformation of the DNA. Enhancers must be either brought into close contact with potentially linearly very distant gene pro-

[1] http://www.roadmapepigenomics.org/overview

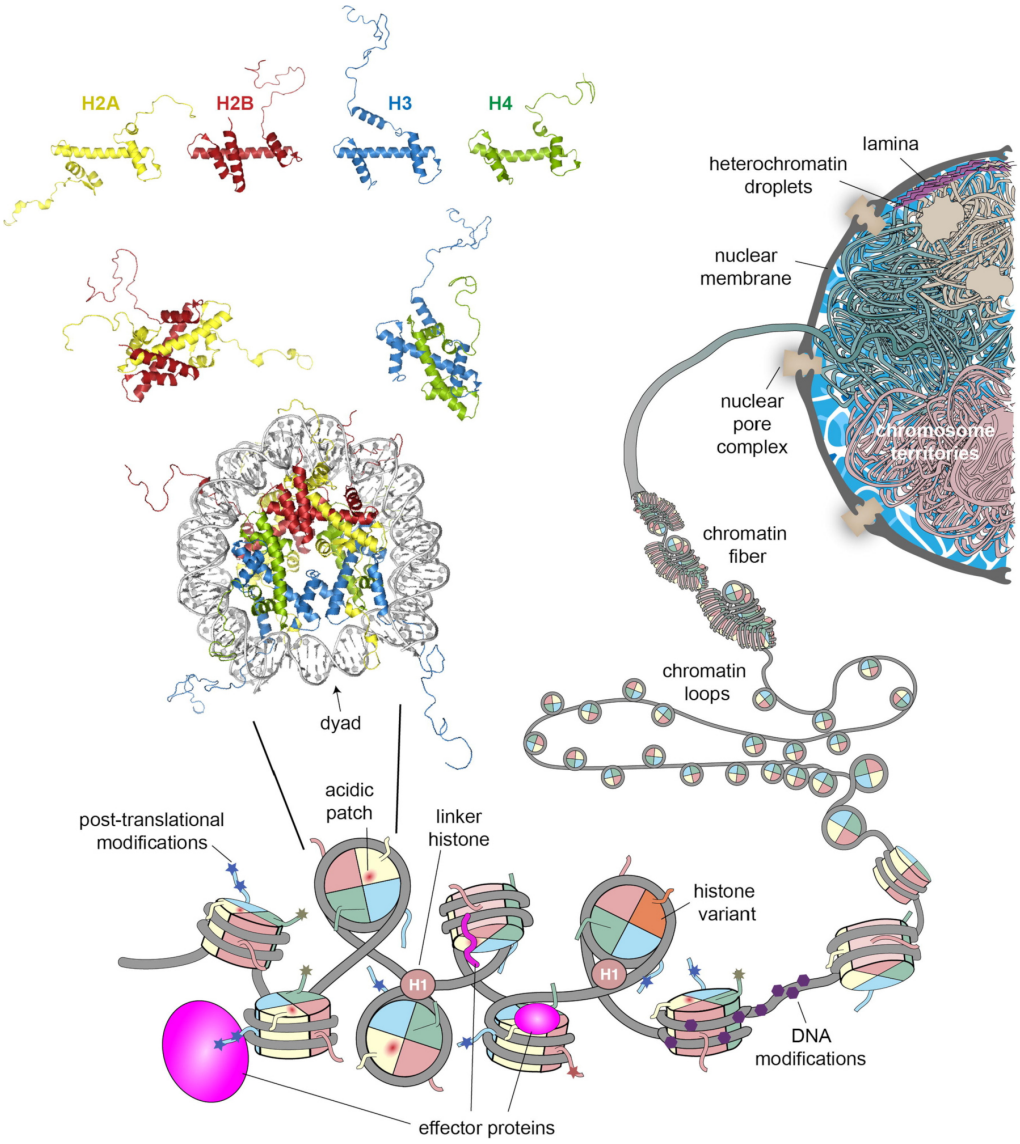

Figure 1.2: Overview of the different organizational levels of chromatin and the nucleosome. The nucleosome is composed of two copies each of the histone protein H2A, H2B, H3 and H4 and 146-147 bp of DNA wrapped around this octamer. Modifications of the DNA and the histone proteins by effector domains establish and maintain the epigenetic marks. The nucleosomes can be further compacted by the linker histone H1, forming the 30 nm chromatin fiber to fit the DNA into the nucleus. Furthermore, loops of chromatin can be extruded to compartmentalize the DNA. This figure is taken from Emmerik and Ingen (2019).

moters or kept insulated from other promoters. Technological progress of chromosome conformation capture (3C) methods has enabled groundbreaking progress regarding this level of regulation via topologically associating domains (TADs) and their role in disease (reviewed in Schmitt, Hu, and Ren (2016), Rowley and Corces (2018), and Schoenfelder and Fraser (2019)).

Chromatin is packed most densely in the metaphase during cell division, when it is visible as individual chromosomes. This is achieved by further compacting and supercoiling the so-called 30 nm fiber. However, this 30 nm fiber itself is already composed of tightly packed nucleosomes compacted by the linker histone H1 (Robinson and Rhodes, 2006). In contrast to this, open chromatin appears as 'beads on a string' formed by accessible DNA with nucleosomes in regular intervals (Khorasanizadeh, 2004) as seen at the bottom of figure 1.2. In open chromatin, binding of transcription factors and the presence of specific epigenetic modifications like DNA methylation and post-translational modifications (PTMs) of histones can then influence the level of gene transcription.

1.1.2 DNA methylation

DNA methylation is the first discovered and most studied epigenetic mark. In mammals, it occurs mostly in the form of cytosine methylated at position 5 in the context of palindromic CpG sites. In humans and other mammals, 70% percent of genome is methylated in most differentiated cell types (Bird, 2002), while stem cells contain significantly less DNA methylation (Berdasco and Esteller, 2011).

The main effect of this epigenetic mark is to repress transcriptional activity. Many genes which are regulated by DNA methylation contain CpG islands (CGI) at their promoters: Stretches of DNA containing highly elevated levels of CpG sites, which are generally depleted in the genome (Bird, 1986). Genes with CGI promoters are generally silenced when methylated but active if the promoter is unmethylated (Jones, 2012). On the other hand, the gene bodies of transcribed genes in open chromatin contain high levels of DNA methylation, likely to repress transcription initiation within the gene body (Neri et al., 2017). Furthermore, transposons and other potentially harmful genetic elements are silenced via DNA methylation (Yoder, Walsh, and Bestor, 1997).

In humans, DNA methylation is set by the maintenance DNA methyltransferase (DNMT) DNMT1 and the *de novo* DNA methyltransferases DNMT3A and DNMT3B. The latter two can form complexes with DNMT3L, which is itself catalytically inactive, but can increase the DNA methyltransferase activity of DNMT3A and B (Chen et al., 2005). It has been shown that DNMT3A (with and without DNMT3L) can form fibers along the DNA (Jia et al., 2007; Rajavelu et al., 2012) which increases its methylation activity (Emperle et al., 2014).

In the classical scheme of DNA methylation outlined in figure 1.3A, DNMT1 is responsible for the reconstitution of fully methylated palindromic CpG sites after DNA replication. Because the newly synthesized DNA strand lacks all modifications, DNMT1 is targeted to

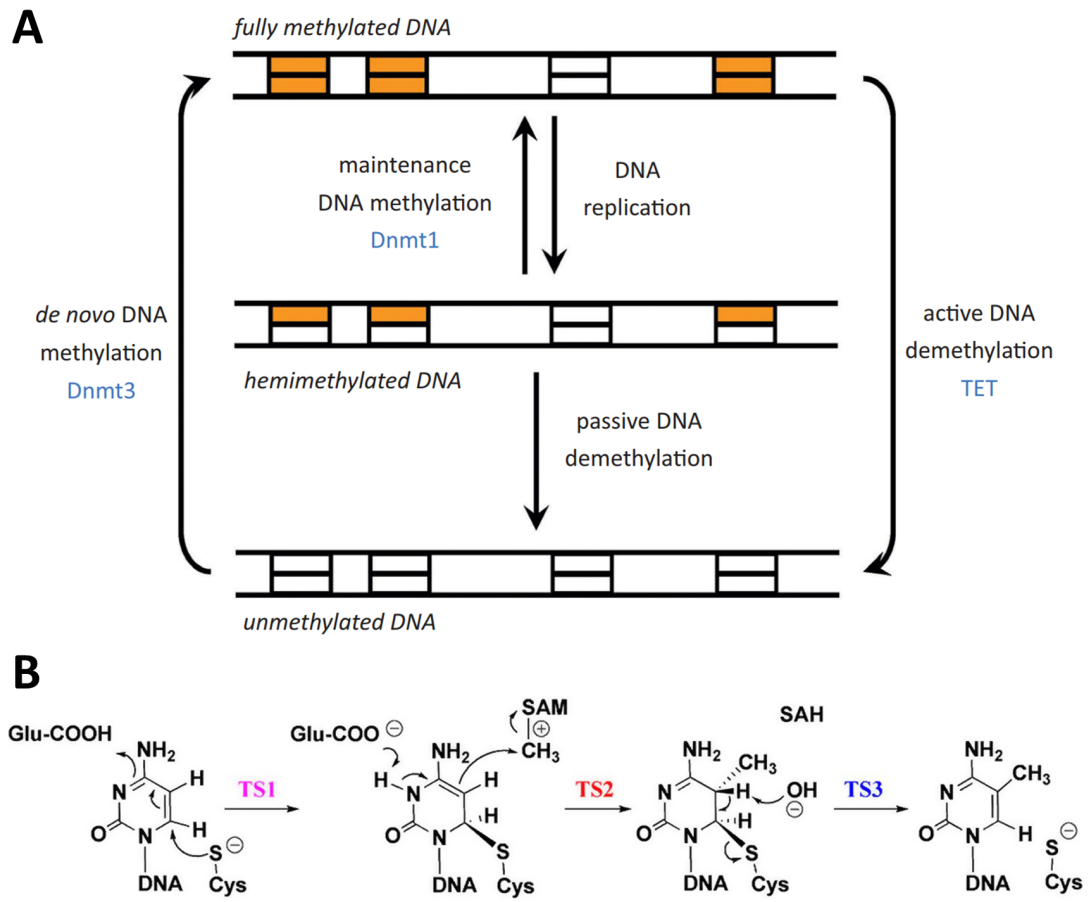

Figure 1.3: Scheme and mechanism of DNA methylation. **(A)** Overview of the DNA methylation and demethylation pathways. Unmethylated CpG sites (white boxes) can be methylated (orange boxes) by the DNMT3 enzymes. After DNA replication, the newly synthesized strand is unmethylated, resulting in a hemimethylated CpG site (half white, half orange boxes). These sites can be either filled up by DNMT1 or lose all methylation after another round of DNA replication. On the other hand, methylated CpGs can be actively demethylated by TET enzymes. This figure is taken from Jeltsch and Jurkowska (2014). **(B)** Transition states of C5 methylation of DNMT1 dependent on the cofactor S-adenosyl L-methionine (SAM). This figure is taken from Du, Wang, and Schramm (2016).

the replication fork via proliferating cell nuclear antigen (PCNA) and ubiquitin like with PHD and ring finger domains 1 (UHRF1) (Iida et al., 2002; Sharif et al., 2007). There, it binds the hemimethylated CpG sites and methylates the CpG site on the unmodified strand. The mechanisms of DNA methylation maintenance will be detailed in section 1.1.5.

The DNMT3 enzymes, on the other hand, methylate transposons and repeats in the genome and establish DNA methylation in differentiating cells (reviewed in Jurkowska, Jurkowski, and Jeltsch, 2011). However, it has become increasingly clear that the functions of the DNMTs are more diverse than anticipated. DNA methylation maintenance by DNMT1 is not perfect and rather a stochastic process, while the DNMT3s also contribute to DNA methylation maintenance (reviewed in Jeltsch and Jurkowska, 2014).

The structures of DNMT enzymes can be broadly divided into regulatory N- and catalytic C-terminal domains (see figure 1.5). The N-terminal part can contain domains binding to specific histone modifications, other proteins and to unmethylated DNA. Furthermore, the

catalytic activity of the enzyme can be controlled allosterically by the N-terminal domains (Jeltsch and Jurkowska, 2016). The C-terminal parts of these DNA methyltransferases methylate the DNA by flipping out the cytosine and transferring the methyl group from *S*-adenosyl *L*-methionine (SAM) to the carbon at position 5 (Lukashevich et al., 2016). This mechanism is shown in figure 1.3B. Site specific analysis of CpG methylation is usually performed via bisulfite conversion and sequencing. This process converts unmethylated cytosines to uracils, while methylated and hydroxymethylated cytosines remain unconverted (see figure 3.14A for the principle of bisulfite conversion). However, novel single-molecule sequencing technologies like SMRT-seq and nanopore sequencing enable the direct detection of modified bases (e.g. Fomenkov et al., 2017; Ni et al., 2019). Unfortunately, these methods have yet to be refined before direct modification sequencing can become a routine technique.

DNA demethylation

DNA methylation was long thought to be a stable mark in the sense that it could only be removed via dilution after DNA replication and failure of the maintenance machinery. However, the discovery that DNA methylation was lost very rapidly and independent of DNA replication in mouse zygotes meant that pathways for active DNA demethylation had to exist (Mayer et al., 2000).

Then, in 2009 it was discovered that a family of enzymes called Ten-Eleven Translocation (TET) can oxidize 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) in cultured cells and *in vitro* (Tahiliani et al., 2009). These enzymes are capable of DNA methylation via multiple oxidative steps and the base excision repair (BER) pathway (reviewed in Pastor, Aravind, and Rao, 2013). Figure 1.4A shows the cycle from 5mC to 5hmC, followed by the next oxidation states 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). The oxidized bases 5fC and 5caC can then be removed by thymine DNA glycosylase (TDG), followed by base excision repair (BER) to reconstitute the unmodified cytosine (He et al., 2011).

TET proteins possess a catalytic C-terminal domain and a regulatory N-terminal part (see figure 1.5). While TET1 and TET3 encode a CXXC domain binding unmethylated DNA, this domain was lost in the case of TET2, where it is encoded by the protein CXXC Finger Protein 4 (CXXC4, also known as IDAX) (Pastor, Aravind, and Rao, 2013). The catalytic mechanism of these dioxygenases shown in figure 1.4B is dependent on α-ketoglutarate (αKG) and Fe(II), which are bound by the C-terminal part of the enzymes. Recently, it has been shown that the vitamin ascorbate is not a cofactor of TET enzymes and instead increases their catalytic activity by reducing Fe(III) to Fe(II) (Hore et al., 2016).

The analysis of 5hmC and the other oxidative states requires multiple steps. Because 5hmC is not converted by bisulfite treatment, it cannot be differentiated from 5mC. Instead, both enzymatic (TAB-seq, Yu et al., 2012a) and chemical treatments (oxBS-seq, Booth et al., 2013) can be used to read 5hmC using bisulfite sequencing. Similarly, the higher oxidative states can also be analyzed by selective treatments (Wu et al., 2014; Song et al., 2013).

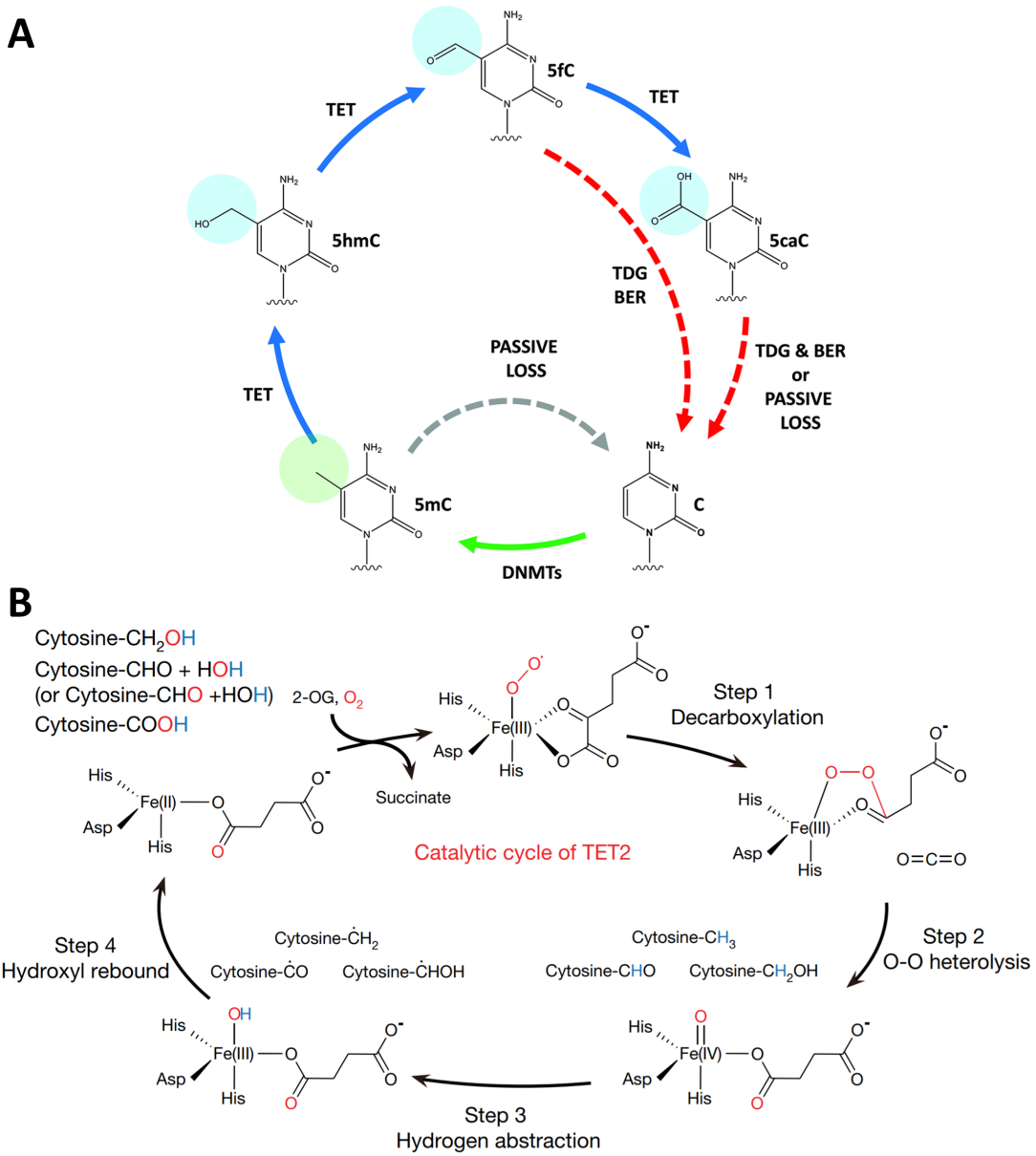

Figure 1.4: DNA demethylation via TET enzymes. **(A)** The step-wise oxidation of 5mC via TETs results in the formation of 5hmC, 5fC and 5acC. The latter two products can then be excised by TDG and repaired via the BER pathway (taken from Ravichandran, Jurkowska, and Jurkowski, 2018). **(B)** Reaction mechanism of TET enzymes dependent on α-ketoglutarate and Fe(II) (taken from Hu et al., 2015).

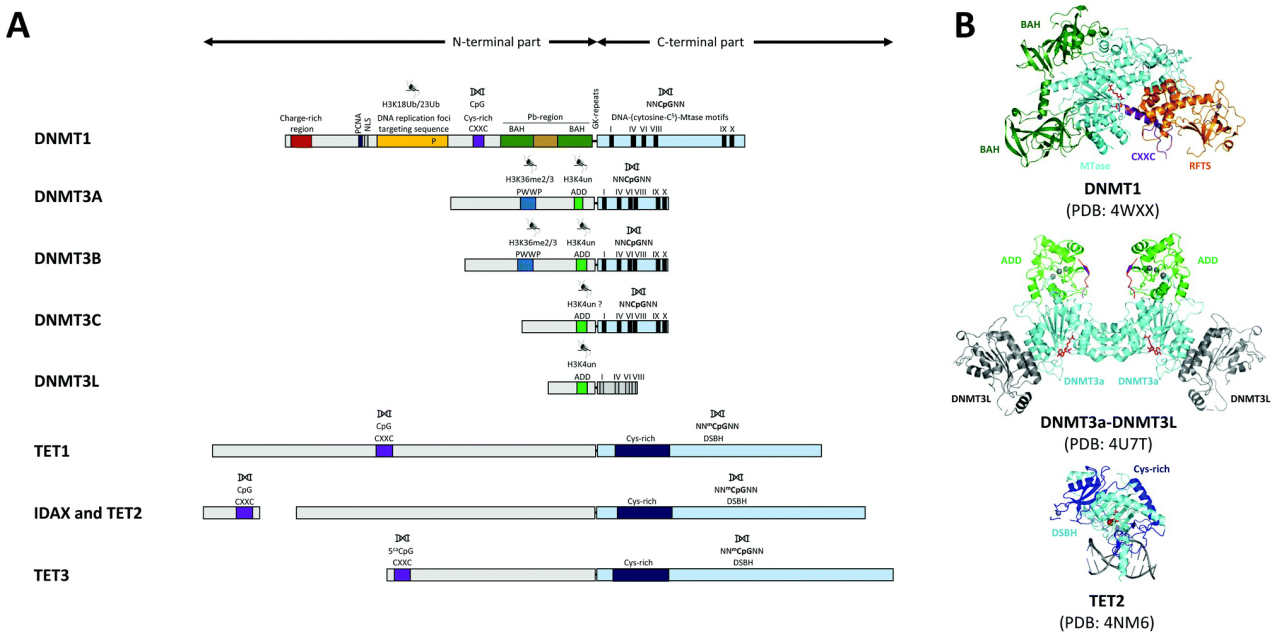

Figure 1.5: Domains and structures of DNMT and TET enzymes. (A) The domain structures of both DNMTs and TETs can be broadly divided into regulatory N-terminal domains and catalytic C-terminal domains. **(B)** The domain-colored crystal structures of DNMT1, DNMT3A-DNMT3L (bound to the H3-tail) and TET2 (bound to DNA) are shown as examples for the structures of the enzyme families. Taken from Ravichandran, Jurkowska, and Jurkowski (2018).

1.1.3 Histone modifications

Histone modifications constitute the second important level of epigenetic control next to DNA methylation. Most of these modifications are placed on the unordered and accessible N-terminal tails of the histone proteins. Methylation, acetylation and phosphorylation are examples for small chemical groups acting as epigenetic signals on histones, but histone tails can be modified with ubiquitin and related proteins as well (reviewed in Meas and Mao, 2015; Cubeñas-Potts and Matunis, 2013). Many of these modifications and the enzymes that can set and remove them are shown in figure 1.6A.

Much work has been done to map the distribution of histone modifications and to discern their correlation with gene expression. This has resulted in the classification of either activating or repressing marks. For example, whereas actively transcribed genes typically contain histone 3 with trimethylated lysines (H3K4me3) around the transcriptional start site, silenced genes are marked with H3K9me3 and H3K27me3 (see figure 1.6B, Kooistra and Helin, 2012). Additionally, lysine acetylation is strongly associated with active transcription and can also influence the strength of DNA binding to the histone cores (Vettese-Dadey et al., 1996). However, especially stem cells can contain marks associated with both active and silenced gene expression at the same promoter (e.g. H3K4me3 and H3K27me3), resulting in bivalent promoters which are poised to quickly differentiate in both directions (Voigt et al., 2012). Detailed maps of which combinations of marks correlate have enabled the definition of many different chromatin states with distinct biological roles (Ernst and Kellis, 2010).

As shown in figure 1.6A, a huge number of enzymes is responsible for the introduction and

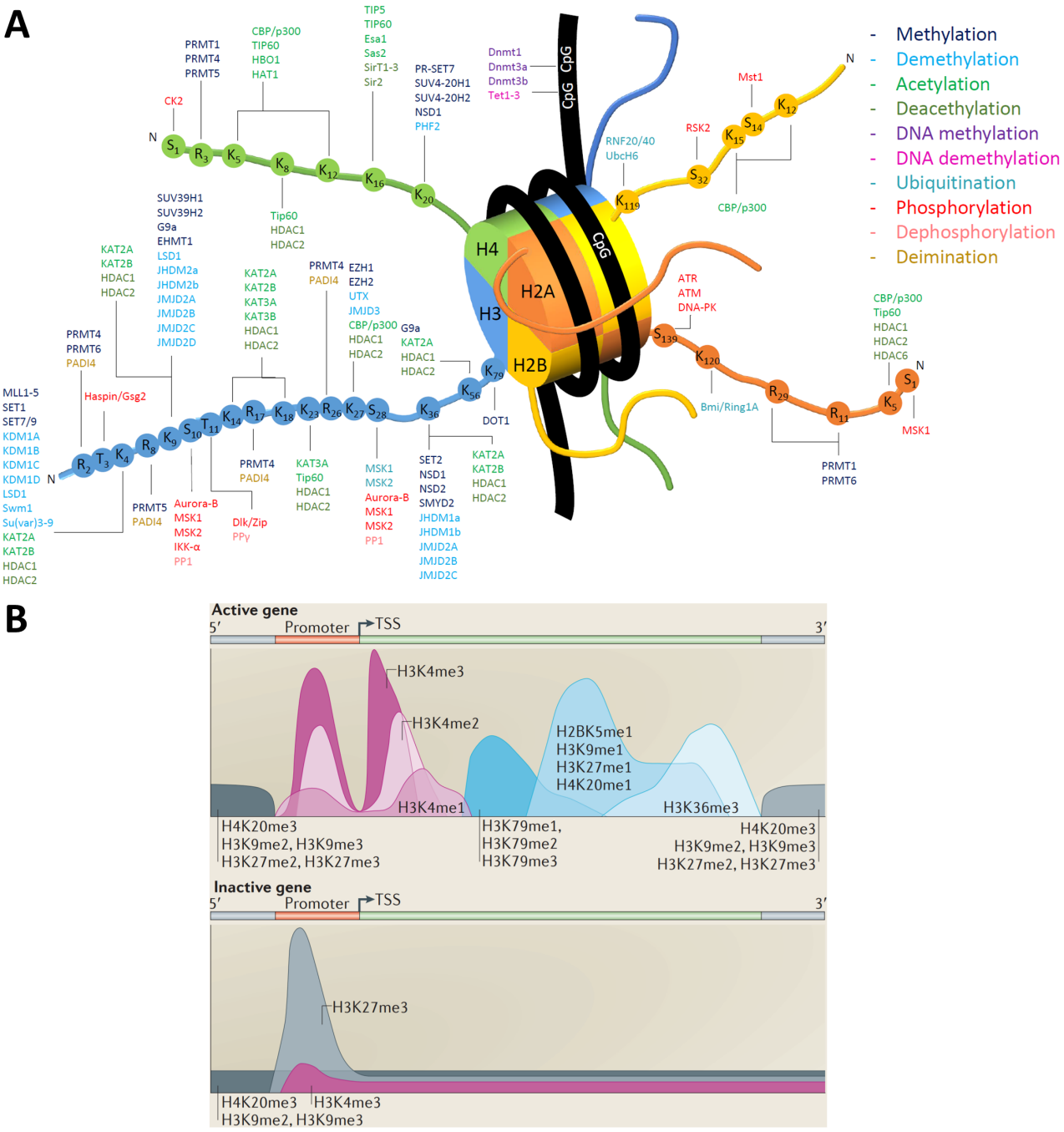

Figure 1.6: Histone post-translational modifications. (A) Overview of epigenetic modifications focusing on many of the histone PTMs and the enzymes responsible for their placement and erasure. This figure was made by Phil Oberacker. **(B)** Typical distribution of histone modifications on active and repressed genes. This figure was taken from Kooistra and Helin (2012).

maintenance of histone PTMs. Many of these enzymes work in complexes: one example is the complex of myeloid/lymphoid or mixed-lineage leukemia (MLL) enzymes such as lysine methyltransferase 2A (KMT2A) with WD repeat domain 5 (WDR5), ASH2L and RPPB5, which together set H3K4 methylation (Nuland et al., 2013). An example for enzymes setting repressive marks is the PRC2 complex around the protein enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2) for methylation of H3K27 (Holoch and Margueron, 2017).

Another histone lysine methyltransferase (HKMT) is also important for this thesis: DOT1 like histone lysine methyltransferase (DOT1L) sets mono, di and trimethylation of H3K79, which is associated with actively transcribed genes and histone acetylation (Hyun et al., 2017; Gilan et al., 2016). Interestingly, for a long time no demethylase had been discovered for these marks (Chory et al., 2019), but recently a first report claimed that lysine demethylase 2B (KDM2B) has H3K79 demethylase activity leading to the recruitment of the sirtuin 1 (SIRT1) histone deacetylase (Kang et al., 2018).

Generally, either the enzyme itself or another protein that is part of the same complex contains reading domains - protein domains which specifically bind a certain modified histone tail. Examples for these binding domains are bromodomains for acetylated lysines and plant homeodomain (PHD) and Pro-Trp-Trp-Pro (PWWP) domains for methylated lysines (Musselman et al., 2012). The enzymes or interaction partners can bind either the same modification they introduce themselves to facilitate spreading of the mark in a positive feedback loop such as for H3K9me3 (Nakayama et al., 2001), or a different modification which either activates or represses catalytic function. The second possibility will be discussed in the next sections.

1.1.4 The epigenetic network

As already indicated in the previous sections, one important feature of epigenetic writers is that they either contain reading domains themselves or interact with reading domains as part of a larger protein complex. This includes not only the interaction of DNA modification-binding domains in DNMTs and TETs and histone PTM-binding domains in histone modifiers, but also interaction between these two layers of epigenetic regulation (Ben-Porath and Cedar, 2001).

This crosstalk is very important for the function of the epigenetic code (Turner, 2007; Bhan, Deb, and Mandal, 2017). Not only does each layer help to direct the correct placement of marks of the other layer, but they also provide a back-up in case one of the layers loses information. Furthermore, the different modifications each have their own distinct stability - both during cell division and regarding changes in gene expression patterns (Bheda and Schneider, 2014). Therefore, they can serve different functions: DNA methylation, as the most stable mark, can act as the final mark cementing a path of cell differentiation (Suelves et al., 2016). DNA methylation also acts as a positive feedback for the re-establishment of

histone modifications after cell division (Hathaway et al., 2012). On the other hand, histone acetylation generally has a short half-life and a high turnover, making it useful for quick changes in gene expression (Zheng, Thomas, and Kelleher, 2013).

A classical example for crosstalk between the different levels is the regulation of DNMT3A activity: The enzyme exists in an autoinhibited state until its ATRX-DNMT3-DNMT3L (ADD) domain binds to unmethylated H3K4, which then enables DNMT3A to bind the DNA (Guo et al., 2015). Similarly, the unmodified H3K4 peptide can also overcome DNMT3A inhibition by methyl-CpG binding protein 2 (MECP2, Rajavelu et al., 2018). This anti-correlation of H3K4me3 and DNA methylation was also shown for DNMT3B at CGI promoters with H3K4me3 in yeast (Morselli et al., 2015). On the other hand, H3K36me3 targets DNMT3s to chromatin via their PWWP domain, which binds the mark in a classical aromatic cage (Dhayalan et al., 2010; Rondelet et al., 2016).

Furthermore, long non-coding RNAs can serve as recruiters or inhibitors of histone and DNA modifying enzymes: for example, Xist and HOTAIR can recruit proteins to introduce marks silencing gene expression, while HOTTIP interacts with MLL-complexes for transcriptional activation and Dali can inhibit DNMT1 (reviewed in Hanly, Esteller, and Berdasco, 2018).

Taken together, these exemplary interactions show that the wide variety of epigenetic modifications and their readers, writers and erasers form an intricate and interdependent network.

1.1.5 Maintenance of DNA methylation

The regulation of the maintenance of DNA methylation is a perfect example for the interplay of many actors and the importance of crosstalk between the different layers of epigenetic marks, forming the epigenetic memory (Bird, 2002; D'Urso and Brickner, 2014). A locus has to contain the correct combination of chromatin marks for functional DNA methylation maintenance.

Firstly, as mentioned in previous section on DNA methylation, the canonical maintenance DNA methyltransferase DNMT1 does not always copy methylation faithfully (Jeltsch and Jurkowska, 2014). Additionally, the activity of DNMT3s is necessary to maintain the high levels of methylated CpG sites at certain loci such as transposons and repeats (Arand et al., 2012).

Secondly, the activity and targeting of the DNMTs is controlled by histone modifications and interaction with other proteins. As stated above, H3K4 has to be unmodified for DNMT3s to be active, whereas H3K36me3 recruits DNMT3s to gene bodies. In the case of DNMT1, most of the targeting and activity control is performed by the protein UHRF1. DNMT1 is recruited and activated by UHRF1 at hemimethylated DNA (Bostick et al., 2007; Bashtrykov et al., 2014). The DNA binding is conferred by the SET and RING-associated (SRA) domain of UHRF1. Together with the PHD, these domains of UHRF1 are required for DNA methylation

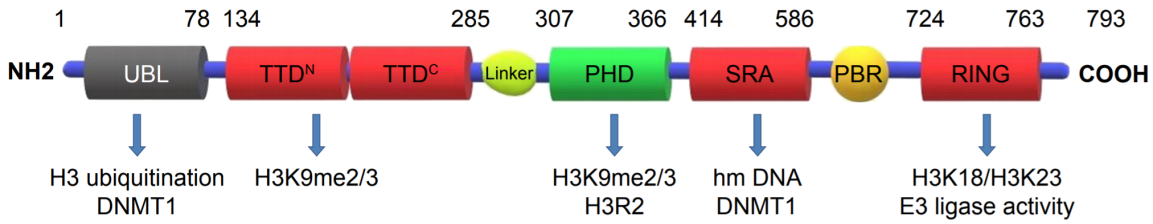

Figure 1.7: Domains of UHRF1 and their functions. The ubiquitin-like (UBL) and really interesting new gene (RING) finger domains ubiquitinate H3K18 and H3K23. The SRA domain binds hemimethylated DNA and the PHD and tandem tudor domains bind H3R2 and H3K9me2/3. This figure was taken from Xue et al. (2019).

maintenance (Kong et al., 2019). An overview of the domains of UHRF1 and their functions is shown in figure 1.7. This targeting function is so important that a knockout (KO) of UHRF1 has a stronger impact on DNA methylation levels than a KO of DNMT1 (Arand et al., 2012). Interestingly, it seems that the DNMT3 enzymes interact with UHRF1 as well (Meilinger et al., 2009), which could also explain this result.

Apart from DNA binding, one of the main functions of UHRF1 is the binding and reading of histone tail modifications. The PHD of UHRF1 binds unmodified H3K4, similar to the DNMTs, and also unmodified H3R2 (Hu et al., 2011). Additionally, H3K9me3 is recognized via its PHD and tandem tudor domains to target DNA methylation maintenance (Rothbart et al., 2013). On the other hand, H3K9 methylation was found to be dependent on DNA methylation in human cancer cells (Nguyen et al., 2002). This interdependence again illustrates the importance of crosstalk between the different layers of epigenetic regulation.

Lastly, UHRF1 can ubiquinate both DNMT1 and histone tails (Citterio et al., 2004; Qin, Leonhardt, and Spada, 2011). This is the final targeting mechanism for DNMT1, as DNMT1 can also read the H3K23 ubiquinated by UHRF1 (Nishiyama et al., 2013). This means that the interaction of UHRF1 and DNMT1 can even be only indirect for DNA methylation maintenance to work.

Taken together, these results show that UHRF1 is the main protein responsible for the regulation of DNA methylation maintenance. UHRF1 reads and integrates the local chromatin information to decide whether or not to maintain DNA methylation at a given locus. This provides a safeguard against spurious DNA methylation, also making UHRF1 an important potential target in cancer therapy (Kong et al., 2019; Xue et al., 2019).

However, as shown in figure 1.3A, loss of DNA methylation can occur not only through a lack of maintenance, but also via active DNA demethylation. Therefore, the other important consideration is the targeting and regulation of TET enzymes, but unfortunately much is still uncertain about this (Melamed et al., 2018; Ravichandran, Jurkowska, and Jurkowski, 2018). They are targeted to both methylated and unmethylated CpG rich loci via their CXXC domains. It appears that in the case of the CXXC domain of TET3, 5caC is the preferred target (Jin et al., 2016).

Furthermore, TET enzymes interact with many other proteins including Nanog, which could target TETs to genes involved in pluripotency (Costa et al., 2013). Another interesting

interaction partner is PRC2, which seems to recruit TETs to bivalent promoters in embryonic stem cells (ESCs) where they accumulate 5hmC (Neri et al., 2013). Additionally, mRNAs of TET enzymes can be regulated by micro RNAs (miRNAs) and the proteins are modified post translation as reviewed by Wu and Zhang (2017), changing their localization and activity. Examples include ubiquitination and acetylation of lysines and PARylation (Ciccarone et al., 2015). Finally, different isoforms with altered domain compositions possessing slightly different functions make the regulation even more complicated (Melamed et al., 2018).

1.1.6 Epigenetics in disease

As epigenetic enzymes and marks regulate gene expression, epigenetic mechanisms are implicated in many diseases (reviewed in e.g. Zoghbi and Beaudet, 2016). Unsurprisingly, many of these diseases are related to failures in differentiation and programming pathways, which is when epigenetic control is most important. However, there are also diseases where localized changes in epigenetic marks and gene control contribute to disease.

An example for localized deficiencies of epigenetic control are imprinting disorders (reviewed by Tucci et al., 2019). Imprinting refers to the DNA methylation and histone modification dependent silencing of either the maternal or paternal allele of a gene, which is set in the germ line. In humans, around 200 genes have been found to be regulated via imprinting. Examples for disorders are the neurological defects in Angelman syndrome (AS) and Prader–Willi syndrome (PWS). Both are caused by loss of imprinting on a stretch of chromosome 15 (15q11-13), for example by deletion of the imprinting center. In the case of AS, maternal expression of *UBE3A* is missing, and in the case of PWS, paternal expression of multiple genes in that region is lost (Tucci et al., 2019).

Another disease with an epigenetic mechanism is the severe neurological disorder Rett syndrome (RTT, reviewed by Lyst and Bird, 2015). The disorder is caused by loss-of-function missense or truncation mutations in *MECP2* (Lombardi, Baker, and Zoghbi, 2015). As MECP2 is one of the bridges between DNA and histone modifications, the mutations can have far reaching consequences and have allowed researches to map the interaction interfaces of MECP2 (Lyst and Bird, 2015).

Cancer is both a genetic and an epigenetic disease with dramatic changes in chromatin structure, epigenetic marks and gene expression (Flavahan, Gaskell, and Bernstein, 2017). In general, tumors exhibit a global loss of DNA methylation but at the same time regional hypermethylation, including at promoters of tumor suppressor genes (Jones and Baylin, 2002). Examples for tumor suppressor genes regularly silenced in cancer include p16/cyclin dependent kinase inhibitor 2A (*CDKN2A*), BRCA1 DNA repair associated (*BRCA1*) and the RB transcriptional corepressor 1 (*RB1*) (Esteller, 2008). The changes in chromatin during cellular differentiation can be partially rolled back in cancer transformation, which can be similar to reprogramming (see figure 1.8, Suvà, Riggi, and Bernstein, 2013). It is becoming clear that these epigenetic alterations are not just passengers but drivers of tumor transforma-

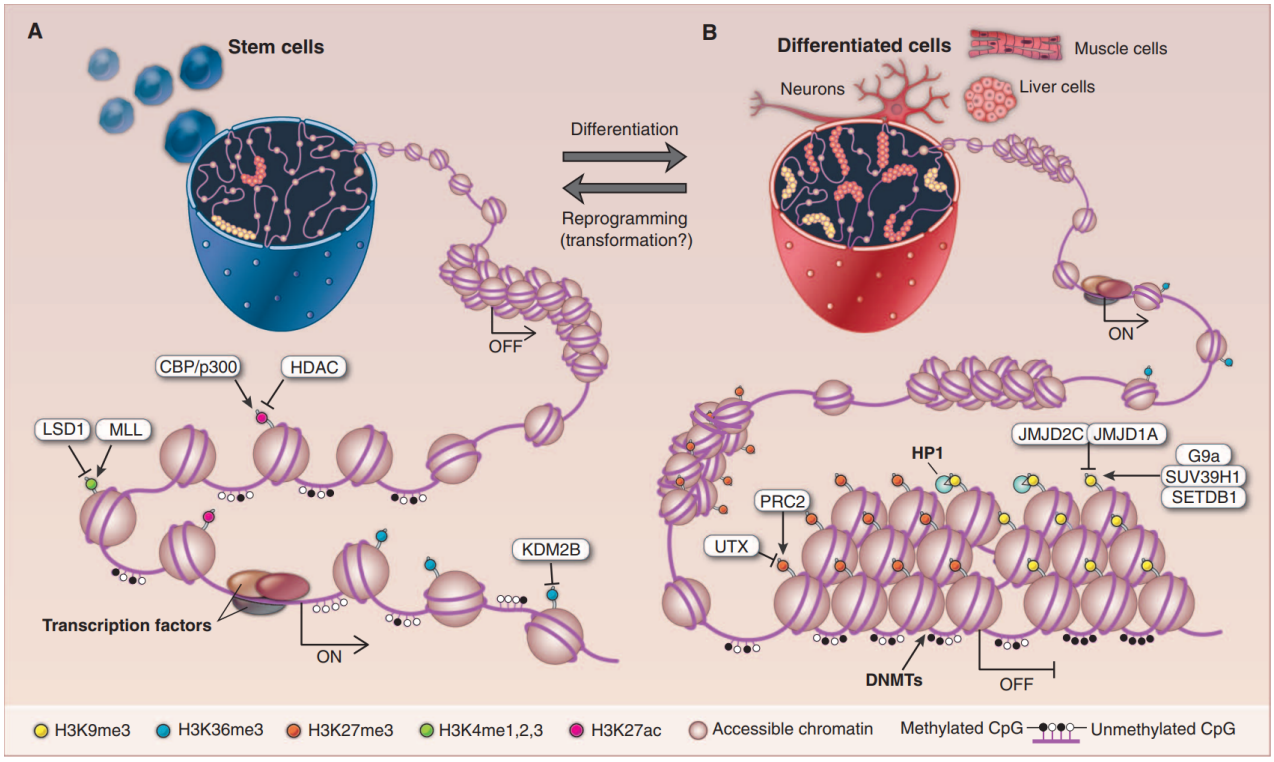

Figure 1.8: Changes in chromatin during cellular differentiation can be partially reprogrammed in cancer transformation. This figure was taken from Suvà, Riggi, and Bernstein (2013).

tion, leading to global changes in gene expression patterns (Chatterjee, Rodger, and Eccles, 2018). Of particular interest are fusion proteins produced by chromosomal translocations as reviewed in Brien, Stegmaier, and Armstrong (2019), e.g. between MLL1/KMT2A and transcription factors such as AF10/MLLT10 or components of the elongation complex. These fusion proteins can be strong oncogenes, leading to oncogenic changes in gene expression (Brien, Stegmaier, and Armstrong, 2019). In some forms of acute myeloid and lymphoblastic leukemia, DOT1L is recruited by the oncogenic fusion proteins MLL1-AF9, MLL1-AF10 and MLL1-ENL, blocking the deposition of repressive histone modifications (Deshpande et al., 2014; Chen et al., 2015).

Interestingly, there has been progress finding interactions between metabolism and epigenetics in cancer. The canonical example are mutations in isocitrate dehydrogenase 1 and 2 (IDH1 and IDH2), which convert isocitrate to αKG. The reaction product of the mutants, (D)-2-hydroxyglutarate, can inhibit αKG dependent demethylases such as TETs, leading to epigenetic deregulation (Dang, Yen, and Attar, 2016). Recently, a link between folate metabolism and BRD4 in cancer was discovered (Sdelci et al., 2019).

Diabetes

Although not a classical epigenetic disease, it appears that epigenetics might play a role in both the autoimmune disorder type 1 diabetes (Jerram, Dang, and Leslie, 2017) and adult-onset type 2 diabetes (Ling and Rönn, 2019). However, more of interest to my studies was the potential generation of insulin-producing cells from glucagon-secreting α-cells. α- and

β-cells are closely related, and it was shown that the genes *Arx* and *Pax4* are master regulators of their cell identities (Collombat et al., 2009; Dhawan et al., 2011); see also section 1.4.1). Interestingly, two studies had suggested that by silencing *Arx* or expressing *Pax4* in α-cells it's possible to induce insulin production in α-cells and potentially even transdifferentiate them to β-cells (Friedman-Mazursky, Elkon, and Efrat, 2016; Zhang et al., 2016). This might provide a potential therapeutic avenue for diabetes treatment.

Epigenetics of aging

Age is the most important risk factor for many deadly diseases such as cancer, neurodegeneration and cardiovascular disease (Sen et al., 2016). Aging correlates with cellular senescence, which is a growth arrested state which cells adopt in response to oncogenic stimuli and after many cell divisions to protect against further damages and potential cancer transformation (Campisi, 2013). During aging, the epigenetic control of gene expression changes and deteriorates, which can also induce cellular senescence. Interestingly, inducing apoptosis of senescent cells with p16 expression increases healthy lifespan by 17% to 35% in mice (Baker et al., 2016) and clinical trials with senolytics are under way (Kirkland et al., 2017; Deursen, 2019).

One hallmark of aging are changes in DNA methylation: while repeat regions lose DNA methylation, promoters show increased DNA methylation levels. These changes have become the basis for the most accurate estimator for biological age so far: Steve Horvath trained a neural network on DNA methylation data from many different donors and tissues and found extremely strong correlations with age for a set of 353 CpG sites (Horvath, 2013; Horvath and Raj, 2018). Another cellular symptom of aging is the loss of heterochromatin and of histones in general, which also increases susceptibility to DNA breaks and transposon activation (Pal and Tyler, 2016).

1.1.7 Epigenetic editing

Most of our knowledge about epigenetics comes from descriptive approaches. Large consortia like Roadmap (Bernstein et al., 2010), ENCODE (Davis et al., 2018) and Blueprint (Fernández et al., 2016), together with initiatives to facilitate sharing and harmonizing datasets like IHEC (Bujold et al., 2016), have revolutionized our understanding of epigenetics with enormous data sets.

However, only describing epigenetics states and their correlation with gene expression seems unable to answer pressing fundamental questions: are epigenetic marks the cause of a specific transcriptional state or merely the byproduct? What is the exact function of each of the many marks and how do they interact with each other and the writers, readers and erasers? Are there any marks and enzymes which represent master switches, or are always combinations necessary to switch and maintain an epigenetic state? That's why intentionally and precisely perturbing the epigenetic network via targeted and precise modifications can be an important and valuable tool to understand the epigenetic code. This method is

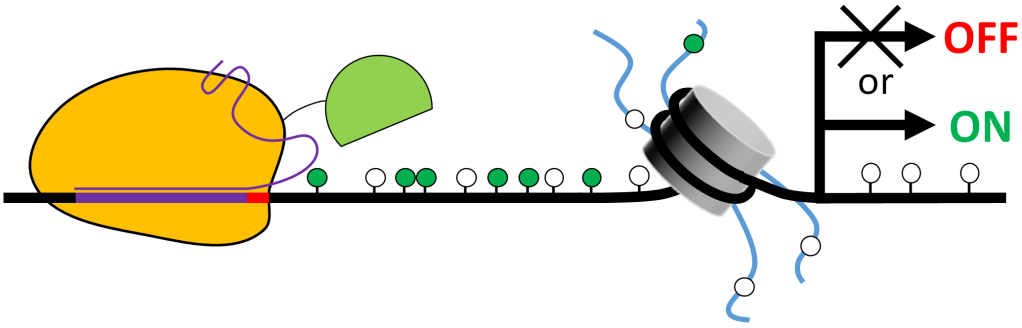

Figure 1.9: Principle of epigenome editing. The method relies on a sequence specific DNA binding domain which can be e.g. dCas9, shown in orange with the gRNA in violet and the PAM site in red (see section 1.2.2). This domain is either fused to or can otherwise recruit an epigenetic enzyme (light green) which can then set or remove specific marks (white and green circles) on the DNA and histones in the vicinity of the binding site, leading to changes in gene expression.

called epigenetic editing or epigenome editing (Kungulovski and Jeltsch, 2016) and can be considered to be part of the larger field of synthetic epigenetics. Synthetic epigenetics is defined broader than epigenome editing and refers to "the design and construction of novel specific artificial epigenetic pathways or the redesign of existing natural biological systems, in order to intentionally change epigenetic information of the cell at desired loci" (Jurkowski, Ravichandran, and Stepper, 2015).

As shown in figure 1.9, epigenetic editing uses a programmable binding domain to recruit epigenetic effectors to a specific genetic locus. Generally, this means that a sequence specific DNA binding domain is fused to enzyme that can then write or erase a certain mark on the DNA or histone tails in the vicinity of the targeted DNA sequence. However, the effector domain (ED) could also be recruited via intermediates for a modular system, or might even not change epigenetic marks itself but instead bind other enzymes or complexes which then write or remove epigenetic modifications.

Obviously an even younger field than epigenetics itself, initial progress for epigenome editing was slow. This was also due to the cumbersome generation of targeting domains, but this changed with the discovery of CRISPR-Cas9 (see section 1.2). To circumvent the construction and validation of DNA binding domains for endogenous target genes, many groups instead decided to use reporter constructs, either on plasmids or integrated into the genome. However, these approaches fail to capture the intricacies of endogenous loci with their interplay of neighboring and distant loci, enhancers and an established, controlled and homeostatic epigenetic state.

Nevertheless, epigenome editing has shown increasing successes over the years. The initial studies have been reviewed by Groote, Verschure, and Rots (2012), but since the first study demonstrating the use of CRISPR-Cas9 for direct epigenome editing with the histone acetyltransferase p300 (Hilton et al., 2015), the field has exploded. Quickly afterwards, DNA methylation with DNMT3A (Vojta et al., 2016) and DNMT3a3L (Stepper et al., 2017) and DNA demethylation with TET1 were demonstrated with dCas9 (Morita et al., 2016; Xu et al., 2016). Similarly, histone methylation was both set with PRDM9 and DOT1L

(Cano-Rodriguez et al., 2016) and removed with LSD1 (Kearns et al., 2015).

Furthermore, genome wide screens for regulatory elements were performed utilizing the modular and high throughput methods available thanks to the modular nature of CRISPR-Cas9 (Joung et al., 2017; Klann et al., 2017). Scientists also engineered novel and highly active combinations of effectors such as VP64-p65-Rta (VPR, Chavez et al., 2015), including modular systems (Cheng et al., 2016; Joung et al., 2017) which also allow stronger signals by recruiting multiple copies (Tanenbaum et al., 2014). There were also studies investigating the single cell dynamics of epigenetic editing, albeit at reporters (Bintu et al., 2016).

A few studies also investigated the stability of the changed epigenetic state, but with contradictory results. Some reported that the introduced marks were lost after days to weeks (e.g. Kungulovski et al., 2015), whereas others reported stable changes to target gene expression (e.g. Amabile et al., 2016; Saunderson et al., 2017). There were also reports that the success and stability of epigenome editing was dependent on the basal state of the local chromatin. Introduced H3K4me3 was lost after less than 10 days when the target gene promoter was DNA methylated (Cano-Rodriguez et al., 2016). Furthermore, multiple studies demonstrated that combinations of multiple epigenetic effector domains achieve stronger effects than using only component: Amabile et al. (2016) combined DNMT3A, DNMT3L and the Krüppel associated box (KRAB), which recruits SETDB1 for H3K9 methylation and the NuRD complex for histone deacetylation (reviewed in Thakore et al., 2016). Yeo et al. (2018) fused KRAB to MECP2, creating a new potent transcriptional repressor.

Taken together, epigenome editing has proven the causal power of epigenetic marks on gene expression and has since continued to instead use this method to elucidate the epigenetic network and move from cells into organisms (e.g. Morita et al., 2016; Williams et al., 2018).

1.2 Targeting proteins

1.2.1 Zinc-finger proteins and TALEs

Cys_2-His_2 (C2H2) zinc-finger proteins were the first sequence-specific programmable DNA binding proteins. As shown in figure 1.10A and B, each finger recognizes 3-4 DNA bases via the major groove with their beta-beta-alpha structure around a zinc ion (Wolfe, Nekludova, and Pabo, 2000). The sequence specificity of C2H2 zinc fingers can be programmed by changing the base-contacting amino acids. However, the code is not perfect and depends on neighboring zinc-fingers and the DNA modification state of the target (Pabo, Peisach, and Grant, 2001). To improve sequence specificity, arrays with multiple zinc fingers can be constructed to recognize longer DNA sequences. Fusing nucleases or transcriptional modifiers to these arrays generated the first tools for genetic (Kim, Cha, and Chandrasegaran, 1996) and epigenetic editing (Kim et al., 1997), but the necessity to validate every new zinc finger array made the work impractical for many applications and widespread adoption.

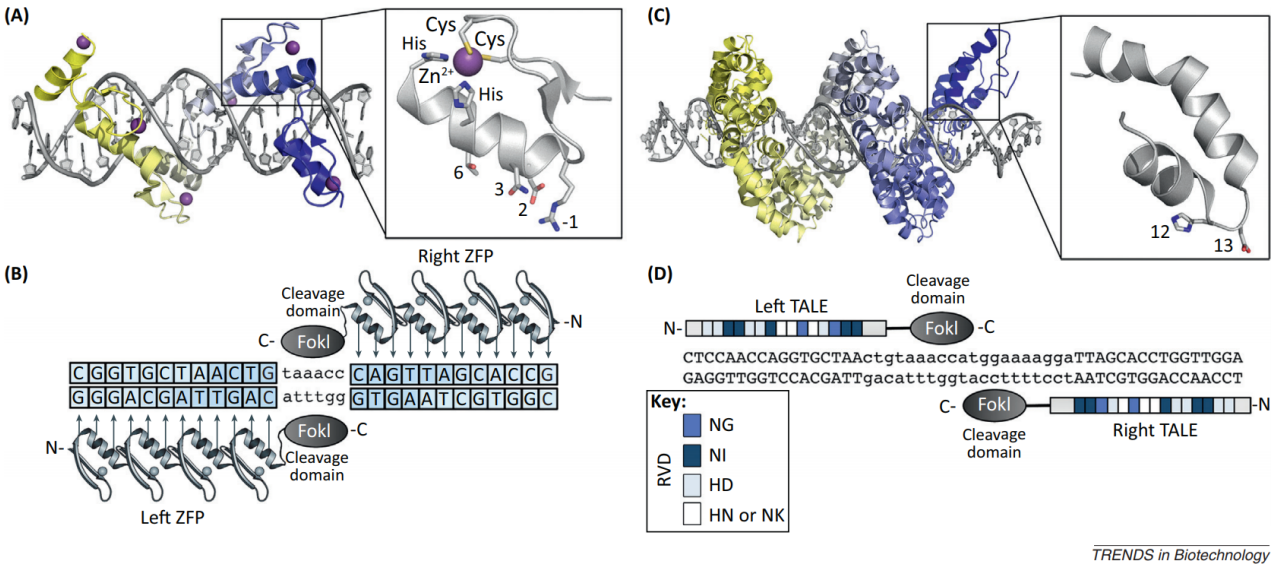

Figure 1.10: Programmable DNA binding via zinc finger arrays and TALEs. **(A)** Structure of a programmed zinc finger array contacting the desired DNA sequence. **(B)** Two zinc finger arrays coupled to the FokI nuclease for sequence specific DNA double strand breaks at a palindromic target site. **(C)** and **(D)** show the same for TALEs. This figure was taken from Gaj, Gersbach, and Barbas (2013).

Transcription-activator-like effectors (TALEs) were discovered later in the bacterial plant pathogen *Xanthomonas* as the next class of programmable sequence specific DNA binding proteins (Boch et al., 2009; Moscou and Bogdanove, 2009). TALE proteins are composed of repeat peptides with 34 amino acids, each repeat contacting one base pair (see figure 1.10C and D). As only two amino acids per repeat are responsible for base recognition and they are governed by a simple code, it is much easier to design new sequence specifies for TALEs than for zinc-finger proteins (Reyon et al., 2012). Just as for zinc-fingers, this capability has been harnessed to edit the genome (Joung and Sander, 2013) and the epigenome (Zhang et al., 2011).

Although both of these programmable and sequence specific DNA binding proteins generated important discoveries and each have their own advantages, they were both eclipsed by the modular nature, speed and flexibility of the CRISPR-Cas9 system.

1.2.2 CRISPR-Cas9

The clustered regularly interspaced short palindromic repeats (CRISPR) in bacteria were already discovered in 1987, but it took twenty more years until their role in bacterial immunity was discovered (reviewed by Hsu, Lander, and Zhang, 2014; Doudna and Charpentier, 2014). A shown in figure 1.11A, fragments of foreign DNA elements like phages are incorporated as spacers into the repeat sections. The region is then transcribed and processed to guide the endonuclease CRISPR-associated 9 (Cas9) to cleave and thereby inactivate the DNA of invading phages (Barrangou and Marraffini, 2014). The target DNA complementary to the CRISPR RNA (crRNA) transcribed from the spacer has to contain a certain protospacer

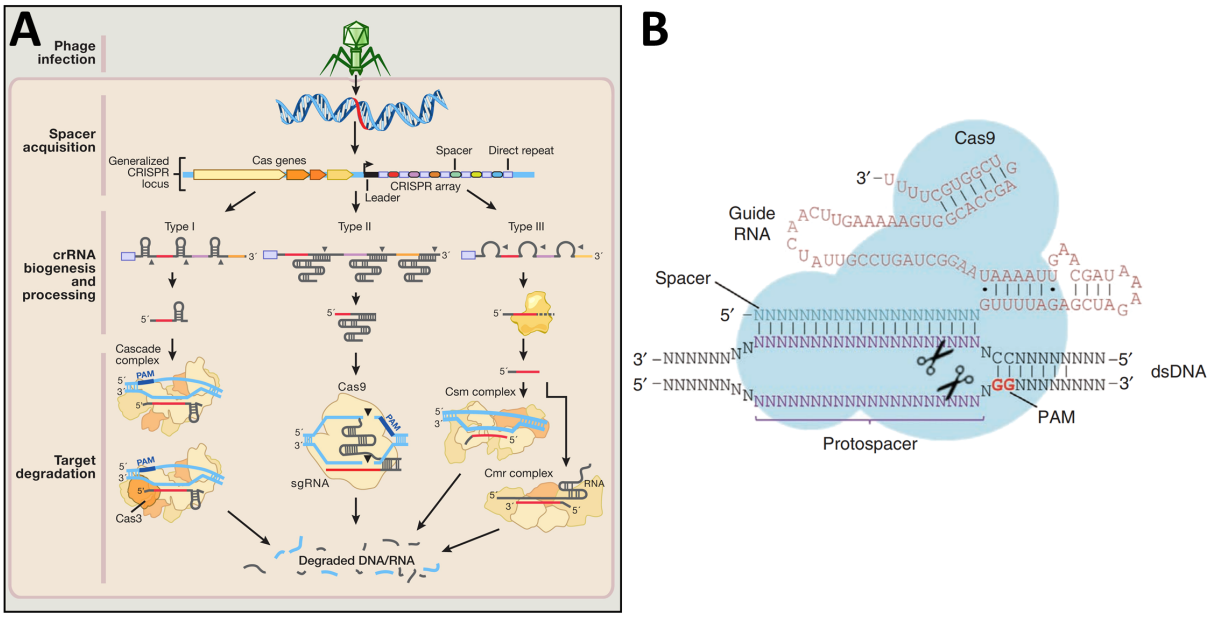

Figure 1.11: Function of CRISPR and Cas9. **(A)** DNA fragments from an infecting phage is incorporated by the bacterial cell into the CRISPR locus, which is then transcribed and processed by one of multiple Cas pathways to cleave the invading DNA. This figure was taken from Hsu, Lander, and Zhang (2014). **(B)** Scheme of Cas9 cleaving the protospacer DNA complementary to the spacer region of the gRNA, next to the PAM site. This figure was taken from Mali, Esvelt, and Church (2013).

adjacent motif (PAM) next to the complementary sequence for binding and cleavage to occur (see figure 1.11B). Furthermore, an additional RNA fragment, the trans-activating RNA (tracrRNA), is necessary for binding and cleavage. This enables the introduction of targeted double-strand breaks (DSBs) in DNA as shown in 2012 *in vitro* using Cas9 from *Streptococcus pyogenes* (Jinek et al., 2012) and one year later in mammalian cells (Cong et al., 2013; Mali et al., 2013). These studies used chimeric guide RNAs which combined both the crRNA and the tracrRNA into a single guide RNA (sgRNA, see figure 1.11B), transforming CRISPR-Cas9 into a powerful modular system with only two components.

Since then, CRISPR-Cas9 has taken the biology world by storm and has revolutionized or birthed many applications (reviewed by Pickar-Oliver and Gersbach, 2019). Examples include genetic screens with the use of gRNA libraries (Wang et al., 2014; Shalem et al., 2014) and model organism generation (Li et al., 2013). CRISPR-Cas9 has enabled gene drives which could modify whole species (Gantz et al., 2015; Hammond et al., 2016), gene editing in human embryos (Liang et al., 2015) and even the birth of the first gene-edited babies (Normile, 2018).

However, mutating the catalytic residues of the two endonuclease lobes of Cas9 generates an easily programmable and sequence specific targeting domain instead of a cleavage protein. This nuclease-deficient dead Cas9 (dCas9) enabled many powerful methods such as multiplexed sequence specific labeling (Ma et al., 2016), the investigation of chromatin composition and interactors at specific loci (Myers et al., 2018), base editing (Eid, Alshareef, and Mahfouz, 2018) or localized mutagenesis (Hess et al., 2016). As already mentioned

above, dCas9 was also adopted as the tool of choice for targeted modulation of the gene expression and the epigenome. Targeted activation and repression of gene expression was shown already in 2013 (Gilbert et al., 2013; Maeder et al., 2013) and direct epigenome editing in 2015 (Hilton et al., 2015). CRISPR-Cas9 has since allowed the field to thrive.

1.3 Magnetic beads for high-throughput handling of nucleic acids

Current methods for the isolation and handling of nucleic acids are still mostly based on silica-matrix columns. Although these are also available in 96-well formats, they are either very expensive or hard to adapt to high throughput workflows. Most commercially available kits require centrifuges and, for single-tube workflows, this increases the handling time linearly with each additional sample until the limit of 24 to 30 slots available in normal tabletop centrifuges, restricting the possible throughput.

Surface coated magnetic beads can be an excellent choice for handling high sample numbers, either by hand or in an automated fashion, as they are suitable for liquid handling stations without centrifuges and can be easily scaled up (Hawkins et al., 1994; Fisher et al., 2011; Rohland and Reich, 2012). Because they are generally used in 8-strips or 96-well plates, many samples can be processed simultaneously using multichannel pipettes.

These beads are nano- to micro-scale particles with a ferrimagnetic core and a functionalized surface. Generally, either a silica- or carboxyl-group based coat is used. Nucleic acids can bind to these surfaces under dehydrating conditions in a reversible manner in a process called solid-phase reversible immobilization (SPRI; DeAngelis, Wang, and Hawkins, 1995). Nucleic acids bound to the beads can be immobilized using magnets for liquid exchange steps such as washing with 70% ethanol and can be eluted in water.

The magnetic nanoparticles as the core of the beads can be synthesized via a wide variety of methods (reviewed in Singh et al. 2014; Wu et al. 2015; Majidi et al. 2016). For life sciences, the co-precipitation of Fe(II) and Fe(III) solutions by the addition of a base is the most common method due to its simplicity (Majidi et al., 2016). To coat the produced nanoparticles, tetraethylorthosilicate (TEOS) is the chemical of choice for silica surfaces (Kim, Kim, and Kim, 2007), while carboxyl-groups are usually achieved via the polymerization of methacrylic acid (MAA) monomers (Aggarwal et al., 1996).

The research community has already established popular protocols to substitute the buffers for the AMPure XP bead based clean-up steps which are part of most sequencing library preparation protocols (Rohland and Reich, 2012; Jolivet and Foley, 2015). However, very few other protocols are available and barely any laboratories take the next step in cost-savings and synthesize their own magnetic beads.

Faced with this situation when we wanted to increase the size of our experiments and scale up the number of samples to answer our experimental questions, we could neither

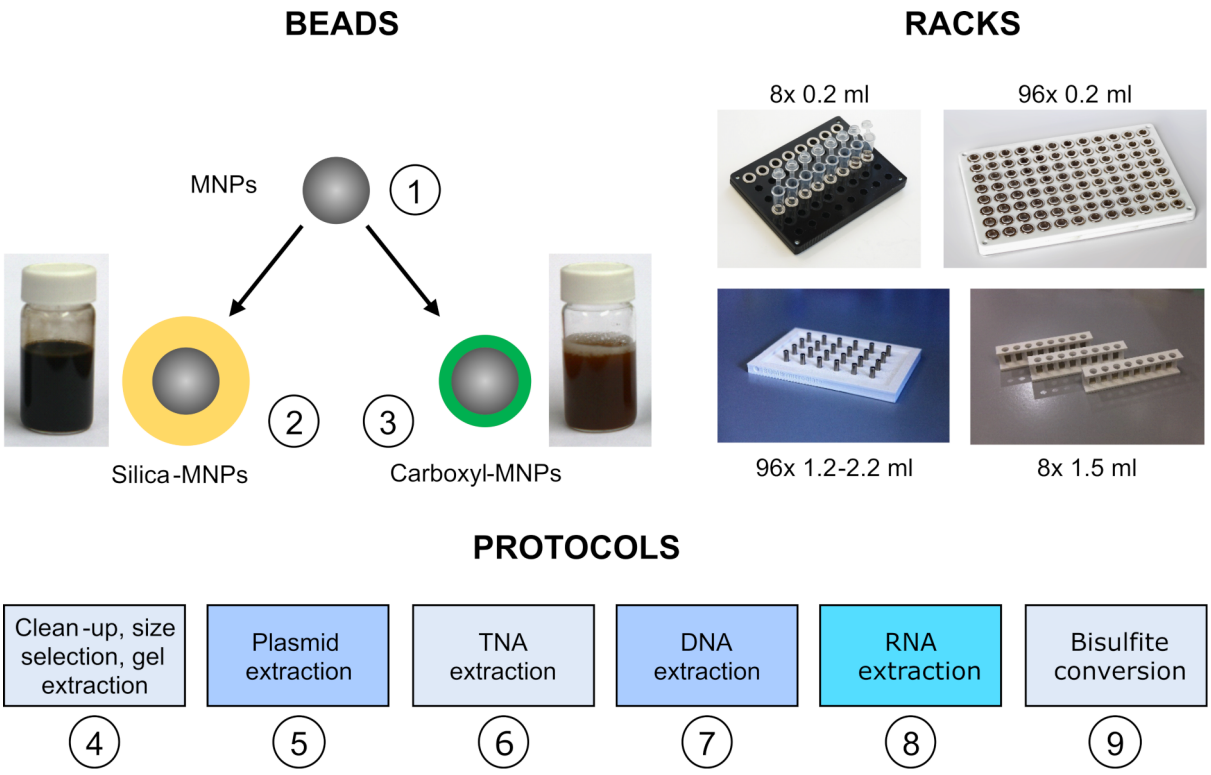

Figure 1.12: The components of the Bio-On-Magnetic-Beads (BOMB) platform (Oberacker et al., 2019). The three pillars of the platform are functionalized magnetic ferrite nano-particles (MNPs, 1) with either a silica (2) or carboxyl surface (3); magnetic racks produced by laser cutting or 3D printing with ring or cylindrical magnets; and protocols for the purification and manipulation of nucleic acids (4-9). The protocols are indicated by the numbers in the circles.

continue using the single tube kits we had used until then, nor afford the commercially available 96-well based high-throughput solutions. Therefore, we sought to expand the available methods for magnetic beads based nucleic acid handling from the published clean-up protocols to cover the daily needs in our work-flows: the isolation of DNA and RNA from bacteria and mammalian cells, the isolation of plasmids from bacteria, and the bisulfite conversion of isolated DNA. As most of these applications follow the principle of lysis, nucleic acid binding, washing followed by an optional manipulation step such as enzymatic digestion and finally elution of the purified nucleic acids, we aimed to establish modular buffer systems and workflows. Additionally, to keep the cost down and further reduce our reliance on commercial providers, we investigated the feasibility to synthesize and functionalize magnetic beads ourselves and were successful with the introduction of simple and flexible protocols.

Together with homemade magnetic racks in multiple formats, these three parts formed the pillars of the platform we then named Bio-On-Magnetic-Beads (BOMB) and released to the scientific community as open-source protocols[2] (Oberacker et al., 2019). The constituents of the platform are illustrated in figure 1.12.

[2]https://bomb.bio/

1.4 Principal aims of the study

Epigenetic regulation of gene expression is crucial for the development of multicellular organisms and its deregulation leads to diseases like cancer and aging. However, purely descriptive efforts appear unable to decipher the large network of epigenetic enzymes and marks. The precise perturbation of this network via epigenome editing seems uniquely positioned to fill this knowledge gap. For this thesis, I wanted to use this method to elucidate the epigenetic network, especially the mechanisms of epigenetic memory with a focus on DNA methylation maintenance.

Additionally, the research community is lacking open-sourced methods for cost-effective high-throughput isolation and handling of nucleic acids. The growing sample numbers in my epigenome editing experiments necessitated the development of solutions for this problem.

Development of a toolbox for epigenome editing

Therefore, the first aim of my thesis was to re-purpose and validate the CRISPR-Cas9 system for epigenome editing. Initially, this work focused on targeted DNA methylation via Dnmt3a and the single-chain fusion Dnmt3a3L following the initial experiments in my MSc thesis. The DNMTs were targeted to four genes in two human cell lines and I studied the effects on target gene transcription and promoter DNA methylation. Furthermore, I investigated the cooperative setting of DNA methylation through Dnmt3a multimerization by targeting mutant proteins with disrupted binding interfaces.

Additionally, to complete the epigenome editing toolbox I also needed a way to activate gene expression. Hence, I used engineered TET enzymes for targeted DNA demethylation and gene activation and studied the combination of the DNA demethylases with other transcriptional activators on multiple target genes.

Elucidate the mechanisms of DNA methylation maintenance

Over the last years it has become clear that artificially introduced DNA methylation via epigenetic editing is often not stable over multiple cell divisions (see also section 1.1.7). This means that some factors which are present at endogenous, non-edited methylated loci are missing at the targeted sites. Therefore, the introduced DNA methylation appears to be recognized as an epigenetic mutation by the cells, which then repair that mutation. Two possibilities exist for that repair: on the one hand, cells could actively demethylate the regions via TET enzymes, or on the other hand, could simply suppress the maintenance of DNA methylation at these loci. To investigate the cause, I studied the kinetics of DNA methylation establishment and loss while inhibiting TET enzymes or reducing the rates of cell division. Furthermore, I aimed to facilitate a complete switch of the epigenetic state of the target genes via simultaneous targeting DNA methylation, H3K9 methylation and histone deacetylation and additionally specifically inhibiting MLL complexes and DOT1L.

Studying tumor transformation with targeted DNA methylation

In a collaboration with the group of Dr. Gabriella Ficz we studied the effects of targeted promoter DNA methylation of tumor suppressor genes on senescence escape and tumor formation.

Generation of insulin-producing cells from pancreatic α-cells

In a collaboration with the group of Dr. Melita Vidaković we investigated a potential therapeutic application of epigenome editing by inducing insulin-production in murine α-cells via targeted DNA methylation of the promoter of the master regulator *Arx*.

Development of an open-source platform for cost-effective high-throughput nucleic acid handling

The ever-increasing number of samples in my studies to answer more complex experimental questions forced us to find solutions to enable high-throughput sample processing while staying on budget. Therefore, we aimed to develop low cost, high-throughput methods based on self-made magnetic beads. My focus in this were the molecular biology methods necessary in my studies: the isolation of RNA and DNA and efficient bisulfite conversion.

1.4.1 Target genes

Several genes were targeted for epigenome editing in this thesis. They were selected based on published results for their capacity to be either activated or repressed by targeting epigenetic effector domains to their promoters. The genes and target sites will be detailed in the coming section.

EPCAM

EPCAM is the gene symbol for the oncogene epithelial cell adhesion molecule. The gene encodes a trans-membrane, Ca^{2+}-independent cell-cell adhesion protein implicated in cell cycle and proliferation control (Münz et al., 2004; Gun et al., 2010) including the β-catenin/Wnt signaling pathway (Yamashita et al., 2007). *EPCAM* gene expression is upregulated in several cancer types, especially breast, colorectal and cervical cancers and has been proposed as a target for therapeutic transcriptional reprogramming (Gun et al., 2010). The utility of this approach has already been validated in several *in vitro* studies with zinc finger fusions (Gun et al., 2013) or small interfering RNAs (siRNAs) (Osta et al., 2004). Furthermore, it was employed as the target gene for epigenetic editing studies (Nunna et al., 2014; Stepper, 2015).

The gene is strongly expressed in the human ovarian cancer cell line SKOV-3 but is mostly repressed in HEK293 cells, with a DNA-methylated promoter. This makes it an ideal target to study both targeted activation and repression, depending on the cell line used for targeting. A map of the *EPCAM* promoter region including the position and direction of the employed gRNAs is shown in figure 1.13A.

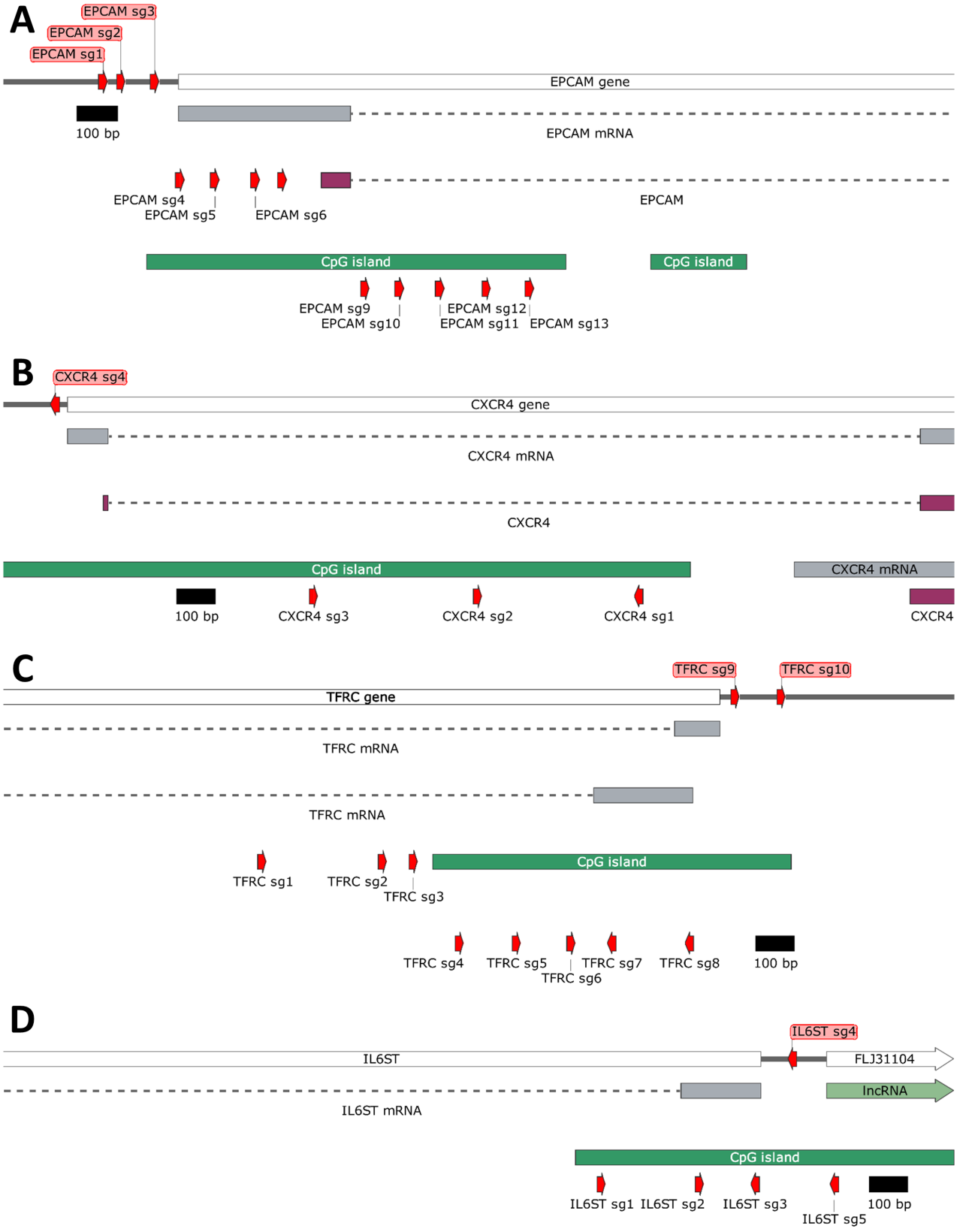

Figure 1.13: Map of the target gene promoter regions. (A) *EPCAM* on chromosome 2 with the gene itself indicated as a white box. Transcribed regions are represented as grey bars with the introns as dashed lines. Protein coding sequences are shown in violet and CpG islands in green. gRNAs targeting the region are represented as red arrows indicating the direction of the PAM site. **(B)** *CXCR4* on chromosome 2. **(C)** *TFRC* on chromosome 3. **(D)** *IL6ST* on chromosome 5. The figures were generated with SnapGene.

CXCR4

The C-X-C motif chemokine receptor 4 with the gene symbol *CXCR4* is a transmembrane, stromal cell-derived factor-1-specific G protein-coupled receptor. Together with CD4, it is used by HIV to gain entry into cells (Feng et al., 1996). Furthermore, it is overexpressed in several cancer types such as colorectal cancer (Cabrero-de Las Heras and Martínez-Balibrea, 2018) and is linked to metastasis and invasiveness (Sun et al., 2010; Susek et al., 2018). *CXCR4* has previously been used successfully as a target for transcriptional reprogramming with dCas9-KRAB (Gilbert et al., 2013) and dCas9 recruiting VP64 (Zalatan et al., 2015).

Except for an only 15 bp long first exon in some splicing variants, it is a monoexonic gene. Thus, expression analysis via RT-qPCR is more cumbersome as potentially remaining gDNA could be amplified as well and would not be distinguishable from complementary DNA (cDNA). A schematic map of the *CXCR4* promoter region including the position and direction of the employed gRNAs is shown in figure 1.13B.

TFRC

TFRC (also known as *CD71*) encodes the transferrin receptor protein, which is responsible for the binding of transferrin loaded with Fe(III). In some cancers, the gene is abnormally upregulated (Shen et al., 2018). *TFRC* was used for targeted repression with dCas9-KRAB by Gilbert et al. (2013). A schematic map of the *TFRC* promoter region is shown in figure 1.13C.

IL6ST

The gene interleukin 6 signal transducer with the symbol *IL6ST* (also known as *gp130*) codes for a transmembrane unit of receptors for multiple cytokines, including interleukin 6 (IL6), and leukemia inhibitory factor (LIF) (Yoshida et al., 1996). Because of its increased expression in various cancer types and correlation with poor prognosis, it is being investigated as a potential cancer therapy target (Xu and Neamati, 2013). *IL6ST* expression was successfully repressed by targeting Dnmt3a for DNA methylation with dCas9 by Vojta et al. (2016). See figure 1.13D for a map of the *IL6ST* promoter region including the position and direction of the employed gRNAs.

ASCL1

The gene achaete-scute family bHLH transcription factor 1 (*ASCL1*) encodes a basic helix-loop-helix transcription factor which binds to the E-box, activating transcription. It acts as a master regulator in the development and differentiation of neurons (Chouchane and Costa, 2018). The gene was activated successfully via targeted effector domains in multiple studies (Konermann et al., 2015; Chavez et al., 2016). See figure 1.14A for a map of the promoter region of the gene *ASCL1*.

LIN28A

LIN28A is the symbol for the gene lin-28 homolog A and encodes a miRNA- and mRNA-binding protein like its homolog *LIN28B*. They are expressed mostly in stem cells and can increase reprogramming success for iPS cells (Yu et al., 2007). Both genes are frequently

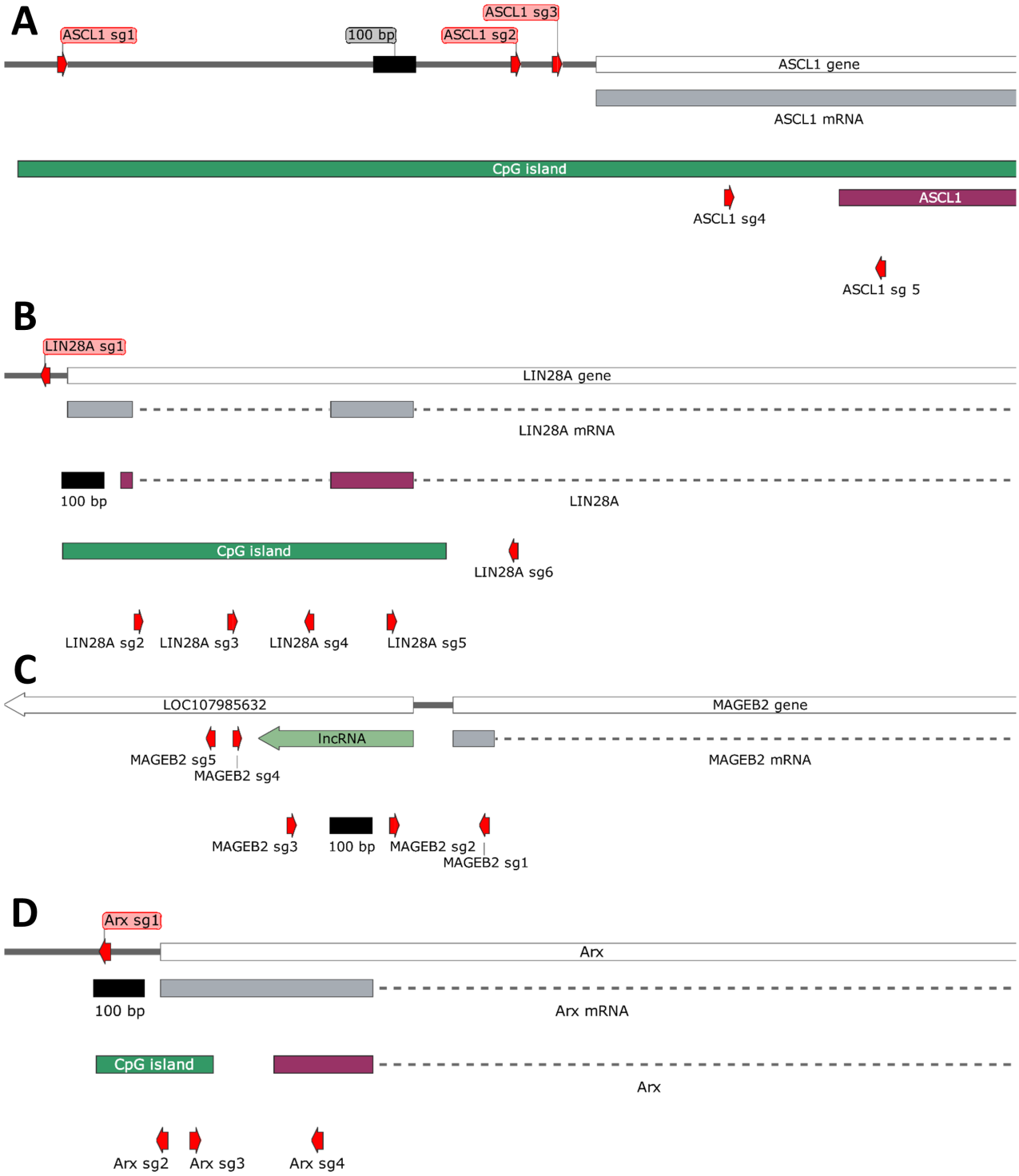

Figure 1.14: Map of the target gene promoter regions. (A) *ASCL1* on chromosome 12 with the gene itself indicated as a white box. Transcribed regions are represented as grey bars with the introns as dashed lines. Protein coding sequences are shown in violet and CpG islands in green. gRNAs targeting the region are represented as red arrows indicating the direction of the PAM site. **(B)** *LIN28A* on chromosome 1. **(C)** *MAGEB2* on the X chromosome. **(D)** *Arx* on the murine X chromosome. The figures were created with SnapGene.

upregulated in a wide variety of cancer types and overexpression is associated with a poor prognosis (Wang et al., 2015). *LIN28A* was demonstrated as a possible target for targeted activation with dCas9 recruiting VP64 (Konermann et al., 2015). A map of the *LIN28A* promoter region including the position and direction of the employed gRNAs is shown in figure 1.14B.

MAGEB2

The MAGE (melanoma antigen) family member B2 (*MAGEB2*) gene encodes a protein of mostly unknown function which seems to enhance E2F activity (Peche et al., 2015). The gene is activated in various tumors and can influence the expression of other genes and promote growth, like many other genes in the MAGE family. These genes are also apparently regulated mainly via DNA methylation and are often overexpressed in tumors with DNA hypomethylation (Van Tongelen, Loriot, and De Smet, 2017). *MAGEB2* was also shown to become actively expressed upon treatment of HEK293 cells with histone deacetylase (HDAC) inhibitors or DNA methyltransferase inhibitor 5-aza-2'-deoxy-cytidine (Milutinovic et al., 2007). This was also achieved by targeting the catalytic domain of Tet1 to the MAGEB2 promoter with dCas9 (Xu et al., 2016). See figure 1.14C for a map of the *MAGEB2* promoter region including the position and direction of the employed gRNAs.

Arx

The gene aristaless- related homeobox (*Arx*) is a master regulator for pancreatic α-cell identity. When this lineage determination gene is activated in pancreatic β-cells, they are converted to α-cells (Dhawan et al., 2011). On the other hand, the gene paired box 4 (*Pax4*) encodes a transcription factor that, when overexpressed, can convert α-cells to β-cells (Collombat et al., 2009). Interestingly, *Arx* downregulation via short hairpin RNA (shRNA) induced insulin production in human β-cell-derived cells. This represents a potential therapeutic intervention for the treatment of β-cell deficiency in diabetes (Friedman-Mazursky, Elkon, and Efrat, 2016). Figure 1.14D shows a map of the murine *Arx* promoter region including the position and direction of the employed gRNAs.

2 Materials and methods

2.1 Plasmid preparation

2.1.1 Fusions of epigenetic effector domains to dCas9

Generation of single domain fusions

To generate the constructs used in this thesis, several plasmids available via Addgene served as the templates (see table 2.1). To fuse the epigenetic effector domains to the human codon-optimized, nuclease-null Cas9 from *Staphylococcus pyogenes*, the M-SPn-VP64 plasmid (Addgene plasmid #48674, a gift from George Church (Esvelt et al., 2013)) was used. The plasmid backbone was amplified with the vector-specific primers dCas9_vector_f and dCas9_vector_r (see appendix table 5.3), exposing the C-terminus of dCas9 for the insertion of a linker and the effector domain. Primers with overhangs complementary to the ends of the vector backbone were used to amplify the effector domains from their respective source plasmids. Table 2.1 lists all dCas9-ED fusions and their source plasmids; and appendix tables 5.3 and 5.4 additionally list all the primers used for their construction. Figure 2.1 shows the most important effector domains fused to dCas9 drawn to scale.

All PCRs were performed using Q5 High-Fidelity DNA Polymerase (NEB). The vector and insert PCR products were then assembled using either Gibson Assembly Master Mix or NEBuilder HiFi DNA Assembly Master Mix (both NEB). After assembling the constructs, the plasmids were transformed into either XL1blue MRF' or TOP10 electro-competent *Escherichia coli* cells. Resulting colonies were screened for the presence of the assembled construct with colony PCR or after plasmid isolation with restriction digestion. Correct plasmid assembly was confirmed by Sanger sequencing via the companies LGC Genomics or Microsynth.

The origin and construction of the effector domains employed in this study will be detailed in the following paragraphs.

The Dnmt3a3L fusion, which combines the catalytic domain of murine Dnmt3a with the catalytically inactive Dnmt3L, was initially constructed and characterized by Siddique et al. (2013). In addition to the fusion of this highly active DNA methylase to dCas9, I also generated several mutants of both Dnmt3a and Dnmt3a3L. The C706A mutant deactivates the catalytic site (Reither et al., 2003), while the R832E substitution stops the formation of lateral Dnmt3a(3L) multimerization (Rajavelu et al., 2012). The two single codon exchange mutants and constructs combining both substitutions were generated using megaprimer mutagenesis with the primers listed in appendix table 5.2.

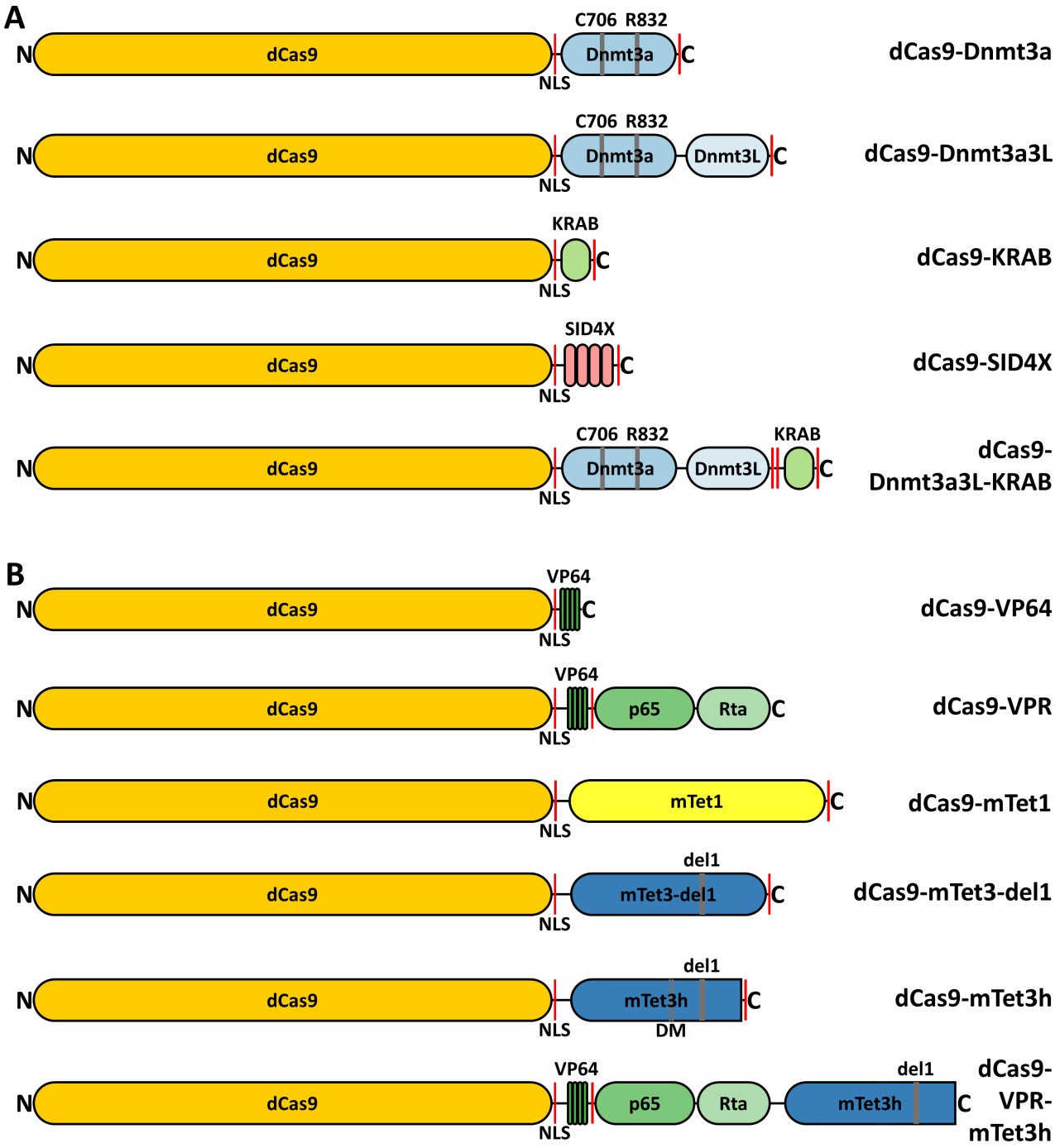

Figure 2.1: Schematic showing the most important dCas9-ED fusion constructs for (A) targeted gene repression and (B) targeted gene activation. The proteins are drawn to scale.

pHAGE EF1α dCas9-KRAB (Addgene Plasmid #50919), the plasmid used as the template for the construction of the dCas9-KRAB fusion, was a gift from Rene Maehr & Scot Wolfe (Kearns et al., 2013).

The template for the dCas9-SID4X plasmid, pAAV_hSyn_TALEBB(HD)-NLS-SID4X_2A-_phiLOV2.1_WPRE_bGHpA (Addgene Plasmid #47450), was a gift from Feng Zhang (Konermann et al., 2013).

dCas9-VPR, a fusion of the effector domains VP64, p65 and Rta, was used unchanged from SP-dCas9-VPR (Addgene Plasmid #63798), a gift from George Church (Chavez et al., 2015). For the fusion of the murine AICDA (also known as AID) protein to dCas9, a synthetic gBlock (IDT) of amino acids 1-169 from NP_033775.1 was amplified.

The engineered variants of the catalytic domain of murine Tet3, mTet3-del1, mTet3ΔC, and mTet3h (short for hyperactive, originally mTet3-del1-ΔC) were designed, constructed and analyzed *in vitro* by Ravichandran (2017). I cloned the domains from the bacterial expression constructs into the dCas9 fusion expression vector for this study as described above. However, after initial experiments, longer linkers composed of $(Gly_4\text{-}Ser)_4$ were introduced in between dCas9 and the Tet catalytic domains via Gibson Assembly. The initial constructs will be referred to as 'sl' (short linker) in this thesis.

The catalytically dead double mutant of mTet3h was generated using megaprimer mutagenesis using the primers listed in appendix table 5.2.

Similarly, the initial dCas9-sl-nTet construct was cloned from a pEXA2 vector containing the sequence XP_002667965.1 (synthesized by MWG Operon), coding for the Tet-like dioxygenase-1 from the heterolobosean amoeboflagellate *Naegleria gruberi*. Its mutants A212V and A212T with drastically reduced oxidation rate of 5hmC (Hashimoto et al., 2014, and Ravichandran et al., manuscript in preparation) were cloned from bacterial expression vectors provided by Dr. Ravichandran.

For all transfection experiments, plasmids were isolated in a larger scale and higher quality with QIAGEN Plasmid Midi Kits. For plasmids containing the gene for the wild-type (WT) Dnmt3a3L protein, adding glucose or using a bacterial growth medium with glucose like SOC for Midipreps and selection plates was necessary as the highly active DNA methyltransferase can cause toxic effects to the bacteria.

Generation of fusions with multiple effector domains

In addition to the primary effector domain fusions described above, I also generated constructs containing multiple effector domains fused to dCas9. Specifically, I combined the repressors KRAB and Dnmt3a3L. KRAB was fused to either the N- or the C-terminus of dCas9-Dnmt3a3L using PCR amplification (see appendix table 5.3 for the primers) and Gibson Assembly.

Furthermore, I fused both the activator VPR and engineered versions of the catalytic domain of mTet3 to dCas9. Several permutations were cloned and studied for efficacy:

Table 2.1: List of plasmids with fusions of dCas9 to various epigenetic effector domains used in this thesis with the source plasmid of the effector domain. The mutants dCas9-Dnmt3a(3L) C706A, R832E and the combination of both and dCas9-mTet3h DM were all created with megaprimer mutagenesis based on the wild-type plasmids (see appendix table 5.2).

dCas9-ED fusion	source plasmid(s)
dCas9-Dnmt3a3L	ZNF-Dnmt3a3L (Siddique et al., 2013)
dCas9-Dnmt3a	ZNF-Dnmt3a CD (Siddique et al., 2013)
dCas9-KRAB	pHAGE EF1α dCas9-KRAB (Addgene Plasmid #50919)
dCas9-SID4X	pAAV_hSyn_TALEBB(HD)-NLS-SID4X_2A_phiLOV2.1_WPRE_bGHpA (Addgene Plasmid #47450)
KRAB-dCas9-Dnmt3a3L	pHAGE EF1α dCas9-KRAB and dCas9-Dnmt3a3L
dCas9-Dnmt3a3L-KRAB	pHAGE EF1α dCas9-KRAB and dCas9-Dnmt3a3L
dCas9-VP64	unchanged from M-SPn-Cas9-VP64 (Addgene plasmid #48674)
dCas9-VPR	unchanged from SP-dCas9-VPR (Addgene Plasmid #63798)
dCas9-AID	synthetic gBlock (amino acids 1-169 from NP_033775.1, murine AICDA)
dCas9-sl-nTet	pEXA2-nTet (synthetic construct from XP_002667965.1)
dCas9-nTet	dCas9-sl-nTet
dCas9-nTet A212T	pEt28a+nTet-A212T and dCas9-nTet
dCas9-nTet A212V	pEt28a+nTet-A212V and dCas9-nTet
dCas9-sl-mTet1	pET-28a-His6-mTet1CD (Ravichandran, 2017)
dCas9-mTet1	dCas9-sl-mTet1
dCas9-sl-mTet2	pET-28a-His6-mTet2CD (Ravichandran, 2017)
dCas9-mTet3	pET-28a-His6-mTet3CD (Ravichandran, 2017) and dCas9-nTet
dCas9-mTet3-del1	pET-28a-His6-mTet3CD-del1 (Ravichandran, 2017) and dCas9-nTet
dCas9-mTet3h (del1-ΔC)	pET-28aHis6-mTet3CD-del1 (Ravichandran, 2017) and dCas9-nTet
VPR-dCas9-mTet3-del1	dCas9-VPR and dCas9-mTet3-del1
VPR-dCas9-mTet3h	dCas9-VPR and dCas9-mTet3h
dCas9-VPR-mTet3-del1	dCas9-VPR and dCas9-mTet3-del1
dCas9-VPR-mTet3h	dCas9-VPR and dCas9-mTet3-del1

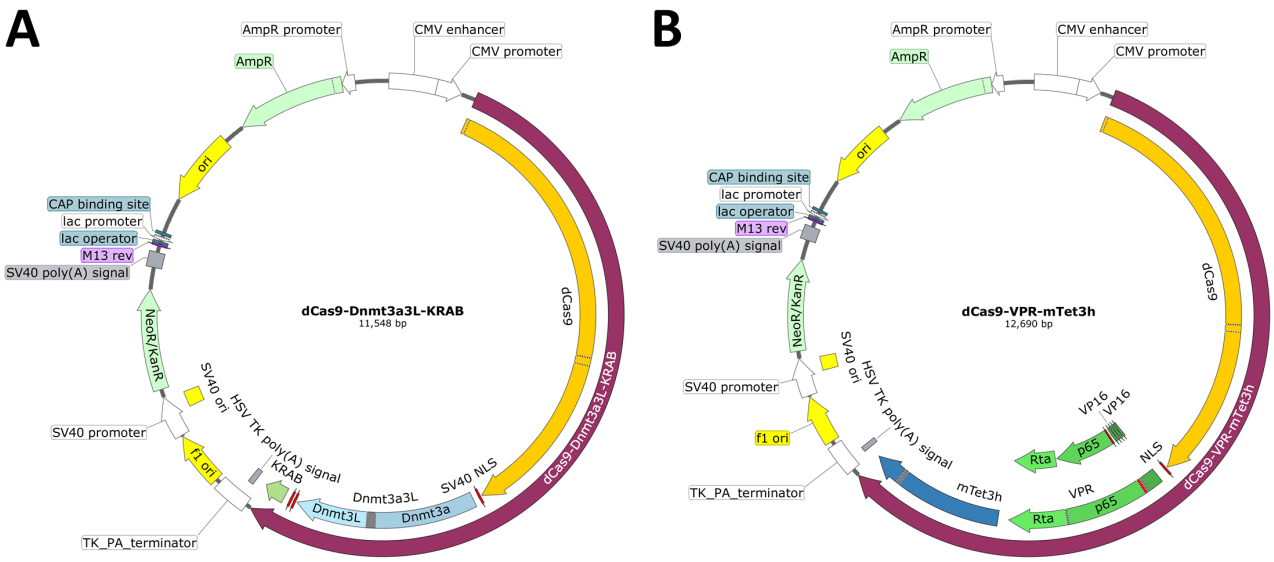

Figure 2.2: Plasmid maps of dCas9-ED plasmids. (A) dCas9-Dnmt3a3L-KRAB; **(B)** dCas9-VPR-mTet3h. Created with SnapGene.

VPR was fused to the N-terminus of dCas9-mTet3-del1 and dCas9-mTet3h and, as another possibility, the mTet3 catalytic domains were cloned to the C-terminus of dCas9-VPR. For the first version, VPR was PCR-amplified and cloned to the XbaI-digested dCas9-mTet3-del1 and dCas9-mTet3h plasmids via NEBuilder HiFi DNA Assembly Master Mix. For the second version, both the mTet3 catalytic domains and the dCas9-VPR plasmid were PCR-amplified and assembled using NEBuilder HiFi DNA Assembly Master Mix. See appendix table 5.4 for the primers used for the PCRs.

The final assembled plasmid maps for the fusions dCas9-Dnmt3a3L-KRAB and dCas9-VPR-mTet3h are shown in figure 2.2.

2.1.2 Generation of gRNA-plasmids

The DNA sequence of the target genes and their promoter regions where extracted using the UCSC genome browser (Kent et al., 2002). Potential binding sites for gRNAs targeting the dCas9-ED fusion proteins to the desired gene promoters were designed with two online services: The E-CRISP platform (Heigwer, Kerr, and Boutros, 2014) was used for gRNAs targeting the *EPCAM*, *CXCR4*, *TFRC*, and *Arx* promoters and the CRISPOR website (Haeussler et al., 2016) for gRNAs targeting *IL6ST*, *ASCL1*, *LIN28A*, and *MAGEB2*. gRNAs were chosen to cover the promoter region and CpG island, if present. Furthermore, gRNAs were selected for high specificity (few off-target binding sites) and high efficiency scores, as available on the CRISPOR service.

Some of the gRNAs were previously validated: ASCL1-sgRNA 2 in this study was used as ASCL1 gRNA -181 by Chavez et al. (2016) and ASCL1-sgRNA 3 was used by Konermann et al. (2015). Several gRNAs targeting the *MAGEB2* promoter in this study were already validated by Xu et al. (2016): MAGEB2-sgRNA 3 in this study corresponds to sgMAGEB2-1,

-sgRNA 4 to sgMAGEB2-3 and MAGEB2-sgRNA 5 to sgMAGEB2-4. Lastly, IL6ST-sgRNA 2 in this study was described as IL6ST-sgRNA 3 by Vojta et al. (2016).

The gRNAs targeting the *EPCAM* promoter were cloned into the gRNA_Cloning Vector (Addgene plasmid #41824, a gift from George Church (Mali et al., 2013)). Overlapping oligos containing the gRNA sequence complementary to the DNA target sequence were annealed after heating them to 95 °C. These DNA fragments were then cloned into the AflII-cut vector using Gibson Assembly.

For all other gRNAs used in this study, ligation cloning was used instead, with a modified gRNA plasmid containing two BbsI cut sites at the gRNA sequence position. For this, the oligos were annealed after heat-treatment as before, but then ligated using T4 DNA ligase (Thermo Fisher) with the BbsI-digested vector. After ligation or Gibson Assembly, the DNA was transformed into electro-competent XL1-Blue MRF' or TOP10 *E. coli* and plasmids isolated from resulting colonies were screened for successful integration by restriction digest using AflII or BbsI, respectively. Correct assembly was verified with Sanger sequencing.

Appendix table 5.5 lists all gRNA binding sites used in this thesis.

2.2 Cell culture

2.2.1 Cell lines and culturing conditions

The main cell lines used in this study were the human ovarian cancer cell line SKOV-3 (a generous gift of Dr. Marianne Rots, University of Groningen) and human embryonic kidney 293 cells (HEK293). The cell lines were maintained at 37 °C with 5% CO_2 in Dulbecco's Modified Eagle's Medium (DMEM) media supplemented with 10% fetal calf serum (FCS), 1× penicillin/streptomycin and 8 mM glutamine (all Sigma-Aldrich) for SKOV-3 cells and 16 mM for HEK293 cells. All microscopy pictures of HEK293 and SKOV-3 cells shown in this thesis were taken using the AMG EVOS fl system.

For passaging or splitting for transfections, the cells were washed with Dulbecco's Phosphate Buffered Saline (DPBS) without calcium chloride and magnesium chloride, detached with Trypsin-EDTA solution (both from Sigma-Aldrich) and seeded in the desired flasks (ranging from six well plates to T175 flasks, from Sarstedt) at dilutions between 1:2 and 1:10.

The murine pancreatic α-cell lines αTC1.6 and αTC1.9 were maintained in a 1:1 mix of DMEM GlutaMAX high glucose (4.5 g/l) and low glucose (1 g/l) (both with pyrovate, Thermo Fisher), supplemented with 10% FCS, 0.02% BSA, and 1x penicillin/streptomycin. The microscopy images of α-cell lines shown in this thesis were taken at 40-fold magnification with an Axiocam digital camera on the Axio Observer Z1 microscope (Carl Zeiss Microscopy).

2.2.2 Transfections

Most of the mammalian cell transfections in these studies were done with MAX PEI (linear, 40 kDa, from Polysciences) at a mass ratio of 3:1 to plasmid DNA as described in Stepper et al. (2017). HEK293 and SKOV-3 cells were seeded in cell culture flasks at usually 40% density. The following day, the plasmid DNA and PEI were mixed in DMEM, incubated for 20 min and then added to the cells. After 12-16 h, the transfected cells were washed with DBPS and immersed in normal growth media. During the course of the experiments, the growth medium was generally exchanged every two days. For more information regarding the transfected DNA, see the result section of the individual experiments.

Reporter plasmids and enrichment of transfected cells

To quickly estimate the transfection efficiency and verify plasmid expression, vectors expressing fluorescent proteins were co-transfected at between 5% to 10% of total DNA: For SKOV-3 cell transfections, mVenus C1 (Addgene plasmid #27794, a gift from Steven Vogel) with a fused NLS (Lungu, 2018) was used, while for HEK293 and αTC1.9 transfections, the original mVenus plasmid without the NLS was transfected. Later experiments exchanged mVenus for pEGFP-puro (Addgene plasmid #45561, a gift from Michael McVoy (Abbate et al., 2001)) to allow the selection of transfected cells with puromycin.

To this end, 2.5 μg/ml puromycin (Thermo Fisher) was added to transfected cells two times: initially around 20 after the transfection, and then again 2 days past transfection (dpt).

As the transfection efficiency for SKOV-3 cells was too low, I had to enrich for transfected cells using magnetic-activated cell sorting (MACS). Therefore, 5% of pMACS LNGFR (Miltenyi Biotec) was included in the transfections and MACS was performed according to the manufacturer's instructions.

Similarly, transfected αTC1.9 cells had to be selected as well. For this, I used fluorescence-activated cell sorting (FACS) with a FACSAria III (BD Biosciences) and the gating strategy shown in figure 3.23B.

Inhibiting and activating small molecule drugs

In several experiments in this thesis, additives were added to cells during transfection experiments to modulate the effects achieved by epigenome targeting. While the general procedures are outlined here, the exact treatment details differed between the experiments and are described again in the results section.

For the time course experiment in section 3.2.1, the inhibitor of alpha-ketoglutarate-dependent dioxygenases dimethyloxalylglycine (DMOG) (Hanson, Rawlins, and Leibold, 2003) was used. DMOG was purchased from Cayman Chemical Company, dissolved as a 1 M solution in DMSO and added to the cells at a concentration of 0.5 to 1 mM, depending on the experiment.

In the experiment in section 3.4.2, I employed two small molecule inhibitors. The inhibitor OICR-9429 disrupts the binding of the MLL-complex member WDR5 to MLL proteins (Grebien et al., 2015). SGC0946, on the other hand, is a selective inhibitor of DOT1L (Yu et al., 2012b). Both drugs were obtained from Sigma-Aldrich and dissolved to 2 mM in DMSO. HEK293 cells were pre-treated with the DOT1L and MLL1 inhibitors starting four days before the transfection at a concentration of 2 μM. During the rest of the experiments, the inhibitors were added to the respective sample with every media change.

For all targeted activation experiments with TET enzymes except for the time course experiment in section 3.6.1, Na-L-ascorbate (VitC, from Sigma-Aldrich) was added to the growth medium to enhance DNA demethylation activity. Ascorbate was used at a concentration of either 25 or 100 μM, depending on the experiment. As cells did not adhere again if ascorbate was added directly after seeding them, it was added 5-14 h later to allow the cells to re-attach (except for the mentioned experiment in section 3.6.1).

Staining of pancreatic α-cells for insulin production

For immunostaining, our collaborator Marija Đordjević (supervised by Dr. Melita Vidaković) seeded transfected αtc1.6 cells on sterile glass cover slips in 24-well plates and fixed them with 4% paraformaldehyde. This was followed by permeabilization with 0.3% Triton X-100 for 10 min and blocking with 3% bovine serum albumin for 1 h. Incubation with the insulin antibody (H-86) (sc-9168, from Santa Cruz Biotechnology) was performed over night at 4 °C with a dilution of 1:50 in PBS containing 0.2% Tween-20. Afterwards, the slides were incubated with the 1:100 diluted Texas Red-conjugated α-rabbit goat antibody sc-3841 (Santa Cruz Biotechnology) for 2 h at room temperature. DAPI staining was performed additionally to visualize DNA.

2.3 Analysis of DNA, RNA and proteins

2.3.1 DNA isolation and sequencing

2.3.1.1 DNA isolation

Initially, genomic DNA was isolated from frozen cell pellets, ranging in cell numbers from tens of thousands of cells (when doing MACS or FACS enrichment) to up to a million cells, using the commercial QIAamp DNA Mini Kit from Qiagen according to the manufacturer's "DNA Purification from Blood or Body Fluids (Spin Protocol)". To test the quality of the isolated gDNA, each sample was quantified using NanoDrop spectroscopy.

Once we had optimized the high-throughput BOMB-methods based on functionalized magnetic beads (see section 3.3), BOMB protocol #7.1[1] was employed for gDNA isolation. Briefly, the cell pellets were lysed in a guanidine thiocyanate (GITC) based buffer after adding RNase A. DNA was then precipitated using isopropanol, bound to silica coated

[1] https://doi.org/10.1371/journal.pbio.3000107.s011

magnetic beads and washed several times with isopropanol and 80% ethanol. After a drying step, the DNA was eluted and stored at -20 °C.

2.3.1.2 Bisulfite conversion of DNA

For site-specific analysis of DNA methylation, around 0.5 µg to 1.0 µg of the isolated DNA was bisulfite converted with Zymo Research's EZ DNA Methylation-Lightning Kit according to the manufacturer's instructions.

After introduction of our optimized BOMB protocols (see section 3.3), BOMB protocol #9.1[1] was used for bisulfite conversion of DNA. Briefly, the DNA strands were melted and bisulfite converted in a sodium metabisulfite based buffer using a thermocycler. The converted DNA was then bound to silica coated magnetic beads with a guanidine-HCl binding buffer, washed and desulfonated with NaOH. After additional washing and drying steps, the bisulfite converted DNA was eluted from the beads and stored at 4 °C until further processing.

2.3.1.3 PCR amplification

Regions of interest were PCR-amplified for each sample for DNA methylation analysis. Usually, these were the promoter regions of the genes targeted with epigenetic effector domains fused to dCas9. PCR was performed on 1/70 to 1/20 of a typical bisulfite conversion elution using Qiagen's HotStarTaq Polymerase with 2.25 mM $MgCl_2$ and 0.033 U/µl polymerase. The touch-down PCR program is detailed in table 2.2 . The annealing temperatures were optimized for each primer pair and ranged from 57 °C to 66 °C for the initial cycles. See appendix table 5.6 for a list of all bisulfite primers.

The PCR products were run on agarose gels for quality control, then combined for each sample and cleaned-up using either the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) or BOMB protocol #4.2 (see section 3.3) with commercial magnetic beads (Sera-Mag SpeedBeads, GE Healthcare). Concentration and purity were measured with NanoDrop.

2.3.1.4 NGS library preparation, sequencing and analysis

The combined bisulfite amplicons for each sample were end-repaired and A-tailed using the components of the SureSelect library preparation kit (Agilent Technologies). The individual samples were then ligated to unique TruSeq HT double indexed adapters, pooled and amplified using PCR for eight cycles with the Q5 polymerase and Kapa-f/r primers (see appendix table 5.1). The library was quantified using the NEBNext Library quantification kit for Illumina (NEB). Clean-ups during the library preparation were performed using magnetic beads as in section 2.3.1.3. The libraries were sequenced on MiSeq machines with

Table 2.2: Touchdown PCR program for amplification of bisulfite converted DNA. The initial annealing temperature "x" was optimized for each primer pair and ranged from 57 °C to 66 °C.

bisulfite PCR program		
95 °C	15 min	
94 °C	25 s	
x °C	30 s	6x
72 °C	40 s	
94 °C	25 s	
x-5 °C	30 s	11x
72 °C	40 s	
72 °C	25 s	
x-10 °C	30 s	36x
72 °C	40 s	
72 °C	4 min	
8 °C	hold	

2x300 PE runs at the Microbiome Core Facility of the UNC School of Medicine (Chapel Hill, USA).

The sequencing results were demultiplexed using Qiime (Caporaso et al., 2010), quality filtered and adapter trimmed with Trim Galore[2] (v0.4.1, using the default parameters with a cut-off at Phred 20 and `--paired`). The sequences were then mapped to the GRCh38 assembly of the human genome using Bismark[3] (v0.14.4; using `--non_directional`). For each CpG site, the percentage of methylation reads were determined with SeqMonk[4] using the "Difference quantification" function and retrieved using the "Annotated probe report" function. The final analysis was performed using Microsoft Excel.

2.3.2 RNA isolation and RT-qPCR

2.3.2.1 RNA isolation

For the first parts of my studies I used the RNeasy Mini Plus Kit (Qiagen) to isolate RNA from frozen pellets of HEK293 cells. For SKOV-3 cells, the Ambion PureLink RNA Mini Kit (Life Technologies) was chosen. For quality control, each sample was quantified and checked for purity using NanoDrop spectroscopy. For additional quality control, samples

[2]Felix Krueger, http://www.bioinformatics.babraham.ac.uk/projects/trim_galore
[3]Felix Krueger, http://www.bioinformatics.babraham.ac.uk/projects/bismark
[4]Simon Andrews, http://www.bioinformatics.babraham.ac.uk/projects/seqmonk

Table 2.3: Thermocycler programs for reverse transcription and qPCR.

Reverse transcription	
42 °C	1 min
25 °C	5 min
37 °C	5 min
42 °C	60 min
90 °C	10 min
8 °C	hold

qPCR and melting curve		
95 °C	3 min	
95 °C	10 s	39x
61 °C	30 s	
plate read		
95 °C	10 s	
65 °C		Δ0.5 °C
5 s, plate read		
95 °C		

were also regularly run on TAE agarose gels after denaturing using a formamide-containing loading buffer and heating to 63 °C for 1 min.

As for DNA isolation and bisulfite conversion, I later switched to our high-throughput protocol for the isolation of RNA based on home-made TRI reagent (BOMB protocol #8.1[1], see section 3.3). Briefly, after cell lysis in TRI reagent, the nucleic acids were bound to silica-coated magnetic beads by the addition of guanidine-HCl and ethanol. After washing with ethanol, the beads were incubated in a DNase I mix to digest DNA. After an additional binding and several washing steps, the beads were dried and the purified RNA was eluted from the beads. RNA samples were stored at -80 °C.

2.3.2.2 RT-qPCR

For the analysis of gene expression of targeted genes via qPCR, the RNA was reverse transcribed into cDNA. Generally, 1 µg of RNA was combined with 0.75 µM dT_{18}-primers, 0.25 µM random octamers, 0.5 µM dNTPs (each), 4 U RNasin Ribonuclease Inhibitor (Promega), 1x RT buffer and 200 U of M-MuLV-RT (obtained from NEB or Enzymatics) in 20 µl. This mix was then incubated in a thermocycler machine with the program in table 2.3.

The generated cDNA was diluted 1:2 and used as the template for quantitative PCR with 0.25 µM primers and SsoFast EvaGreen Supermix in a CFX96 Connect Real-Time detection system (both Bio-Rad) running the program in table 2.3. Technical triplicates were run per primer pair for each sample.

For the analysis of relative transcription levels, the Cq values were extracted for each well. Initially, actin beta (*ACTB*) expression was used to normalize the expression of the genes of interest. However, I later switched to receptor accessory protein 5 (*REEP5*) (Eisenberg

and Levanon, 2013) expression for normalization as this gene exhibited more consistent expression. Appendix table 5.1 contains a list of all qPCR primers.

The Pfaffl-method (Pfaffl, 2001) was used for the calculation of the abundance of the target transcript relative to the reference gene transcript, taking into account the PCR efficiencies of the two primer pairs as calculated from ten-fold dilution series. The relative expression was calculated for each well relative to the mean Cq of both the other primer pair of the same sample, and the averages of both primer pairs for the normalization sample (usually mock transfection). The formula is shown in equation 2.1 (E is the PCR amplification efficiency of the target or reference gene; ΔCq is the difference between the Cq of the normalization sample and the sample of interest). For each sample, the average expression was then calculated by taking the mean of all six values generated as stated above (3 per primer pair).

$$ratio = \frac{(E_{target})^{\Delta Cq_{target}(control-sample)}}{(E_{ref})^{\Delta Cq_{ref}(control-sample)}} \tag{2.1}$$

2.3.3 Western blot

To verify the similar expression levels of the dCas9-ED fusion proteins, HEK293 cells transfected in T25 flasks with 10% mVenus, 20% dCas9-ED and 70% TFRC sg9 were harvested 3 dpt. Cells were lysed using RIPA-buffer with cOmplete Protease inhibitor (Roche) and sonicated for seven cycles with 20 s at 25% amplitude and 30 s break using a Q120AM sonicator (ActiveMotif). The sonicated samples were centrifuged at 4 °C at 22,000 g for 20 min. The protein content of the supernatant was determined with Bradford Ultra (Expedeon) normalized to BSA (AppliChem) dilutions. 8 µg protein of each sample were loaded on 4% to 12% gradient gels (Expedeon) in RunBlue LDS sample buffer (Expedeon) after boiling at 85 °C for 5 min and centrifugation.

Proteins were then blotted in 192 mM glycine, 25 mM Tris NaCl, 20% methanol, 0.02% SDS onto a nitrocellulose membrane (GE Healthcare). After blocking over night with 5% skimmed milk in PBST (137 mM NaCl, 2.7 mM KCl, 10 mM NaH_2PO_4, 2 mM K_2HPO_4, pH7.4) , the membrane was cut in half horizontally at the 70 kDa band of the PageRuler Plus Prestained (Thermo Fisher) marker. The top half was incubated with anti-Cas9 (ActiveMotif #61578, 1:5000) and the bottom half with anti-ACTB (ab8227, abcam) for 1 h. After washing with PBST, the membranes were incubated for 2 h with anti-mouse (GE Healthcare NA931V) and anti-rabbit (GE Healthcare NA934) for the top and bottom halves, respectively. The horseradish peroxidase-linked antibodies were visualized using 1:10 diluted SuperSignal West Femto (Thermo Scientific) using the Fusion-SL3500 system (Peqlab).

3 Results

In my MSc thesis (Stepper, 2015), I demonstrated targeted DNA methylation using the nuclease-null version of the novel CRISPR-Cas9 system as the targeting device. Thanks to the use of the single-chain fusion Dnmt3a3L (Siddique et al., 2013), I achieved strong and widespread DNA methylation of the targeted *EPCAM* promoter in human SKOV-3 cells, leading to target gene repression. Targeting dCas9-Dnmt3a3L with only a single gRNA lead to a peak of introduced DNA methylation next to the binding site, but also to an increase of DNA methylation over the whole *EPCAM* promoter. These preliminary studies confirmed that the CRISPR-Cas9 system is suitable for epigenetic editing.

For my PhD thesis, I set out to build upon my previous work to develop a versatile toolbox for epigenetic editing with CRISPR-Cas9, for both targeted gene repression and activation. I then used this toolbox to dissect the epigenetic code with an emphasis on the regulation of DNA methylation maintenance. Furthermore, I applied the technology to study tumor formation and potential therapeutic applications.

3.1 Efficient targeted DNA methylation with a chimeric dCas9–Dnmt3a–Dnmt3L methyltransferase

The CRISPR-Cas9 system makes it possible to easily target a different promoter or multiple at once by simply changing the gRNA sequence. Analyzing the effects of targeted DNA methylation on multiple target genes is important, as the precise epigenetic state of each gene might lead to different responses. For this expansion in scope, I used the lessons I learned during my MSc thesis: Instead of choosing new targets in the SKOV-3 cells, which are difficult to transfect, I targeted DNA methylation to *TFRC* and *CXCR4* (see section 1.4.1) in HEK293 cells. This allowed me to skip the cumbersome MACS procedure that was necessary to select for transfected SKOV-3 cells, as HEK293 are much easier to transfect with high efficiencies. Figure 3.1 shows a comparison of the transfection efficiencies for SKOV-3 and HEK293 cells, and a flow cytometry (FC) analysis of MACS enrichment.

Another lesson was to switch from subcloned bisulfite amplicon sequencing via colony picking, plasmid isolation and Sanger sequencing, to index-based multiplexed libraries and Illumina MiSeq 2x300 bp paired-end sequencing. This vastly improved the amount and quality of DNA methylation data I was able to obtain. Furthermore, these changes allowed me to increase sample numbers significantly and to address more experimental questions. Firstly, does the response of the target gene differ between genes and cell lines? Secondly,

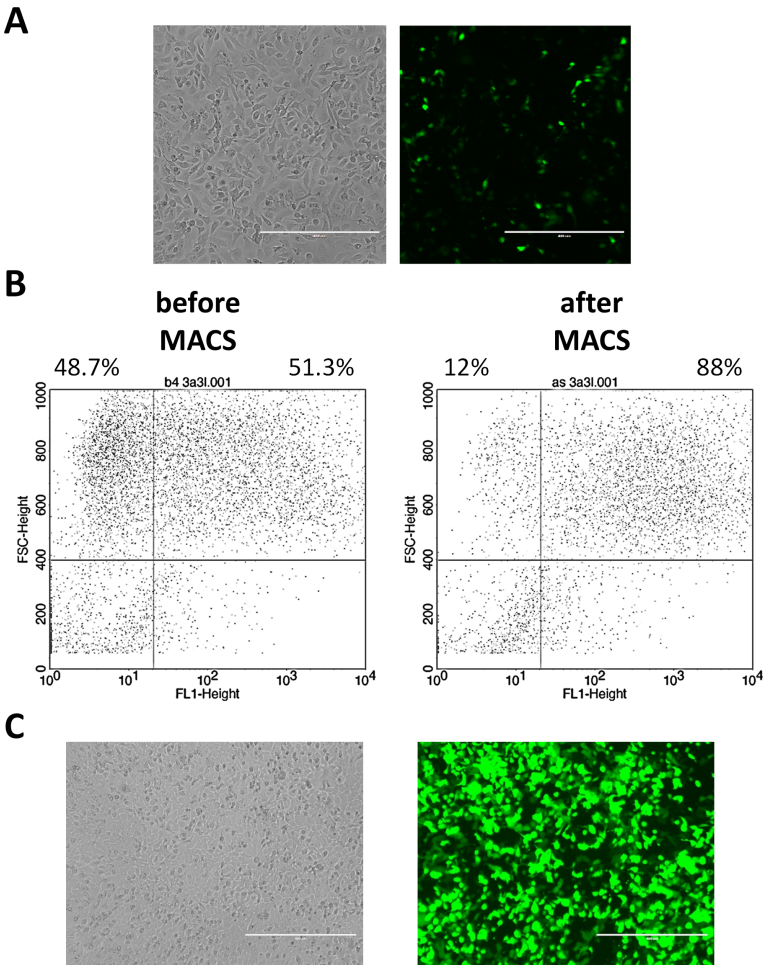

Figure 3.1: Typical transfection results in SKOV-3 and HEK293 cells. (A) Because the transfection efficiency in SKOV-3 cells was too low as seen here in a comparison between bright field and fluorescence microscopy, the transfected cells were enriched using the MACS method. This boosted the fraction of transfected/fluorescent cells from 51.3% to 88%, as assayed by FC **(B)**. In contrast, transfections of HEK293 cells resulted in a high fraction of fluorescent cells **(C)**, making enrichment unnecessary.

are the effects of targeting DNA methylation with the most effective gRNAs additive? And thirdly, how does Dnmt3a3L compare to just Dnmt3a or mutated versions of the effector?

3.1.1 Targeting promoters with dCas9-Dnmt3a3L leads to widespread DNA methylation

EPCAM in SKOV-3 cells

I targeted the established *EPCAM* promoter in SKOV-3 cells to test the system and analyze the effects of dCas9-Dnmt3a3L at this locus. Therefore, the cells were transfected using PEI with a mixture of 5% mVenus for transfection control, 5% pLNGFR for MACS enrichment, 20% of dCas9-Dnmt3a3L and 70% of either a single gRNA plasmid or an equimolar mix of multiple gRNAs. Five days after the transfection, transfected cells were harvested following MACS enrichment. Genomic DNA was isolated from the harvested and frozen samples and bisulfite converted. I then amplified the *EPCAM* promoter locus for DNA methylation

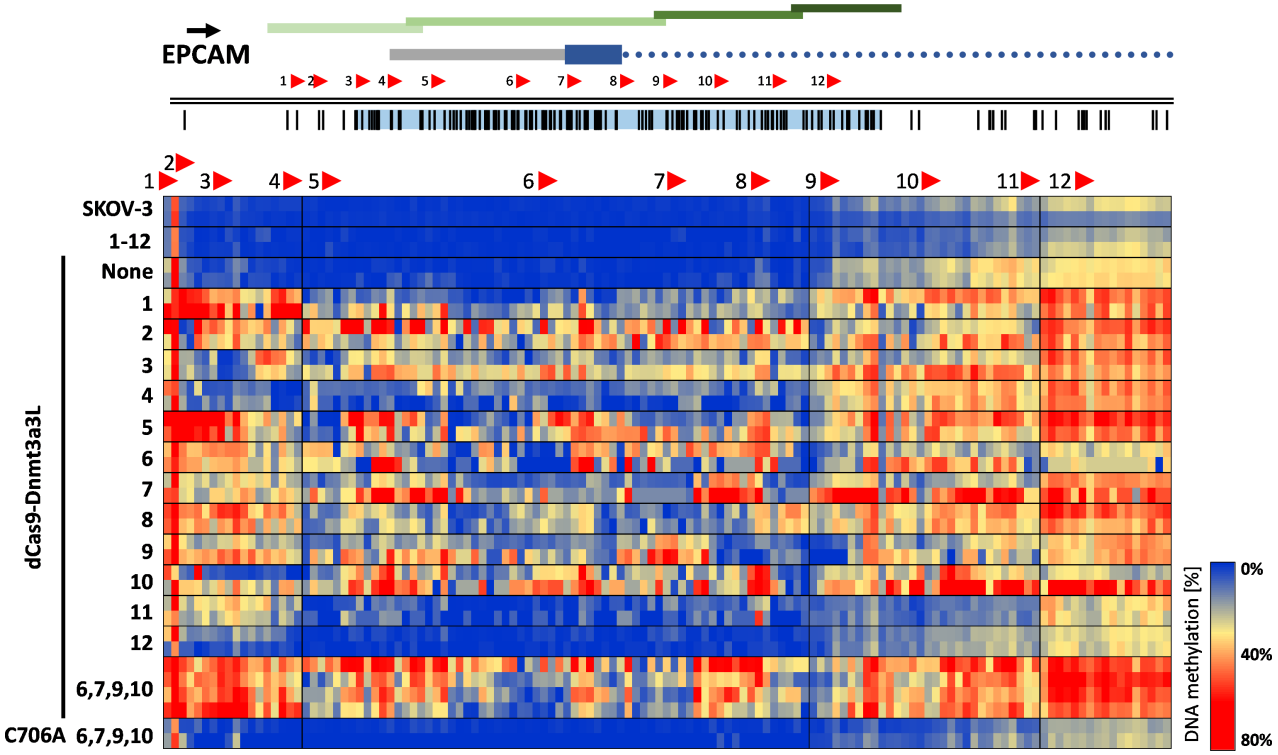

Figure 3.2: Targeted DNA methylation at the endogenous *EPCAM* promoter in human SKOV-3 cells (Stepper et al., 2017). The genomic locus pictured at the top (Chr. 2: 47368894–47370157) depicts this promoter, which was targeted by the fusion protein dCas9-Dnmt3a3L using a total of 12 gRNAs, whose location and orientation are shown as red arrows. Methylation was investigated in four consecutive amplicons covering the CpG island (indicated with green bars). The CpG island itself is represented as a light blue box, encompassing most of the CpG sites in the region (depicted as black vertical lines). The 5' UTR of the *EPCAM* gene is shown as a grey bar, with the first exon in blue and the following intron indicated with a blue dotted line. The average DNA methylation of each CpG site is represented in a heat map according to the colors in the legend shown on the right. Each row shows the result of one biological repeat of the experiments.

analysis and prepared the samples for sequencing as an indexed library. The analysis of this sequencing data of SKOV-3 cells confirmed the results of my MSc thesis (Stepper, 2015). The analyzed promoter region of *EPCAM* was mostly unmethylated for untreated SKOV-3 cells, except for the downstream part of the CpG-island towards the gene body, as is characteristic for expressed genes. In contrast, targeting the dCas9-Dnmt3a3L construct with either single gRNAs or a mix of the most effective (6,7,9 and 10) resulted in the introduction of DNA methylation (see figure 3.2). Interestingly, the DNA methylation was introduced not only next to the gRNA binding site where the Dnmt3a3L enzyme fused to dCas9 can directly approach the DNA, but over the complete promoter region. Furthermore, the CpG sites covered by dCas9 bound to the complementary gRNA were almost completely protected from the introduced DNA methylation. This suggests that dCas9 stayed bound to the target site very strongly and for long periods of time. Except for sgRNA 12, the targeted DNA methylation was introduced efficiently right next to the gRNA binding site and in the direction in which the Dnmts were presented, as seen in the heat map analysis in figure 3.6.

Both the heat map analysis in figure 3.2 and the average DNA methylation over the complete analyzed region in figure 3.3A demonstrated that some of the single gRNAs were

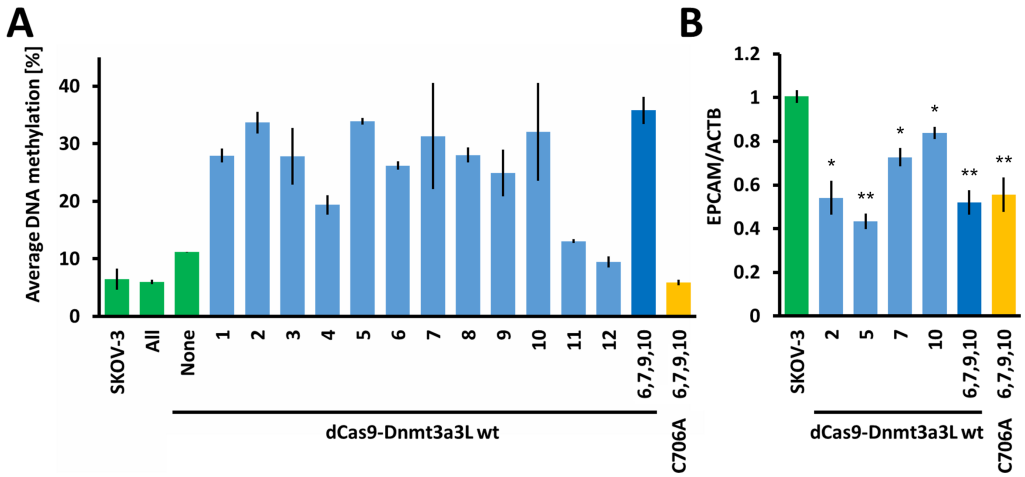

Figure 3.3: Influence of epigenome editing on DNA methylation and *EPCAM* expression (Stepper et al., 2017). Average DNA methylation of all 131 analyzed CpG sites of the *EPCAM* promoter after targeting DNA methylation **(A)** and corresponding changes in gene expression **(B)**. The numbers below the bars denote the gRNAs that were used to target the dCas9–Dnmt3a3L. Green bars highlight control experiments. The error-bars represent the SEM from at least two biological repeats performed in technical triplicate). *P < 0.05, **P < 0.0001, unpaired t-test relative to the 'SKOV-3' control sample.

more effective in depositing methylation than others. The highest average DNA methylation resulted from targeting dCas9-Dnmt3a3L with either only sgRNA 5 (33.9%±0.5%) or the combination of sgRNAs 6, 7, 9, and 10 together (35.8%±2.3% DNA methylation over the whole analyzed region). Interestingly, targeting dCas9-Dnmt3a3L with gRNA 4 resulted in less DNA methylation compared to its neighboring gRNAs. Analyzing the promoter structure revealed that the gRNA binding site overlaps with the annotated TSS, where the transcriptional machinery might compete with dCas9-Dnmta3L for binding.

When targeting the EPCAM promoter, gRNAs 2, 5 and the combination of the four select gRNAs also achieved the strongest downregulation of *EPCAM* expression, decreasing the expression to only 0.43±0.04-times the expression of untreated cells in the case of sgRNA 5 (see figure 3.3B). However, even targeting the C706A mutant led to a significant reduction in expression. This indicates that either sterical effects (blocking transcriptional machinery) or other effects of the transfection itself. These effects could also be in part responsible for the observed repression with the WT Dnmt3a3L.

TFRC and CXCR4 in HEK293 cells

The subsequent experiments were performed in HEK293 cells, as the transfection efficiency of around 80% (see figure 3.1C) was high enough to skip the MACS enrichment steps. Apart from the missing MACS enrichment, the experiments targeting of targeting dCas9-Dnmt3a3L with both single and multiple gRNAs to *TFRC* and *CXCR4* in HEK293 were performed similarly as before. However, the pLNGFR plasmid was omitted from the transfections as no MACS was performed; I instead used 10% of the total transfected DNA for the mVenus plasmid.

The effects on promoter DNA methylation are shown in figure 3.4A and B. The assayed

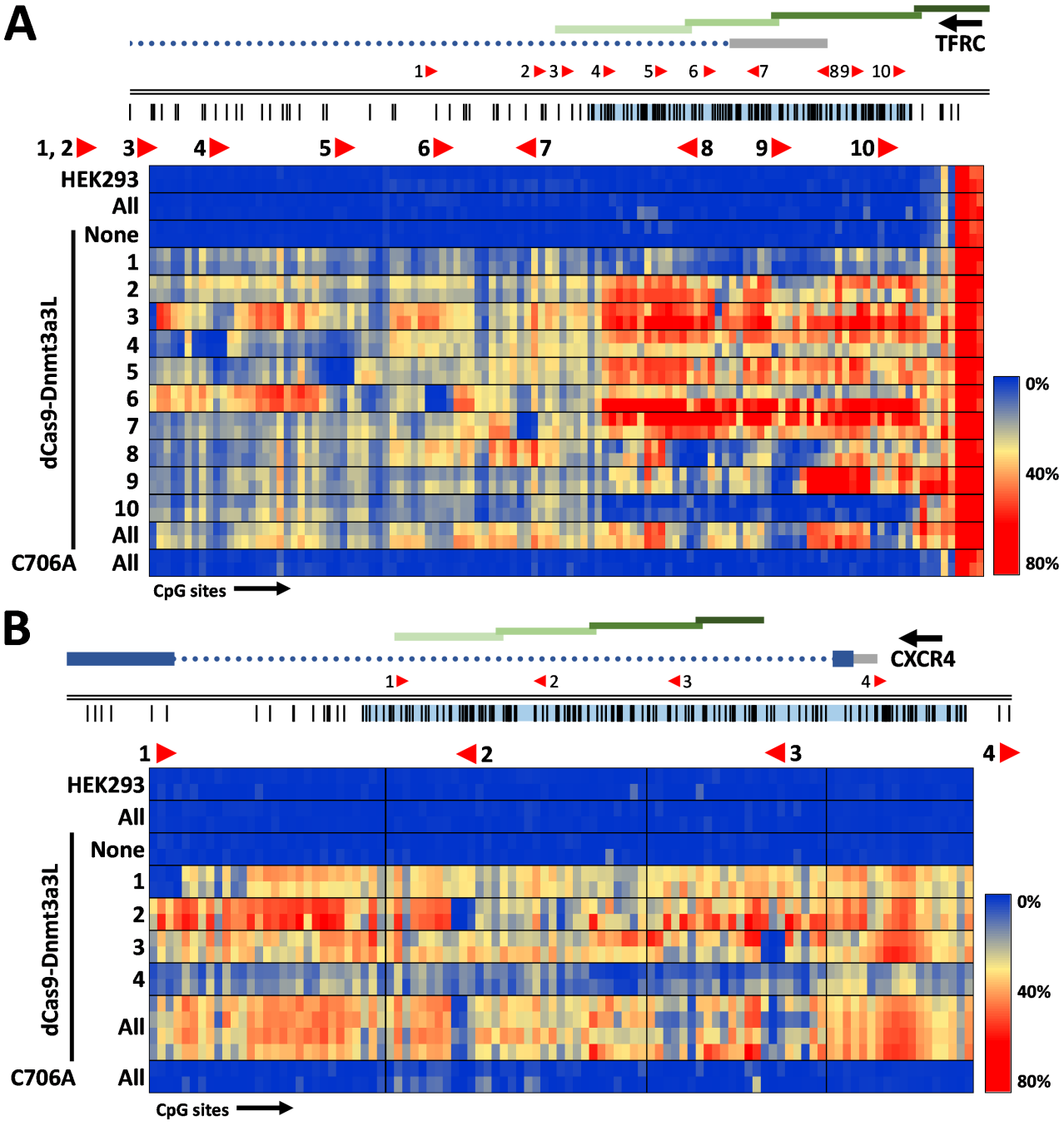

Figure 3.4: Effects on DNA methylation of targeting dCas9-Dnmt3a3L to the promoters of **(A)** *TFRC* (Chr. 3: 196081310–196082794) and **(B)** *CXCR4* (Chr.2: 136116618–136117805) in HEK293 cells. The average DNA methylation level of each CpG site is represented as a heat map. See figure 3.2 for a more detailed description. (Stepper et al., 2017).

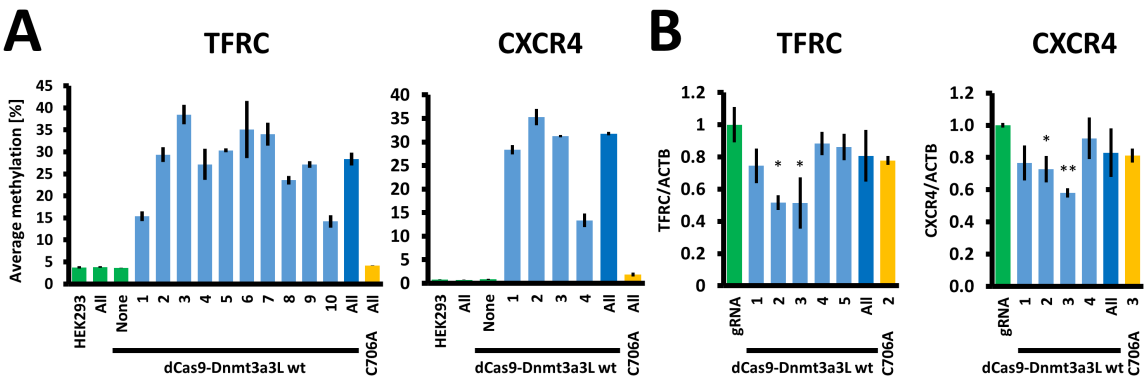

Figure 3.5: Influence of epigenome editing on DNA methylation and expression of *TFRC* and *CXCR4* (Stepper et al., 2017). **(A)** Average DNA methylation of all analyzed CpG sites at the *TFRC* and *CXCR4* promoters for the samples as denoted below the bars and **(B)** change in gene expression. See figure 3.3 for a more detailed description.

promoter regions were unmethylated in untreated HEK293 cells, except for a few CpG-sites at the CpG-island shores of the 5'-UTR of *TFRC*. This was also the case when transfecting only the gRNAs or dCas9-Dnmt3a3L without the gRNAs ('untargeted'). When including a single gRNA, DNA methylation was deposited with all the tested gRNAs to levels between 13.3%±1.5% and 38.5%±2.2% (see figure 3.5A). Interestingly, this DNA methylation often covered the whole locus; even more so than for *EPCAM* in SKOV-3 cells. As at the *EPCAM* promoter, the CpG sites in gRNA binding sites covered by dCas9 showed almost no DNA methylation; and the introduced DNA methylation usually peaked right downstream of the target PAM site.

Targeting dCas9-Dnmt3a3L to *TFRC* and *CXCR4* with mixes of all gRNAs lead to strong DNA methylation, reaching 28.3%±1.5% for *TFRC* and 31.8%±0.4% for *CXCR4*. However, the best single gRNAs were at least as good at introducing DNA methylation and achieved up to 38.5%±2.2% for *TFRC* with gRNA 3 and 35.3%±1.8% for *CXCR4* with sgRNA 2. Comparing the distribution of DNA methylation over the whole locus, the samples targeted with a mix of all gRNAs set DNA methylation less localized and more evenly spread across the target region. As expected, the catalytically dead mutant of Dnmt3a3L, C706A, did not introduce DNA methylation when targeted to the promoters.

For *TFRC*, targeting dCas9-Dnmt3a3L with sgRNA 2 led to a reduction in *TFRC* expression to 0.52±0.05 relative to the gRNAs only control(see figure 3.5B). Similarly, sgRNA 3 achieved a repression of TFRC expression to 0.51±0.16-times. The repression was less pronounced for the other single gRNAs, including the sample with the mix of all *TFRC* gRNAs, which did not reach a significant level of target genes repression. At the *CXCR4* promoter, targeting dCas9-Dnmt3a3L with gRNAs 2 or 3 resulted in significant target gene repression to 0.73±0.08 and 0.58±0.03, respectively. Like for *TFRC*, targeting all four gRNA binding sites simultaneously did not result in a downregulation of *CXCR4*. For both target genes in the HEK293 cells, the C706A mutant of Dnmt3a3L resulted in neither DNA methylation nor repression, even when targeted with the most effective gRNAs.

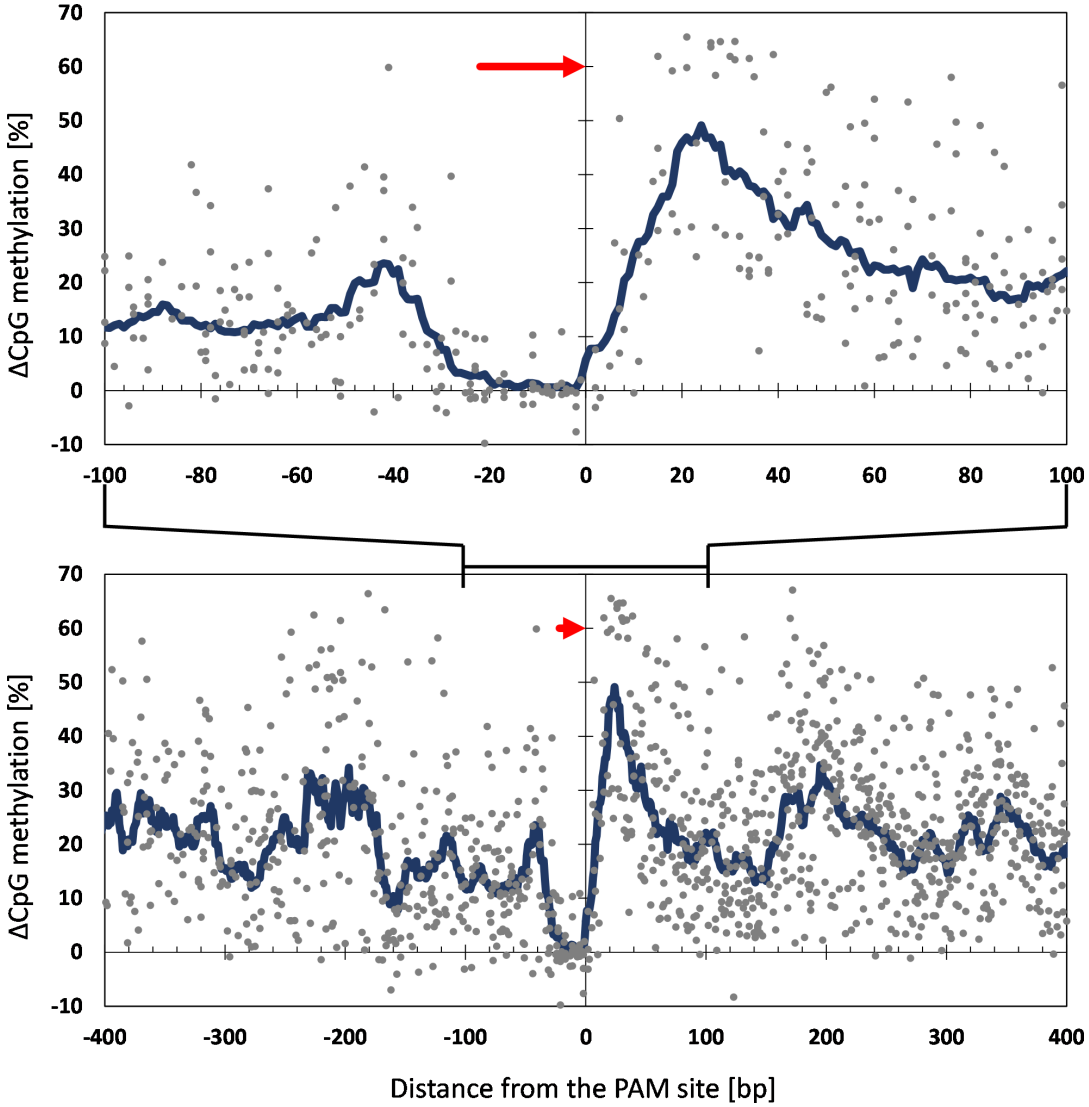

Figure 3.6: Introduced DNA methylation relative to the PAM site of the gRNA binding site (Stepper et al., 2017). This scatter plot shows the DNA methylation of each sequenced CpG site (grey dots) of all experiments where dCas9-Dnmt3a3L was targeted with a single gRNA. The y-axis shows the difference in DNA methylation with the methylation in the untreated cells subtracted. The gRNA orientation and binding site is represented by the red arrow, whereas the blue line represents the average introduced DNA methylation in a 15 bp sliding window.

Distribution of targeted DNA methylation

As seen in the heat map visualizations in figures 3.2 and 3.4, DNA methylation was introduced most strongly at the CpG sites right downstream of the gRNA binding site in the direction of the PAM site, which was itself protected from DNA methylation. This observation was confirmed by an analysis of all CpG sites when targeted dCas9-Dnmt3a3L with a single gRNA relative to that gRNA binding site (see figure 3.6). For this analysis, the DNA methylation present in untreated cells was subtracted from the edited cells, yielding only the effect of targeting dCas9-Dnmt3a3L (ΔCpG methylation). A moving average of DNA methylation of 15 bp revealed patterns in the strength of targeted DNA methylation. Firstly, the 20 to 30 bp upstream of the PAM site were almost completely protected. Secondly, the strongest increase in DNA methylation was observed downstream of the gRNA binding site,

with a peak at around 25 bp downstream. Additionally, both down- and upstream of these initial peaks, the average introduced DNA methylation decreased again. However, at longer distances a potential periodic pattern emerged, with methylation peaks and valleys spaced by about 170 bp. This might suggest effects of nucleosome positioning, as nucleosomal DNA could be protected from acquiring DNA methylation.

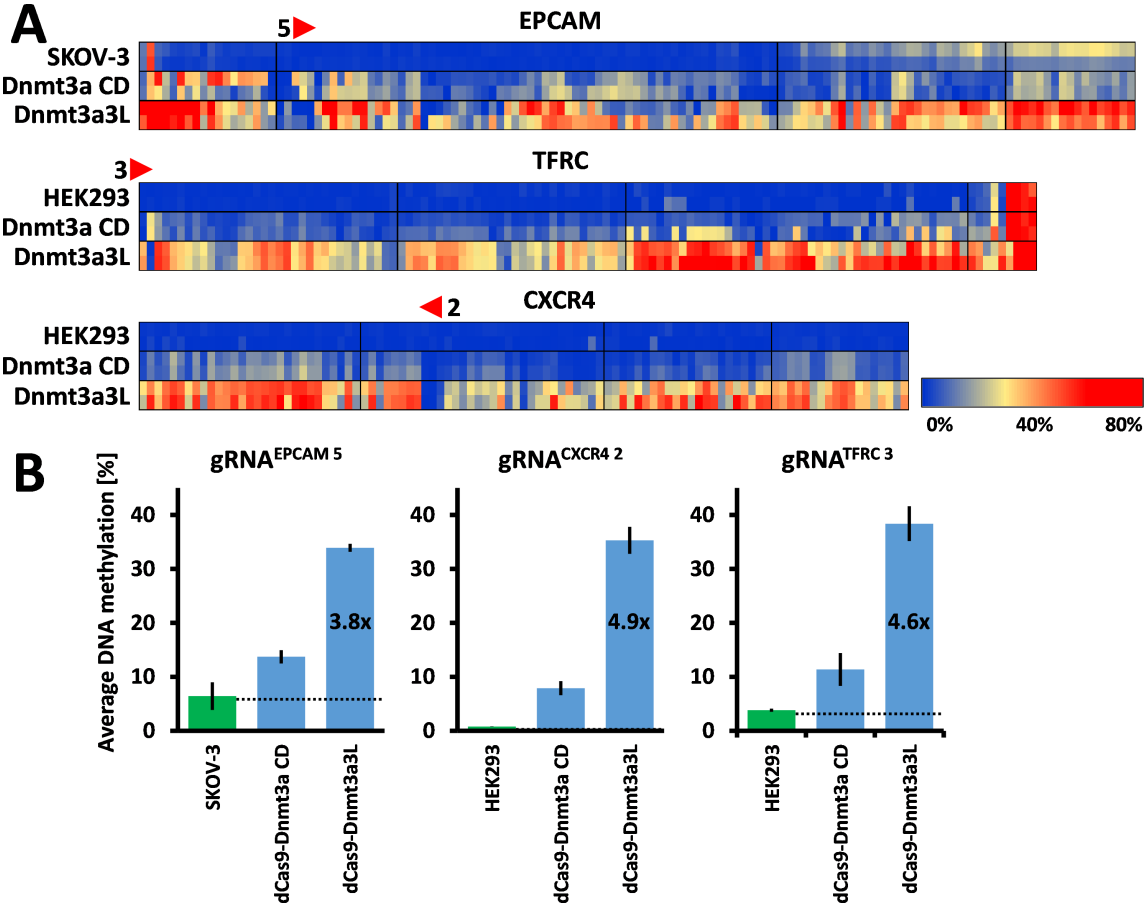

Figure 3.7: Comparison of targeting DNA methylation efficiency by dCas9-Dnmt3a and dCas9-Dnmt3a3L (Stepper et al., 2017). **(A)** Average DNA methylation per CpG site for single gRNA targeting experiments to the promoters of *EPCAM*, *TFRC* and *CXCR4* (see figure 3.2 for a more detailed description). **(B)** Averaged DNA methylation over all analyzed CpG sites for each experiment with the fold increase in methylation efficiency calculated after subtracting the background methylation level already present in the untreated cells.

dCas9-Dnmt3a3L vs dCas9-Dnmt3a

Furthermore, I compared the effect of targeting dCas9-Dnmt3a3L to only the catalytic domain of dCas9-Dnmt3a, without the activity-enhancing Dnmt3L domain. The results are shown in figure 3.7. This comparison demonstrated that dCas9-Dnmt3a3L was 3.8 to 4.9 times more efficient at depositing DNA methylation when targeted to the promoters of *EPCAM*, *TFRC* and *CXCR4* with these selected gRNAs. Differences in dCas9 fusion protein expression can be excluded as an explanation for these differences, as Western blotting revealed similar protein expression levels for dCas9-Dnmt3a and dCas9-Dnmt3a3L (see figure 3.8). Interestingly, targeting dCas9-Dnmt3a lead to widespread DNA methylation as well, but not to the same extent as for the fusion protein dCas9-Dnmt3a3L.

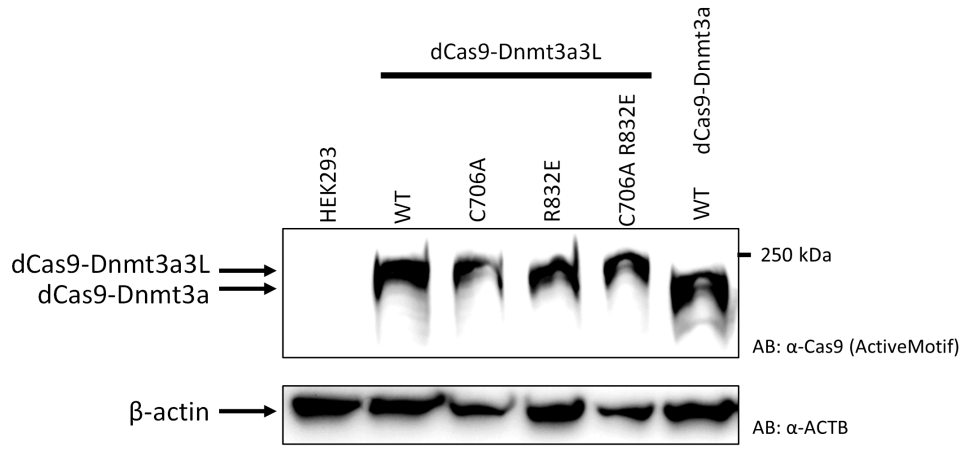

Figure 3.8: Western blotting demonstrated similar protein levels of the transfected dCas9-ED fusion proteins (Stepper et al., 2017). Expected protein masses are 224 kDa for dCas9-Dnmt3a3L and 197 kDa for dCas9-Dnmt3a. The primary antibody was anti-Cas9 (ActiveMotif #61578, 1:5000 dilution for all antibodies), anti-mouse (GE Healthcare Life Sciences NA931V) was used as the as secondary antibody. For the actin loading control, anti-ACTB (ab8227, abcam) and anti-rabbit (GE Healthcare NA934) were used.

3.1.2 DNA methylation spreads via Dnmt3a3L fiber formation

In addition to the catalytically dead mutant of Dnmt3a3L, C706A, I also targeted another mutant which disrupts the binding interface necessary for multimerization of multiple Dnmt3a units, R832E (Rajavelu et al., 2012). Interestingly, targeting this mutant of dCas9-Dnmt3a3L with a combination of the best gRNAs to the promoters of *EPCAM* (in SKOV-3 cells) and *TFRC* and *CXCR4* (in HEK293 cells) resulted in drastically lower methylation levels (see figure 3.9) than when targeting the WT protein. Strikingly, the R832E mutant was not able to direct DNA methylation further than close to its target site. In contrast to this, the WT Dnmt3a3L introduced DNA methylation not only next to the gRNA binding site, but all over the analyzed region. The difference was especially obvious for the promoters of *TFRC* and *EPCAM*. Nevertheless, gRNAs 1 and 2 targeting the R832E mutant to the *CXCR4* promoter resulted in some DNA methylation occurring along the DNA. Furthermore, the mutant deposited lower levels of DNA methylation even right next to the binding site when compared to WT dCas9-Dnmt3a3L. This result can be partially explained by the fact that the R832E mutation demonstrated a decreased catalytic activity by a factor of ~2 *in vitro* (Rajavelu et al., 2012). Differences in dCas9 fusion protein expression were similar as tested by Western blotting (see figure 3.8) and were therefore not responsible for the observed difference. Taken together, these results together with previous biochemistry data (Emperle et al., 2014) showed that the multimerization of multiple units of Dnmt3a3L, forming a fiber along the DNA, was necessary to achieve the highly efficient and widespread DNA methylation observed in these experiments.

However, it could also be argued that targeting the inactive mutant C706A of Dnmt3a3L could recruit active endogenous Dnmt3a protein. This would also result in fiber formation and spreading DNA methylation along the DNA, even if not right next to the gRNA binding

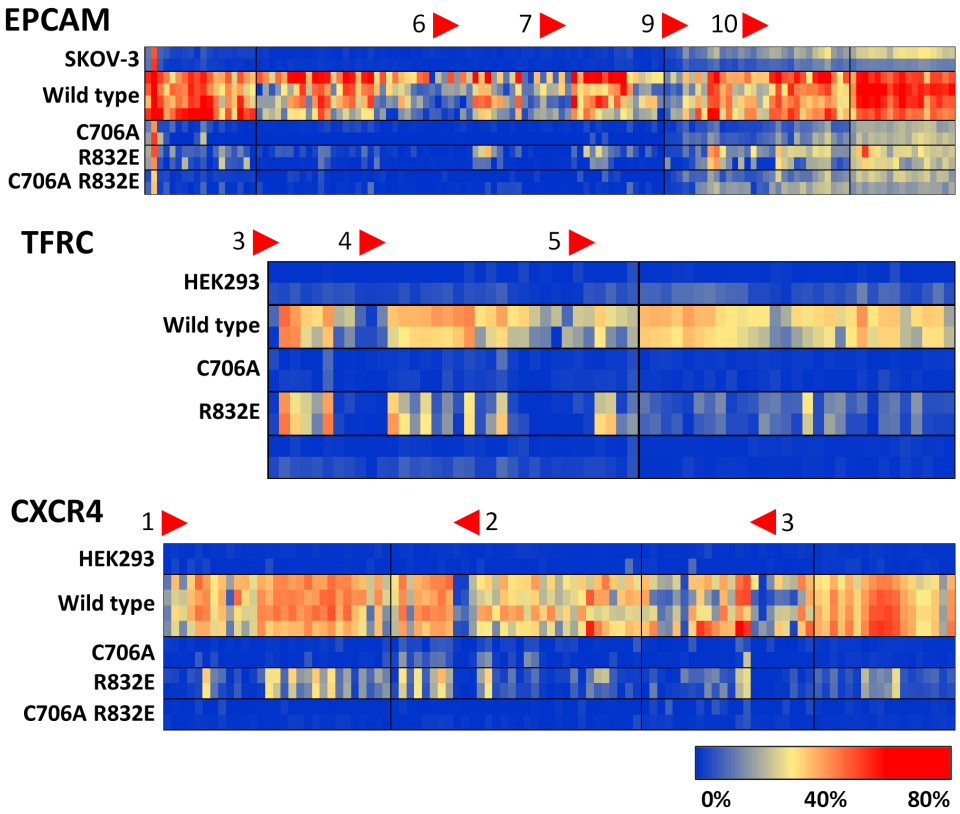

Figure 3.9: Spreading of DNA methylation along the promoter is dependent on the multimerization of Dnmt3a3L (Stepper et al., 2017). DNA methylation levels displayed as a heat map (see figure 3.2 for a detailed explanation) comparing the ability of Dnmt3a3L with its mutants to methylate the DNA of the promoters of *EPCAM*, *TFRC* and *CXCR4*.

site. However, this was not observed, potentially because the transfected cCas9-Dnmt3a mutant outnumbered the active endogenous Dnmt3a units. As expected, the double mutant dCas9-Dnmta3a3L C706A, R382E did not lead to the introduction of any DNA methylation when targeted to these loci.

3.1.3 Targeted DNA methylation leads to limited off-target effects

Even though the gRNAs used in this study were both designed and selected for high specificity and few off-target binding sites, it was still a possibility that dCas9-ED fusion proteins might bind and modify DNA at other sites than intended. To check for potential off-target effects, additional potential binding sites for four gRNAs (TFRC sg2 and 3, CXCR4 sg2 and 4) were predicted using the CRISPOR tool (Haeussler et al., 2016). The four highest scoring off-target sites (according to the MIT or CFD scores) were selected for analysis via bisulfite sequencing. The PCR primers used for amplification are listed in the appendix table 5.6. As shown in figure 3.10A, cells transfected with dCas9-Dnmt3a3L and the respective gRNA showed increased DNA methylation in these short and selected amplicons for the high-scoring off-target binding sites of TFRC gRNAs 2 and 3. The 9.2 and 11.1 percentage points (pp) increase in DNA methylation at off-target sites were much lower compared to an increase of 25.6 and 34.7 pp at the intended target sites for TFRC gRNAs 2 and 3,

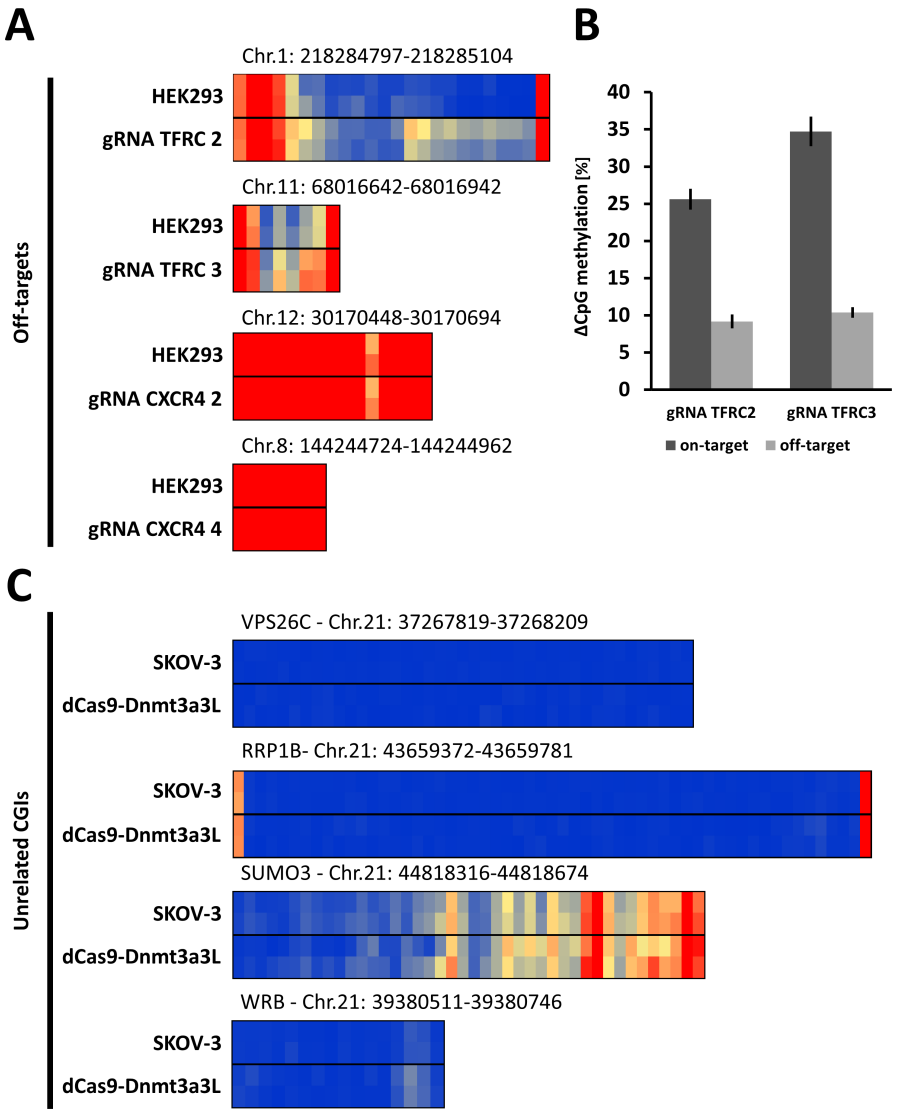

Figure 3.10: Analysis of potential off-target effects of targeted DNA methylation by dCas9-Dnmt3a3L (Stepper et al., 2017). **(A)** Top off-target sites were selected for four gRNAs and analyzed for increased DNA methylation when transfected. **(B)** These sites exhibited an increase on DNA methylation of 9.2% and 11.1%, respectively (light grey), compared to 25.6% and 34.7% at the intended target sites (dark grey). The off-target sites for CXCR4 gRNAs 2 and 3 were already fully methylated. **(C)** The four analyzed promoters with CpG sites but without off-target sites did not show an increase in DNA methylation.

respectively, as shown in figure 3.10B. Therefore, the effect size at the indented target sites was still around three times higher. As the two off-target sites for the CXCR4 gRNAs were already almost fully methylated, they exhibited no increase in methylation.

To specifically analyze the potential increase of DNA methylation at other untargeted promoters with CpG islands, four amplicons in CpG islands were chosen and analyzed. As can be seen in figure 3.10C, these CpG islands did not show increased DNA methylation levels. This indicates that unbound dCas9-Dnmt3a3L does not lead to a general increase in DNA methylation levels genome wide.

3.2 Stability of targeted DNA methylation

3.2.1 Targeted DNA methylation is lost over time

DNA methylation is generally passed on faithfully to the daughter cells after cell division. However, it was unclear whether this was also the case for transfected and edited cells in the absence of the targeted dCas9-Dnmt3a3L fusion construct. Therefore, I probed the development of DNA methylation levels at the *TFRC* promoter at multiple time points ranging from 2 dpt to 16 dpt.

Additionally, this experiment was performed with the addition of the dioxygenase inhibitor DMOG to half of the samples. As DMOG inhibits TET enzymes, this should make it possible to discern whether the potential loss of the introduced DNA methylation is caused by active demethylation via the TET-enzyme dependent pathway. The inhibitor was added to one set of samples starting on the first day after transfection at a concentration of 1 mM. The other set served as a control with 1:1000 DMSO, as DMOG was dissolved in DMSO before adding it to the cell culture media.

The results of this experiment targeting the *TFRC* promoter with dCas9-Dnmt3a3L and TFRC gRNA 2 in figure 3.11 showed that the introduced DNA methylation increased from 3.7%$\pm$0.1% in untreated HEK293 cells to 16.7%$\pm$1.7% for the DMSO-treated samples and 16.1%$\pm$0.6% for the DMOG-treated samples 5 dpt. But after this initial peak, the level of DNA methylation dropped back almost to the initial level within just three days. As the speed of setting and losing DNA methylation and its location were the same between the DMOG and DMSO-treated samples, these results suggested that TET enzymes are not the main drivers responsible for the loss or removal of the introduced DNA methylation in this cell type.

3.2.2 Combining Dnmt3a3L with KRAB leads to stronger target gene repression

The previous experiment in section 3.2.1 showed that the targeted DNA methylation was lost quickly over the span of only 3-5 days. We then wanted to firstly, improve the stability

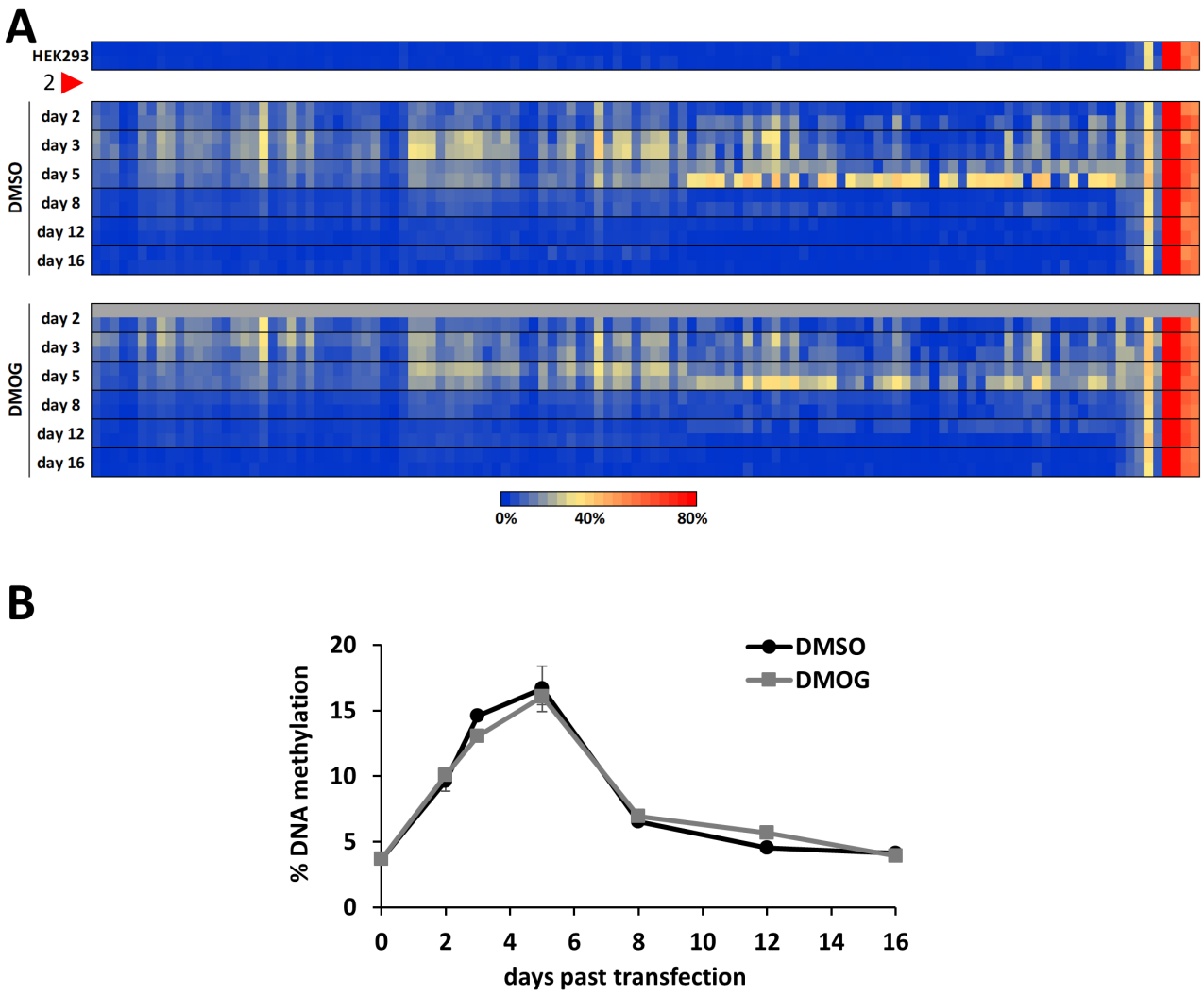

Figure 3.11: The introduced DNA methylation was not maintained over time in the cells. (A) Heat map results for a time course targeting dCas9-Dnmt3a3L with TFRC sg2, which binds 75 bp upstream of the first CpG site of the analyzed region. No data was available for the first biological replicate of the DMOG-treated cells 2 dpt (colored in grey). See figure 3.2 for a more detailed explanation of the heat map. **(B)** The introduced DNA methylation reached a peak five days after the transient transfection, but then dropped quickly to the initial level at the last analyzed time point 16 days past transfection. The TET inhibitor DMOG made no difference regarding the stability of the targeted DNA methylation. Error bars denote SEM of the average DNA methylation of the analyzed region for the two biological replicates.

of the introduced mark and target gene repression, and secondly, study the mechanisms responsible for the fast removal or loss of DNA methylation.

To this end, I planned to replicate these results with more samples and target genes and study the influence of the complete epigenetic state of the target promoter and how this state changes when one mark is edited. As DNA methylation is just one layer of epigenetic control, histone modifications are also present and will influence both the initial and long-term outcome of epigenome editing. Furthermore, I wanted to identify which factors are necessary to achieve a complete and stable switch of the epigenetic and transcriptional state of the target gene. One possible avenue to test this is to combine the effects of multiple effector domains (EDs) working via different pathways to edit the epigenetic state on multiple levels.

My student Roman Zeiger and I cloned the Krüppel associated box (KRAB) effector domain to dCas9-Dnmt3a3L. The repressive effects of KRAB work mainly via the recruitment of complexes responsible for the methylation of H3K9 (Schultz et al., 2002) and deacetylation of H3K27ac (Schultz, Friedman, and Rauscher, 2001). In conjunction with the targeted DNA methylation introduced by Dnmt3a3L, these effects should lead to stronger and more sustained target gene repression. We constructed plasmids fusing KRAB to either the N- or C-terminal end of dCas9-Dnmt3a3L. Preliminary experiments showed that the C-terminal fusion dCas9-Dnmt3a3L-KRAB produced stronger target gene silencing than the N-terminal fusion protein (data not shown).

To further analyze the efficacy of these new fusion proteins and the kinetics and mechanisms of DNA methylation loss, we transfected HEK293 cells in T25 flasks with mVenus, the dCas9-ED fusions and the plasmid encoding the CXCR4 sg3. The samples were harvested 3, 5, 8 and 10 days past transfection. Furthermore, to begin elucidating the influence of cell division on the removal of the introduced DNA methylation over time, some samples were not passaged until harvesting 10 dpt. These samples were kept in 6-well plates and, in one instance, starved of serum to further decrease the number of cell divisions.

We then isolated gDNA and performed bisulfite conversion followed by the amplification of the targeted *CXCR4* locus and Illumina MiSeq sequencing. The resulting changes in DNA methylation are shown as heat maps in figure 3.12A. The changes in average DNA methylation are plotted in figure 3.12B.

As expected, the mock transfected samples showed no DNA methylation at the investigated *CXCR4* promoter. The lack of introduced DNA methylation for the dCas9-KRAB construct demonstrated that repression via KRAB was not sufficient to recruit and activate the DNA methylation machinery. The N-terminal fusion KRAB-dCas9-Dnmt3a3L did introduce DNA with similar efficacy as dCas9-Dnmt3a3L (7.1% and 6.9% DNA methylation 5 dpt, respectively), but the sample retained less DNA methylation on day 10 (2.6% to 4.0%). Both constructs achieved DNA methylation not only right next to the protected gRNA binding site, but also across the whole analyzed region, similar to the results in section 3.1.

However, the C-terminal fusion dCas9-Dnmt3a3L-KRAB was much more effective compared to dCas9-Dnmt3a3L alone with 11.3% DNA methylation on day 5 past transfection.

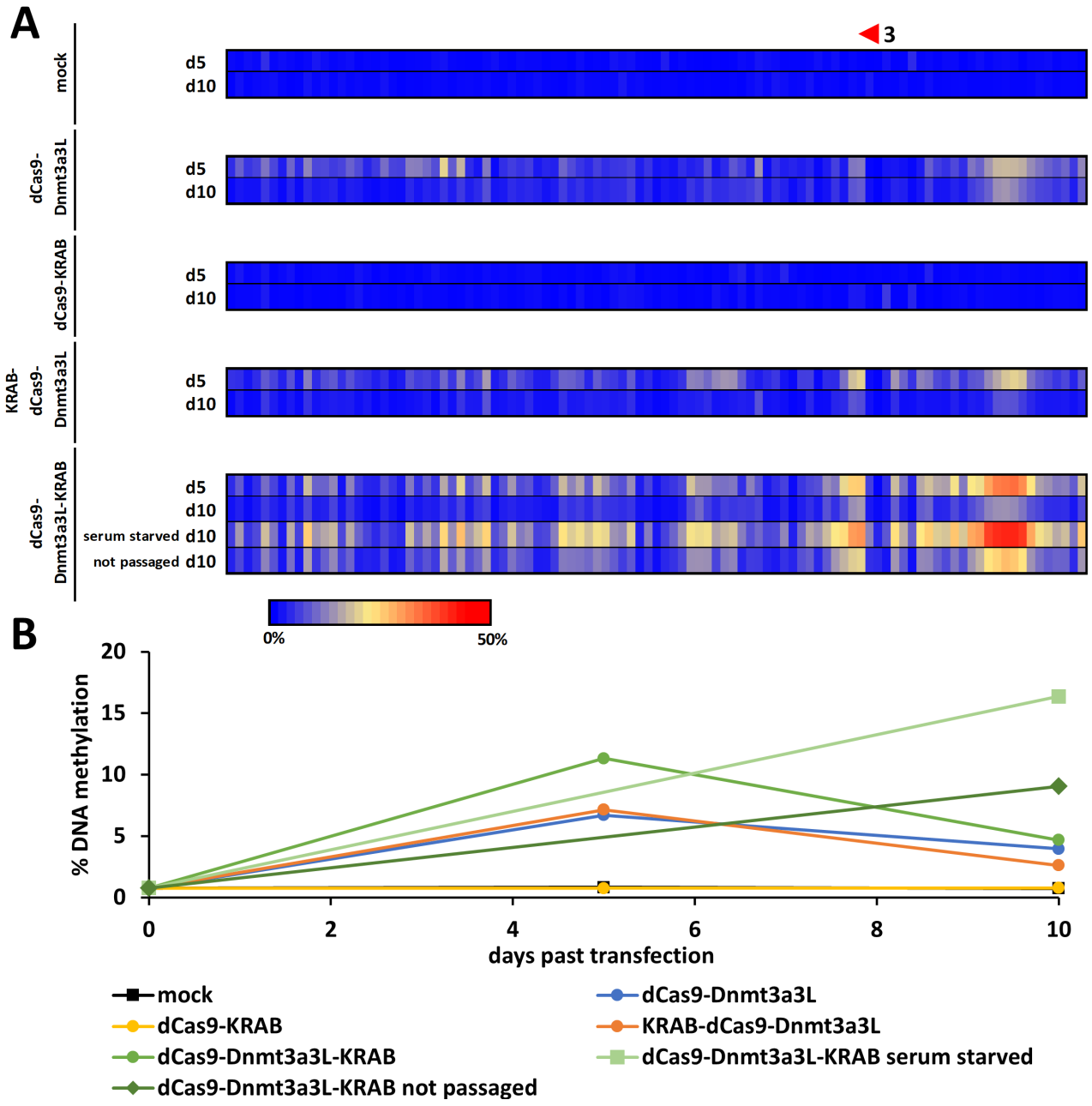

Figure 3.12: The fusion dCas9-Dnmt3a3L-KRAB introduced more DNA methylation and reduced cellular proliferation enhances DNA methylation targeting. (A) Heat map representation of the *CXCR4* promoter showing the increased DNA methylation when targeting dCas9-Dnmt3a3L and its N-and C-terminal fusions with KRAB. See figure 3.2 for a more detailed description. **(B)** Line chart of the changes in DNA methylation at the *CXCR4* promoter for each sample over time.

Nevertheless, the amount of DNA methylation left 10 dpt was again just about half as strong as on day 5 (4.7%). Interestingly, the dCas9-Dnmt3a3L-KRAB samples that were not passaged between transfection and harvesting 10 dpt showed increased DNA methylation 10 dpt (9.1%). Depriving these not-passaged samples of serum (starting 3 dpt) lead to an even higher DNA methylation content 10 dpt of 16.4% on day 10.

This superior efficacy of the new dCas9-Dnmt3a3L-KRAB fusion protein compared to the single effector domains demonstrated the importance of editing the epigenetic state of the target gene promoter along multiple pathways. Furthermore, the much higher amount of deposited and potentially retained DNA methylation when cellular proliferation was impaired either by contact inhibition or serum starvation highlighted the influence of cell division on the maintenance of the introduced DNA methylation. However, another potential explanation would be that reduced cellular growth leads to more sustained expression of the transfected plasmids.

3.3 Bio-On-Magnetic-Beads (BOMB)

Based on the results of the experiments shown so far, it was clear that further experiments probing multiple factors influencing stability were necessary. To determine the factors influencing the efficacy and stability of the introduced epigenetic marks, more target genes sampled at multiple time points to probe the temporal dimension of epigenetic editing had to be investigated. The number of resulting samples would be enormous and unfeasible to process with the methods employed until then. Therefore, we had to find or develop new techniques for downstream processing and analysis that would enable us to process many samples quickly, reliably and at a low cost.

We had already used homemade buffers for DNA clean-up and size selection using Sera-Mag SpeedBeads (GE Healthcare) as an alternative for the expensive AMPure XP (Beckman Coulter), especially for the library preparations. These buffers were based on the protocols published by Faircloth and Glenn (2011) and Jolivet and Foley (2015). The possibility to easily scale up both the volume and number of samples when working with them and to automate sample processing were especially helpful. Therefore, we chose to adopt and develop methods based on magnetic beads. With the help of published protocols and methods, we started to synthesize magnetic nanoparticles and coat them with either silica or carboxyl groups ourselves from very cheap raw materials. More detailed information regarding the synthesis of magnetic beads can be found in our publication Oberacker et al. (2019).

Additionally, we developed methods and buffers for the isolation of nucleic acids using these self-made magnetic beads. I focused on the development of protocols for the extraction of TNA, RNA or gDNA from cultured mammalian cells. Jule Focken and Vivien Meyer, two

students whose BSc theses I supervised, helped with the development and testing of these methods.

For the isolation of total RNA, we were successful by combining the robust TRI-buffer method (Chomczynski and Sacchi, 1987; Chomczynski and Sacchi, 2006) with binding of the nucleic acids to silica-coated beads (before the addition of chloroform). After several washing steps with 90% ethanol, DNA can be removed with an on-bead DNase I digest followed by more washing steps. This BOMB protocol 8.1 yielded highly pure and intact total RNA, retaining both the ribosomal RNAs as well as mRNA and tRNA (see figure 3.13A). The isolated RNA had a $A_{260\,nm}/A_{280\,nm}$ ratio of around 1.9 when measuring adsorption with Nanodrop (see figure 3.13D). Furthermore, the isolated RNA was suitable for sensitive downstream processes like RT-qPCR, as demonstrated in figure 3.13F via a serial dilution of cDNA in qPCR. With an additional extraction step with TRI reagent to remove RNases, this method can be adopted to isolate TNA or RNA from *E. coli* (see figure 3.13G). In the resulting publication (Oberacker et al., 2019) we also provide a protocol for the preparation of the TRI reagent.

Another possibility to isolate nucleic acids was initially devised by our collaborators Dr. Tim Hore and Dr. Ferdinand von Meyenn. This method functions with a lysis buffer based on guanidinium isothiocyanate (GITC) and sodium lauroyl sarcosinate (sarkosyl) and uses 2-propanol (isopropanol) to precipitate nucleic acids to the magnetic beads. My students and I continued to improve this protocol and adopt it for additional applications. It has the benefit of being far less toxic than the phenol-based TRI system used in the paragraph above, removing the need to work in a fume hood. By default, the method isolates total nucleic acid, though it is possible to isolate only RNA by adding an on-bead digest with DNase I after the first washing steps. Conversely, including RNase A for the pellet resuspension step leads to the isolation of only DNA with a $A_{260\,nm}/A_{280\,nm}$ ratio of around 1.8 (see figure 3.13B, C and E). Furthermore, this method is suitable to isolate gDNA including potential plasmids from *E. coli* and other organisms (see figure 3.13H).

In addition to these methods for the isolation of nucleic acids, I also developed a fast and efficient method for bisulfite conversion of DNA utilizing magnetic beads based on Shiraishi and Hayatsu (2004). Initially, purified DNA is converted with a conversion reagent containing sodium metabisulfite, sodium hydroxide (NaOH), dimethylformamide and hydroquinone. DNA melting and conversion is enhanced by repeatedly cycling between 54 °C and 95 °C in a thermocycler. Afterwards, the DNA is bound to silica-coated magnetic beads with guanidine-HCl and washed with 90% ethanol. The bisulfite conversion is finished with a NaOH-containing desulfonation buffer while the DNA is still bound to the beads, followed by an additional wash step and recovery of the converted DNA. DNA converted with this BOMB protocol 9.1 is suitable for amplification via PCR and achieved conversion rates of unmethylated cytosines to uracils (and thymines after PCR amplification) as high as commercial kits (98% and higher) as demonstrated by Sanger sequencing (see figure 3.14B and C).

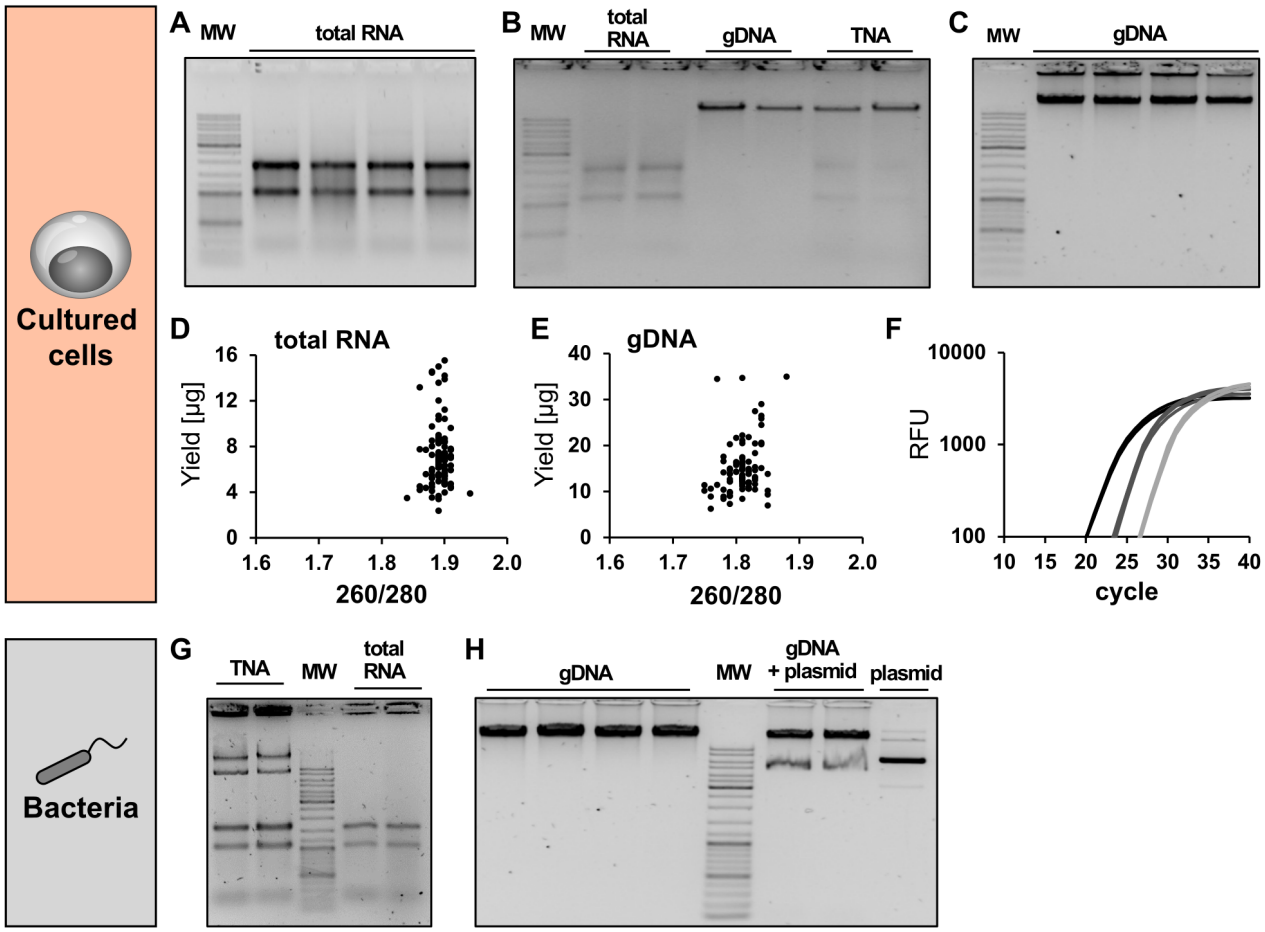

Figure 3.13: BOMB protocols achieve high-throughput isolation of high-quality nucleic acids (Oberacker et al., 2019). **(A)** RNA isolated from HEK293 cells according to BOMB protocols 8.1. MW: GeneRuler DNA Ladder Mix (Thermo Fisher Scientific). **(B)** TNA isolated from HEK293 cells with BOMB protocol 6.1 in the rightmost two lanes, followed by either RNase A digest or DNase I digest yielding gDNA and RNA, respectively. **(C)** gDNA isolated from HEK293 cells according to BOMB protocols 7.1. **(D, E)** Total yield plotted over the ratio of $A_{260\ nm}/A_{280\ nm}$ for each sample of the RNA and gDNA isolations from HEK293 cells, respectively. **(F)** RT-qPCR amplification curve of a 10-fold serial dilution (black: undiluted; dark grey: 1:10 diluted; light grey: 1:100 diluted) of RNA from HEK293 isolated according to BOMB protocol 8.1. **(G)** TNA isolated from *E. coli* using BOMB protocol 6.6 (left two lanes), followed by a DNase I digest yielding only RNA as shown in the two lanes on the right. **(H)** Agarose gel electrophoresis of gDNA isolated from TOP10 *E. coli* (left four lanes) compared with DNA isolated from *E. coli* transformed with a plasmid (pHAGE-EFS) in lanes 6 and 7. The purified plasmid is shown as reference in the rightmost lane. DNA was isolated using BOMB protocol 7.1.

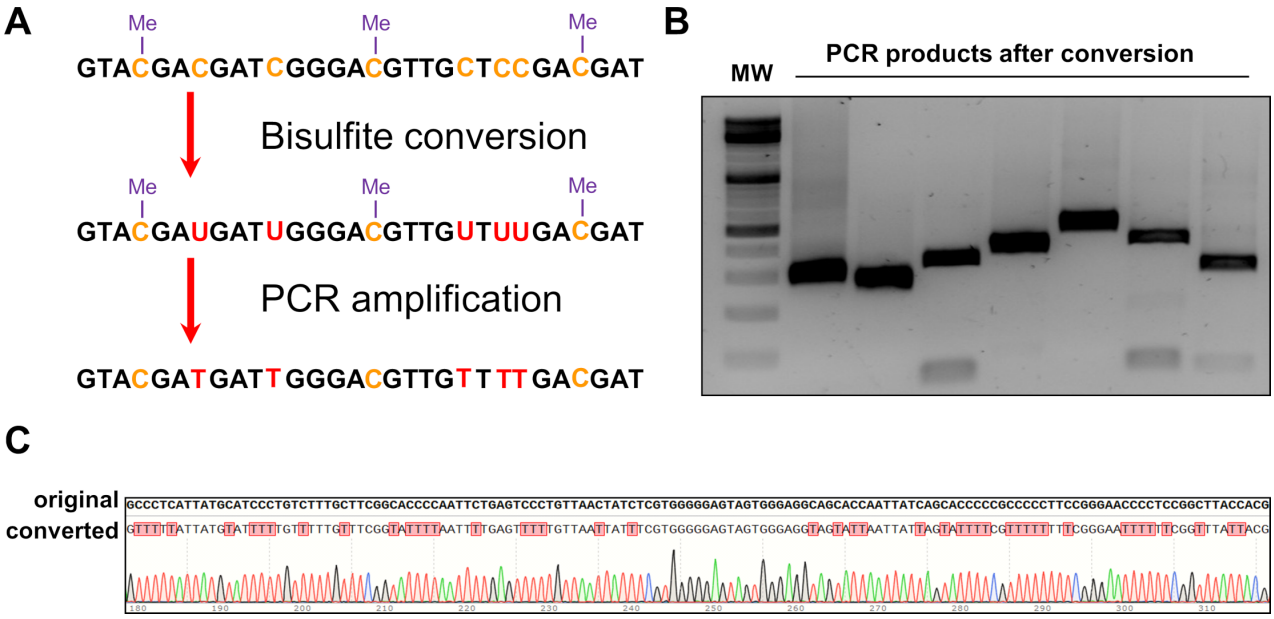

Figure 3.14: Bisulfite conversion with BOMB (Oberacker et al., 2019). **(A)** Bisulfite converts unmethylated cytosines to uracils, which result in thymines after PCR amplification. **(B)** Agarose gel electrophoresis of PCR amplification products of DNA converted with BOMB protocol 9.1. **(C)** Comparison of the original and the converted DNA sequence of PCR amplified bisulfite converted DNA showing the conversion of all non-CpG cytosines.

In addition to the methods and protocols described in this chapter, the resulting publication Oberacker et al. (2019) also contains detailed instructions for the synthesis of magnetic nanoparticles and their coating with either silica or carboxyl groups. Furthermore, we provide several methods for purifying DNA, including the extraction from agarose gels, and a method for the removal of nucleotides while retaining short DNA fragments. Additionally, we developed protocols for the isolation of plasmids from *E. coli* both with and without the need for a centrifuge. The TNA isolation protocol 6.1 described above was also adopted for specialized samples like mammalian tissues, plants, yeast and environmental samples.

These methods are suitable for high-throughput workflows as it is possible to process multiple 96-well plates simultaneously. However, speed, the possibility to easily scale the number of samples and even automation are only one part of the suitability of a method for high-throughput work. It is also important that the cost of the method including all consumables are not prohibitively expensive as the lab must also be able to afford processing so many samples.

As the home-made functionalized magnetic beads are synthesized using very cheap chemicals (costing less than 0.01 € per sample) and most other chemicals necessary for the BOMB protocols are also very affordable, extreme cost-savings are possible compared to commercial kit solutions. Additionally, the most expensive components of our methods are generally enzymes and high-quality ethanol, which often need to be supplied by the user even when using commercial kits. Table 3.1 lists the costs per 96 samples of the most important BOMB protocols and of commercial alternatives from three major suppliers. The BOMB methods enable cost savings between 10 and 32-fold, depending on the application.

Table 3.1: Cost comparison of BOMB protocols and commercial solutions (Oberacker et al., 2019). The complete costs for 96 samples were calculated taking into account plastics, solvents, and enzymes like DNase I. For the kit content replacement cost column, these costs were omitted. The range of prices for commercial kits were taken from three major suppliers.

Protocol	BOMB: complete cost per 96 samples	BOMB: Kit content replacement per 96 samples	Price of commercial kits per 96 samples	Price advantage (kit/BOMB complete cost)
Clean up	5 €	0.34 €	155 - 218 €	31 - 43x
Gel extraction	11 €	1.50 €	155 - 228 €	14 - 21 x
Plasmid DNA extraction	14 €	5.00 €	146 - 191 €	10 - 14 x
TNA extraction from cells	11 €	2.20 €	349 - 1332 €	32 - 121x
TNA extraction from tissues	23 €	6.50 €	349 - 471 €	15 - 20x
gDNA extraction	12 €	2.20 €	183 - 354 €	15 - 30x
Total RNA extraction	32 €	2.50 €	440 - 592 €	14 - 18x
Bisulfite conversion	16 €	5.00 €	207 - 655 €	13 - 41x

Another cost factor are the magnetic racks which are necessary to employ magnetic beads. These racks can cost several hundred euros, especially for 96-well plates. We developed 3D-printed or laser-cut magnetic racks, both for single 1.5-2.0 ml microcentrifuge tubes and 96-well plates (see figure 1.12). The plate magnetic racks work for both (PCR) microtiter plates and deepwell plates (2.2 ml per well). As these magnetic racks use inexpensive magnets and material costs are almost negligible, they enable even more cost savings.

3.4 A combination of factors is necessary for a stable switch of the epigenetic state of the target

Targeting dCas9-Dnmt3a3L to the promoter of an actively expressed gene resulted in the introduction of widespread DNA methylation and target gene repression (see section 3.1). However, the first time course experiment showed that the DNA methylation was lost quickly once the cells lost expression of the transfected plasmids. This process could not be slowed down by inhibiting TET enzymes via the dioxygenase inhibitor DMOG (see section 3.2.1). Interestingly, the amount of DNA methylation at the target gene promoter was higher when the rate of cell division and DNA replication was reduced via contact inhibition or serum starvation (see section 3.2.2). Furthermore, the effects on DNA methylation and repression

were enhanced when targeting a fusion protein of the DNA methyltransferase Dnmt3a3L and the repression domain KRAB acting on the level of histone modifications. These results pointed to the possible explanation that the effect size both shortly after transfection and its stability depend on the complete epigenetic state of the target gene and how it influences the maintenance of DNA methylation (see section 1.1.2).

I wanted to continue investigating the stability of targeted DNA methylation using the high-throughput methods for nucleic acid handling (see section 3.3). However, I first had to test co-targeting multiple promoters to reduce the number of samples for subsequent time course experiments. In addition, a high-throughput method to enrich transfected cells would make it easier to identify the effect of epigenome editing.

3.4.1 Multiplexed target gene repression

In preparation for subsequent long-term experiments I tested whether it is possible to co-target multiple target genes with several gRNAs each and what ratio of transfected DNA of dCas9-ED-plasmid and gRNA plasmids achieves the strongest and most stable target gene repression. Additionally, a new target gene, *IL6ST*, was introduced in this experiment. To achieve strong effects, the fusion protein dCas9-Dnmt3a3L-KRAB was chosen, together with the gRNAs demonstrating the highest efficacy for each target: TFRC sg3, 6 and 9; CXCR4 sg2 and 3 (see figure 3.5); and IL6ST sg1, 3 and 5 (data not shown).

I seeded 0.7 million HEK293 cells into 6-well plates with two biological replicates per sample in the morning and transfected the cells in the evening. The fraction of transfected DNA allocated for the dCas9-Dnmt3a3L-KRAB plasmid and the equimolar gRNA plasmid mixes ranged from 5% to 40% and 90% to 45%, respectively. Conversely, the amount of the fluorescence reporter plasmid remained the same at 5%, but was changed from the mVenus plasmid to eGFP-Puro (Addgene #45561) to allow the enrichment of transfected cells by puromycin selection. For the gRNA mixes, I pooled either the two or three gRNAs for each sample separately, or all eight together to test how co-targeting influences the effect strength. This would reduce the number of samples for future experiments.

Puromycin was added at a concentration of 2.5 µg/ml 20 h and 2 days after the transfection. Previous experiments had determined this to be a sufficient concentration and time to kill HEK293 cells not carrying the PuroR-coding plasmid. The cells were harvested on days 3, 5 and 9 past transfection. Changes in target gene expression were analyzed with RT-qPCR for the time points 3 and 9 dpt after isolating RNA according to BOMB protocol 8.1 (see section 3.3). As seen in figure 3.15, targeting dCas9-Dnmt3a3L to the promoters of *TFRC*, *IL6ST* and *CXCR4* led to a strong repression of gene expression to between 0.4 or even 0.2 for *CXCR4* relative to mock transfected cells 3 dpt. There was no clear relationship between targeting a single versus all three promoters and the ratio of dCas9-ED to gRNAs for all samples, as the differences between the different samples were small for each target gene.

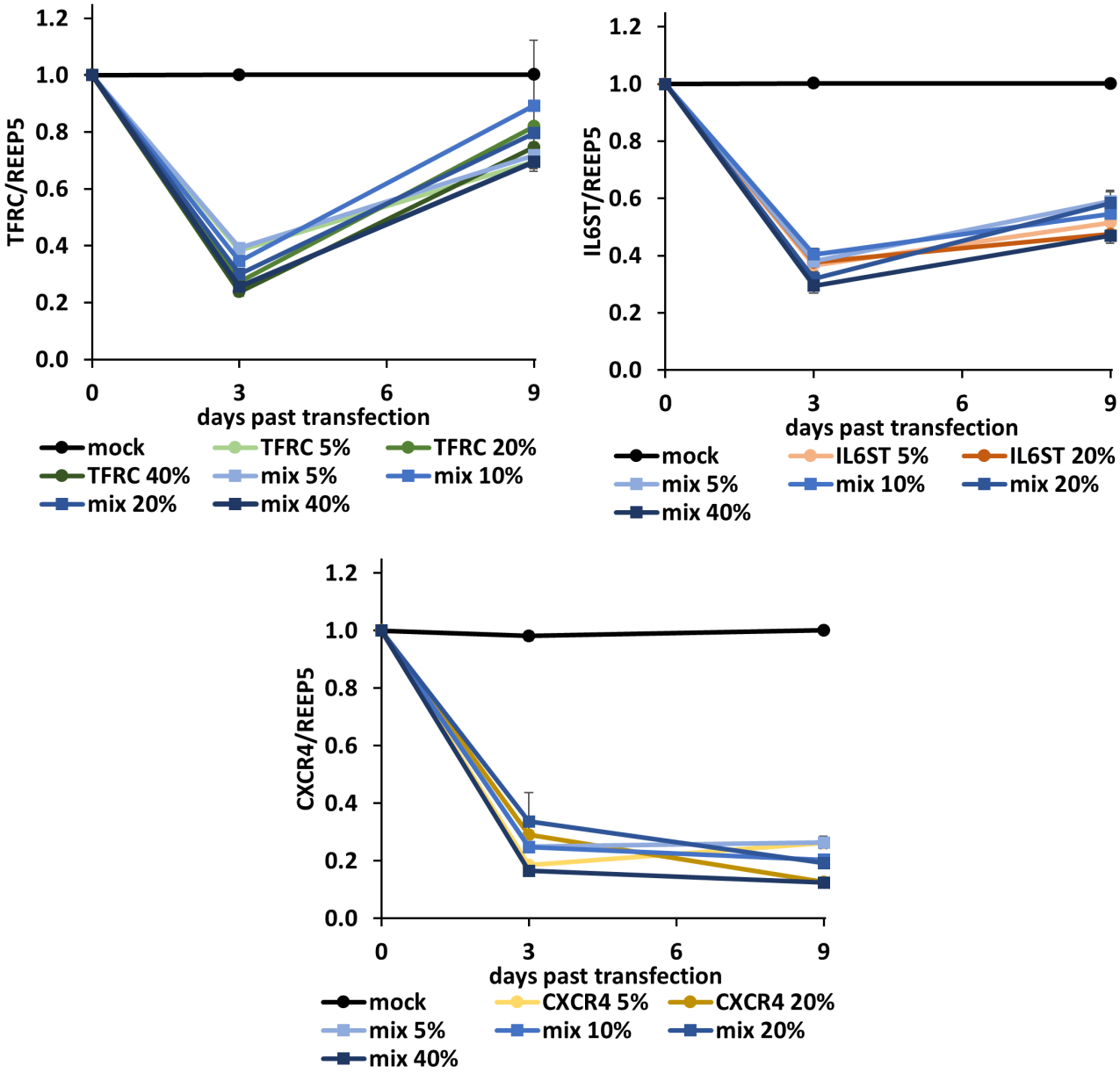

Figure 3.15: Co-targeting dCas9-Dnmt3a3L-KRAB to three target genes leads to comparable effects as targeting a single gene. RT-qPCR results from two biological replicates (error bars denote SEM) 9 dpt in HEK293 cells. The sample name "mix" means that this sample was transfected with a mix of a total of 9 gRNAs targeting the promoters of TFRC, IL6ST and CXCR4, whereas only the gene name means that this sample received only the two or three gRNAs targeting that single gene. The number next to the sample name denotes the fraction of total transfected DNA belonging to the dCas9-Dnmt3a3L-KRAB plasmid.

The three target genes were all repressed on the first time point 3 dpt. However, their responses differed substantially 6 days later: while *TFRC* expression returned to between 0.7-fold to 0.9-fold, *IL6ST* stayed repressed with 0.5-fold to 0.6-fold expression compared to the mock transfected sample. Strikingly, the expression of *CXCR4* was even lower on day 9 compared to 3 dpt for all samples except for both only the CXCR4 gRNAs and the mix with 5% dCas9-Dnmt3a3L-KRAB. The most effective samples contained gRNAs for all three targets and either 20% dCas9-Dnmt3a3L-KRAB, which went from 0.3-fold to 0.2-fold expression from day 3 to 9 dpt, and with 40% dCas9-Dnmt3a3L-KRAB plasmid (from 0.2-fold to 0.1-fold).

These results demonstrated that the transfection of nine gRNAs co-targeting three different promoters worked well without a loss of effect size. The sample with 20% dCas9-ED and 75% gRNA mix was among the best for all target genes. As the presence of even more epigenetic effector domains in the cells might lead to increased off-target effects, I used this combination for subsequent experiments.

3.4.2 Combining targeted epigenome editing with inhibitors against epigenetic enzymes

I wanted to find a method to further facilitate the switch of the target genes epigenetic state besides targeting DNA methylation with Dnmt3a3L and repressive histone marks with KRAB. Therefore, I employed two small molecule inhibitors against the H3K79 methyltransferase DOT1L (SGC0946) and the H3K4 methyltransferase MLL1 (OICR-9429). We had hypothesized that both of these marks, which are associated with actively transcribed genes, were repressing the endogenous DNA methylation maintenance machinery. The epigenetic marks H3K79me1/2/3 set by DOT1L have no known demethylase, making it a potential candidate for a stable mark that could block DNA methylation maintenance at actively transcribed loci. Furthermore, DOT1L increased gene expression when targeted (Cano-Rodriguez et al., 2016). For H3K4me3, which is set by MLL1 in complex with other proteins, the link to DNA methylation is clearer: The ADD domain in DNMTs binds unmethylated H3K4, making this a necessity for *de novo* DNA methylation with the endogenous full length DNMT3 proteins.

We then tested whether the combination of DNA methylation (via Dnmt3a3L), H3K9me3 (KRAB), less H3K4me3 (MLL1 inhibitor) and less H3K79me2/3 (DOT1L inhibitor) would suffice to achieve a stable switch of the epigenetic state of the target gene promoters. Presumably, the repressive marks would then recruit and activate the DNA methylation machinery consisting of UHRF1 and DNMT1 to propagate the introduced DNA methylation and the repressed transcriptional state after cell division to the daughter cells.

Furthermore, for this experiment we again included the general dioxygenase inhibitor DMOG to multiple samples. These samples again tested the result from section 3.2.1 that inhibiting TET enzymes does not change the kinetics of DNA methylation loss after the

transfected plasmids are not expressed anymore. For this experiment, DMOG was added to the respective samples at a concentration of 0.5 mM. Treatment started after the collection of the 3 dpt time point, as DMOG is already effective within less than 24 h (Marchbank et al., 2011).

Four dCas9-ED fusion proteins were used for this experiment: dCas9-Dnmt3a3L, dCas9-Dnmt3a3L-KRAB, dCas9-KRAB and the histone deacetylase recruiting (Adams, Chandru, and Cowley, 2018) fusion of the Sin3-interaction domain (SID, used in four repeats) dCas9-SID4X. Targeting these fusion proteins to three different gene promoters with the addition of the inhibitors mentioned above meant that a lot of samples would have been required to test the possible combinations independently. I previously showed that co-targeting these three gene promoters did not diminish the effect sizes (3.4.1), therefore I could reduce the number of samples by a factor of three by co-targeting the three promoters. Additionally, I again used puromycin selection to enrich the fraction of transfected cells.

Transfection

As described in section 2.2.2, HEK293 cells were pre-treated with the DOT1L and MLL1 inhibitors starting four days before the transfection. DMOG, on the other hand, was added to the respective samples starting after the collection of the 3 dpt time point at a concentration of 0.5 mM, as it is already effective within less than 24 h (Marchbank et al., 2011).

The samples in this experiment are listed on the left side in figure 3.16. All samples were transfected with 5% eGFP-Puro to enable puromycin enrichment of transfected cells. As before, mock transfected samples additionally contained 20% pUC19 plasmid and 75% empty gRNA vector. This empty gRNA vector was also used for the untargeted control samples, which used 20% of the respective dCas9-ED plasmids just like the targeted samples. All other samples were transfected with a gRNA mix consisting of an equimolar mix of TFRC sgRNAs 3, 6 and 9; IL6ST sgRNAs 1, 3 and 5; and CXCR4 sgRNAs 2 and 3. Puromycin was added to the cells 20 h and 2 days after the transfection at a concentration of 2.5 μg/ml. The experiment was carried out with biological triplicates in 6-well plates using HEK293 cells.

The cells were harvested in three aliquots of around 0.25 million cells and passaged on days 3, 6, 9, 12, 15, 18 and 21 past transfection. The cell pellets were then used for the isolation of RNA and gDNA to analyze the effects on target gene transcription and promoter DNA methylation. Vivien Meyer, a student I supervised for her BSc thesis, isolated RNA according to BOMB protocol 8.1 and performed RT-qPCR. The expression results were normalized for each set of samples to the mock transfected control with the same inhibitor and are presented as heat maps for all three target genes in figure 3.16.

Effects on *TFRC* transcription

On the first analyzed time point 3 days past transfection, all four effector domains achieved repression of all three target genes for all inhibitor combinations: no inhibitors, DOT1L inhibitor, MLL1 inhibitor, both combined, and DMOG. However, especially without any inhibitors, the samples transfected with dCas9-KRAB or dCas9-SID4X quickly lost their

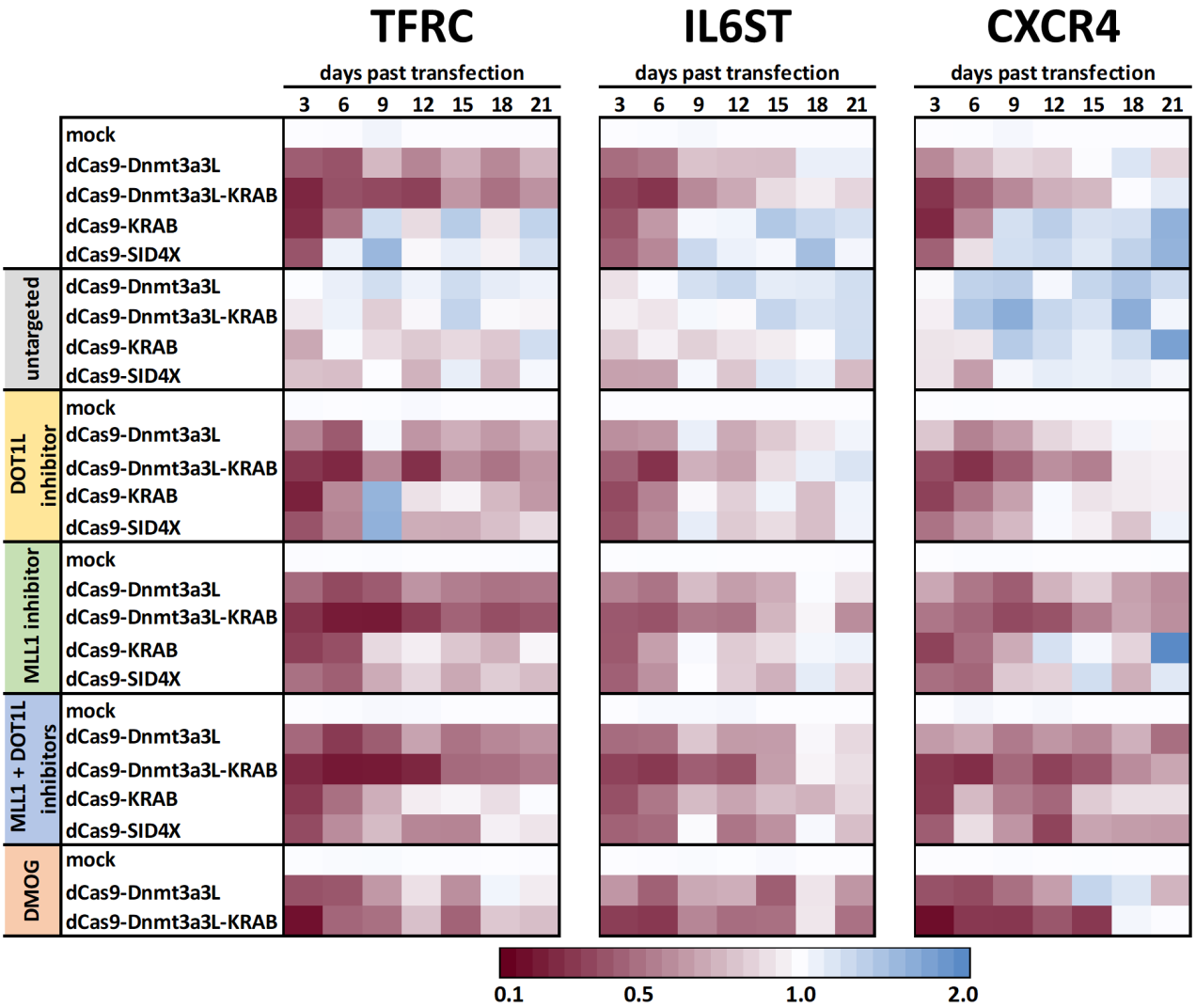

Figure 3.16: Heat map representation showing the RT-qPCR results of targeting four effector domains to three target genes with multiple added inhibitors. The table on the left lists the different dCas9-EDs transfected into HEK293 cells co-targeting the promoters of *TFRC*, *IL6ST* and *CXCR4* and the inhibitors added to the cell culture medium. The color of each heat map cell denotes the average expression of this target gene for the time point indicated at the top with the color legend located at the bottom of the figure. Generally, the average expression was taken from three biological replicates except for a few time points for which one replicate was removed for technical reasons.

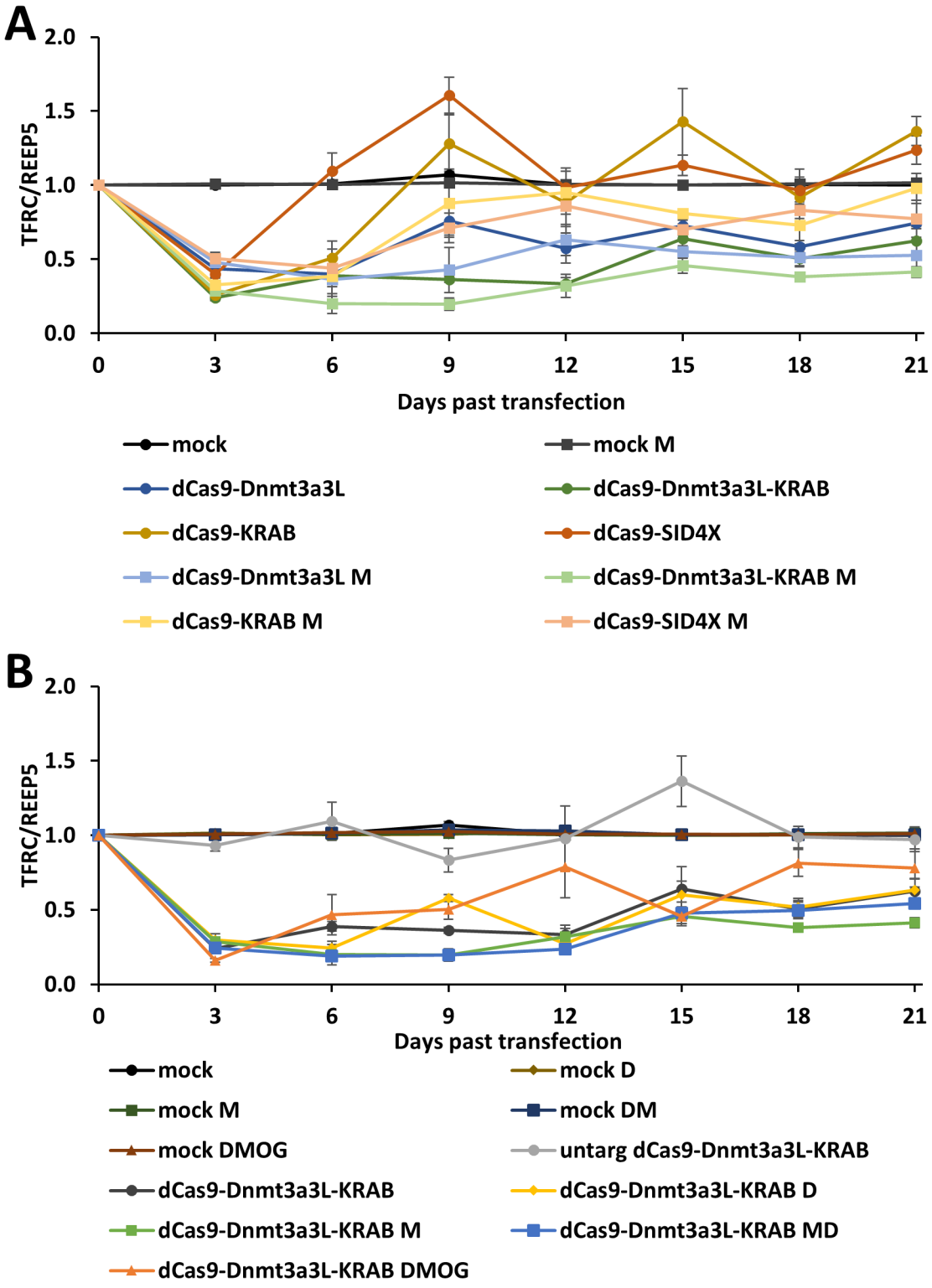

Figure 3.17: Kinetics of gene expression when targeting four effector domains to *TFRC* with multiple added inhibitors. (A) Line chart showing the effects of all four effector domains both without inhibitors and with the MLL1 inhibitor. **(B)** Comparison of effects when dCas9-Dnmt3a3L-KRAB to TFRC with all tested inhibitors. Error bars denote SEM of at least two biological replicates.

repressive effects over the next time points. Interestingly, the inhibitors against MLL1, DOT1L, or both enzymes together, led to lower target gene expression at later time points compared to the samples without inhibitors.

On the other hand, dCas9-Dnmt3a3L and especially dCas9-Dnmt3a3L-KRAB repressed the target genes much longer. In the case of *TFRC*, repression lasted even until the last time points 18 and 21 dpt (0.59±0.04 and 0.75±0.02-fold expression for dCas9-Dnmt3a3L, and 0.51±0.06 and 0.62±0.08-fold expression compared to mock for dCas9-Dnmt3a3L-KRAB, respectively). The MLL1 inhibitor increased the stability of the target gene repression for these effectors just as for KRAB and SID4X. The line chart in figure 3.17A illustrates the differences between the four effector domains and between the sample with no inhibitors and with the MLL1 inhibitor for the *TFRC* target gene.

The fusion effector domain dCas9-Dnmt3a3L-KRAB was the most effective one at inducing target gene repression. This was measurable both for the initial time points, when the construct was still expressed strongly, and the later time points, when the repressed transcriptional state had to be maintained by the endogenous epigenetic machinery. Therefore, this effector domain is well suited to illustrate the influence of the different inhibitors as shown in figure 3.17B: Comparing the inhibitor samples on days 18 and 21, the MLL1 inhibitor achieved the strongest repression at these time points with just 0.41±0.04-fold expression. This was closely followed by the MLL1 and DOT1L inhibitor combination with 0.54±0.01-fold expression, the samples with only the DOT1L inhibitor and no inhibitor at all with 0.63±0.08 and 0.62±0.08-fold expression, respectively. The DMOG sample with just 0.78±0.13-fold expression compared to the mock transfected controls was least effective at maintaining repression. This demonstrates both the stability-enhancing effects of MLL1 inhibition and the lack of influence of dioxygenases like TET enzymes on *TFRC* expression over time. Similarly, inhibiting DOT1L did not affect the kinetics of *TFRC* expression for dCas9-Dnmt3a3L-KRAB.

Effects on *IL6ST* and *CXCR4* transcription

The results for the other two target genes, *IL6ST* and *CXCR4*, were similar to those for *TFRC* as described above. However, variability was slightly higher for these two target genes, potentially due to technical difficulties with the RNA isolations. This might also explain the outlier time point 18 dpt for *IL6ST* and the different trends observed for *CXCR4* between this experiment and the results in section 3.4.1. Whereas *CXCR4* was still repressed just as strongly by dCas9-Dnmt3a3L-KRAB 9 dpt as on day 3 past transfection in section 3.4.1 even without the addition of any inhibitors, in this experiment repression had already declined from the initial 0.29±0.03-fold to 0.59±0.04-fold expression 9 dpt. Also compared to *TFRC* in this experiment, the repression effects on *CXCR4* were less stable even when the inhibitors were added. However, at least until 12 or 15 days past transfection, respectively, the dCas9-Dnmt3a3L and dCas9-Dnmt3a3L-KRAB samples with DMOG stayed more repressed compared to both no inhibitor and the MLL1 inhibitor, in contrast to *TFRC*.

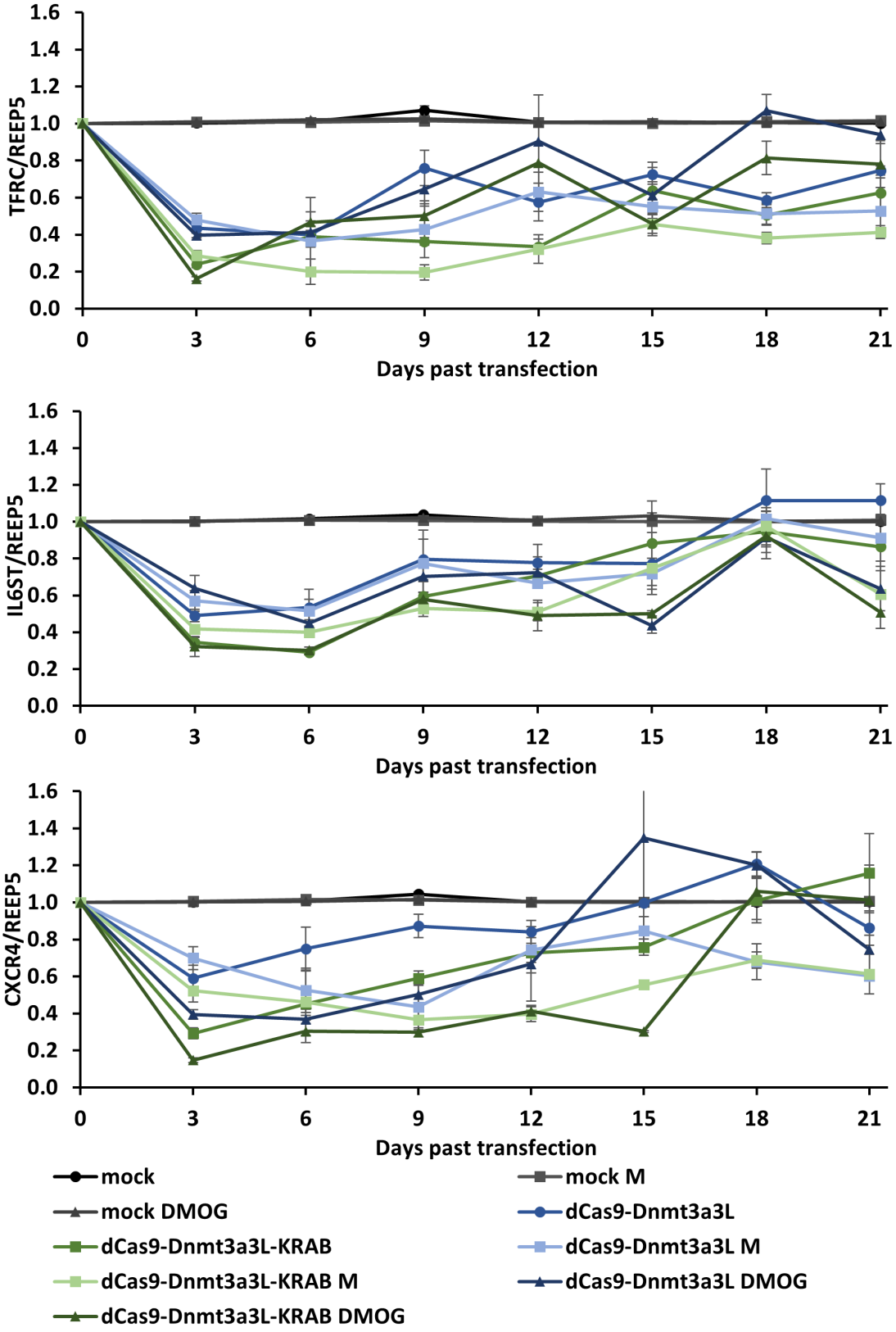

Figure 3.18: Comparison of target gene expression kinetics for dCas9-Dnmt3a3L and dCas9-Dnmt3a3L-KRAB with no added inhibitors, the MLL1 inhibitor and DMOG. Error bars denote SEM of at least two biological replicates.

This small stability-increasing effect of DMOG was also measured for the target gene *IL6ST*. Figure 3.18 shows the effects of dCas9-Dnmt3a3L and dCas9-Dnmt3a3L-KRAB with no inhibitors, the MLL1 inhibitor and DMOG for all three target genes for better comparability.

Summary

This experiment confirmed the synergistic effects of the different domains of the dCas9-Dnmt3a3L-KRAB fusion protein both on the initial target gene repression and the stability of that downregulation. This means that the introduced marks have to influence the activity of the DNA methylation machinery at the edited locus. It seems that the reduction of H3K4me3 abundance via the MLL1 inhibitor is beneficial for achieving stability at a promoter targeted for repression. In contrast, the H3K79me mark, which should be influenced by the DOT1L inhibitor, did not modulate long-term target gene repression.

Similarly, the influence of TET enzyme activity did not affect the stability of targeted repression. The DMOG-treated samples mostly lost the repressed transcriptional state even faster than without any added inhibitors. However, this response varied between the different target genes, highlighting the importance of the individual epigenetic state of the target promoter and the necessity of taking a tailored approach. To answer the open questions properly, it will be necessary to analyze the effects of the transfection and inhibitor treatment combinations on both the promoter DNA methylation and histone modifications. However, at the time of writing this thesis, these results are not yet available.

3.5 Applications of the targeted DNA methylation technology

While my primary studies focused mostly on the development and characterization of a toolbox for targeted epigenetic editing via CRISPR-Cas9 and the elucidation of the epigenetic network, I was also interested to apply these tools to study other important biological questions. Thankfully, I had the opportunity to use epigenetic editing in collaborations with groups in the UK and Serbia. The results of these collaborative studies will be detailed in the following sections.

3.5.1 Studying tumor initiation in primary breast cells using targeted DNA methylation

In cooperation with the group of Dr. Gabriella Ficz at the Barts Cancer Institute of the Queen Mary University of London, we employed targeted DNA methylation with dCas9-Dnmt3a3L as a tool to study the effects of DNA hypermethylation in breast cancer formation. However, it is important to note that my part in this study was limited to contributions to the experimental work. I constructed many of the plasmids used in this study, especially the

dCas9-Dnmt3a3L plasmid and the gRNA plasmids, and performed the targeted bisulfite work shown in figure 3.21.

While cancer development coincides with dramatic changes of the epigenetic landscape, in most cases the question of whether these changes are drivers or merely results of cancer initiation or development remains unanswered (Suvà, Riggi, and Bernstein, 2013; Easwaran, Tsai, and Baylin, 2014). Often, promoters with CpG island become methylated (Costello et al., 2000) which can lead to the repression of tumor suppressor genes. The possibility to precisely and efficiently target DNA methylation via the dCas9-Dnmt3a3L toolbox we developed in section 3.1 allowed us to study the isolated effects of tumor suppressor gene hypermethylation and to determine whether these changes can suffice to drive tumor initiation.

Therefore, the group of Dr. Gabriella Ficz transiently transfected primary myoepithelial cells isolated from healthy human breast tissue using jetPrime. The fusion protein dCas9-Dnmt3a3L was targeted by 25 gRNAs to the promoters of *HIC1*, *PTEN*, the three transcripts of *RASSF1*, and the two transcripts *p14* and *p16* of *CDKN2A* (see table 5.5 for a list of all gRNAs). Furthermore, they included the plasmid pMACS for MACS enrichment of transfected cells, which was generally performed two days past transfection. Employing the EPIC array for DNA methylation analysis, we found that 10 dpt all four target genes exhibited strongly elevated levels of DNA methylation when transfected with targeted dCas9-Dnmt3a3L compared to an inactive deletion mutant dCas9-Dnmt3a3LΔ (see figure 3.19A-D). Furthermore, the edited cells showed significantly lower expression of the targeted genes compared to the dCas9-Dnmt3a3LΔ sample 10 dpt, with the exception of *HIC1* and *PTEN*, as assayed via both RT-qPCR and RNA-seq (see figure 3.19E and F).

Additionally, the DNA methylation targeted cells remained phenotypically similar to early passage cells, whereas the untransfected or mutant transfected cells stopped proliferating and exhibited typical signs of senescence entry. They stained positive for beta-galactosidase and became large and flat (data not shown here, but can be found in the publication Saunderson et al., 2017). In contrast to this, edited cells continued growing past 15 dpt as assayed by colorimetric and counting assays (see figure 3.20A and B). This proliferating phenotype past 15 dpt could not only be achieved by targeting dCas9-Dnmt3a3L to the four genes, but also with dCas9-Dnmt3a. However, neither the inactive mutants dCas9-Dnmt3a3LΔ and dCas9-Dnmt3a3L C706A, nor dCas9 without a fused effector nor untargeted dCas9-Dnmt3a3L led to a strong increase in cell growth (see figure 3.20C). This prevention of senescence entry and the continued proliferation of DNA methylation edited cells could constitute a first step in tumor initiation.

We also found that the introduced DNA methylation via dCas9-Dnmt3a3L was maintained for multiple weeks despite the loss of the editing construct after a few days due to the transient transfection. This was already suggested by the different growth rates of edited cells 15 to 25 dpt mentioned above, but was also confirmed by the targeted bisulfite sequencing performed by me. We compared untransfected early passage myoepithelial cells from the

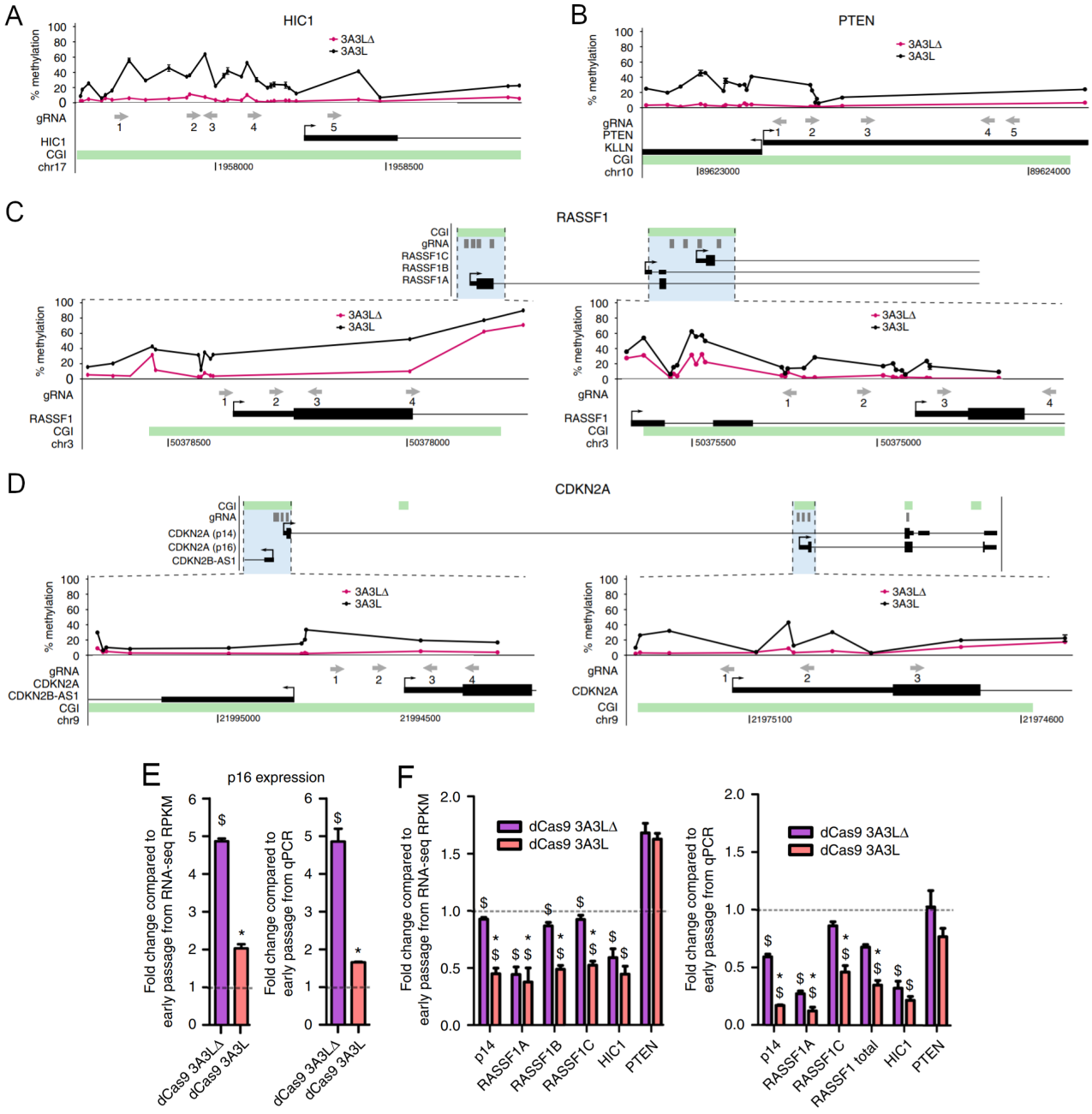

Figure 3.19: Successful targeted DNA methylation and repression of tumor suppressor gene promoters using dCas9-Dnmt3a3L (Saunderson et al., 2017). Primary myoepithelial cells from donor 1 were transfected, followed by MACS 2 dpt. The analysis of DNA methylation via the EPIC array of cells harvested 10 dpt shows increased DNA methylation with dCas9-Dnmt3a3L compared to an inactive mutants at the promoters of **(A)** *HIC1*, **(B)** *PTEN*, ***(C)*** the three transcripts of *RASSF1*, and **(D)** the two transcripts *p14* and *p16* of *CDKN2A*. CpG islands are shown in green, gRNA binding sites as grey arrows, DNA methylation of cells transfected with dCas9-Dnmt3a3L in black and an inactive deletion mutant dCas9-Dnmt3a3LΔ in red. **(E)** *p16* and **(F)** most of the other target genes transcripts are significantly repressed 10 dpt when targeting dCas9-Dnmt3a3L compared to dCas9-Dnmt3a3LΔ. n=3, mean$\pm$SEM. T-tests. Stars denote $p<0.01$ for RT-qPCR and $p<0.001$ for RNA-seq for dCas9-Dnmt3a3L compared to -Dnmt3a3LΔ, dollar signs denote $p<0.01$ for RT-qPCR and $p<0.001$ for RNA-seq for dCas9-Dnmt3a3L or -Dnmt3a3LΔ compared to early passage cells. Significance for RT-qPCR is t-test and for RNA-seq is from DESeq2 analysis with BH correction.

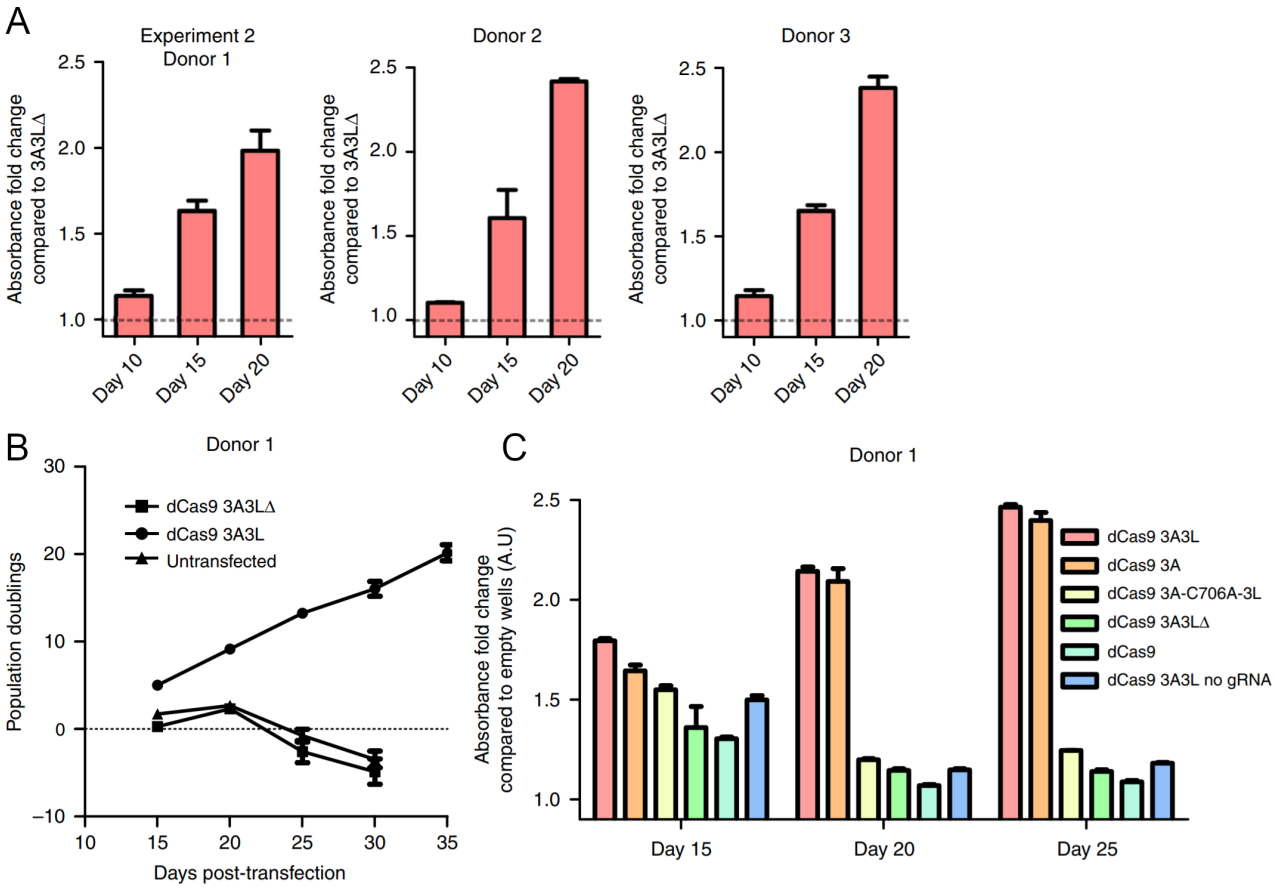

Figure 3.20: Epigenome edited cells show increased proliferation (Saunderson et al., 2017). **(A)** Colorimetric proliferation assay of transfected and sorted cells from three donors. **(B)** Cumulative population doublings of edited compared to untransfected myoepithelial cells. **(C)** Comparison of multiple effector domains and mutants on cell growth with a colorimetric assay. n=3, mean±SEM.

three donors with edited and selected cells harvested 37 to 41 dpt. For the early passage cells, the promoters of *HIC1*, *PTEN*, *RASSF1* and *CDKN2A* were almost completely unmethylated except for the shores of the CpG islands. In contrast to this, the edited cells had maintained the targeted DNA methylation which had spread in clusters across the analyzed regions (see figure 3.21A-D). Interestingly, while the gRNA binding site itself was still protected for example for PTEN sgRNA 1 and RASSF1B/C sgRNA 3 and 4 as also observed in section 3.1, most of the other gRNA binding sites were methylated by that time. As observed before with dCas9-Dnmt3a3L, DNA methylation was observed spread over the complete promoters, even several hundred base pairs from the next gRNA binding site.

Whereas DNA methylation was well maintained for multiple biological replicates with cells from donor 1 and for donor 3, the edited cells from donor 2 lost DNA methylation at the HIC1 and PTEN promoters (see figure 3.21A and B). They also resumed expressing these two genes, including the CDKN2A transcripts despite showing DNA methylation at the latter locus (data not shown here, but can be found in the publication Saunderson et al., 2017).

Furthermore, RNA-seq of cells transfected with dCas9-Dnmt3a3L or dCas9-Dnmt3a3LΔ from 10 dpt revealed that especially multiple p16 regulated transcripts were expressed

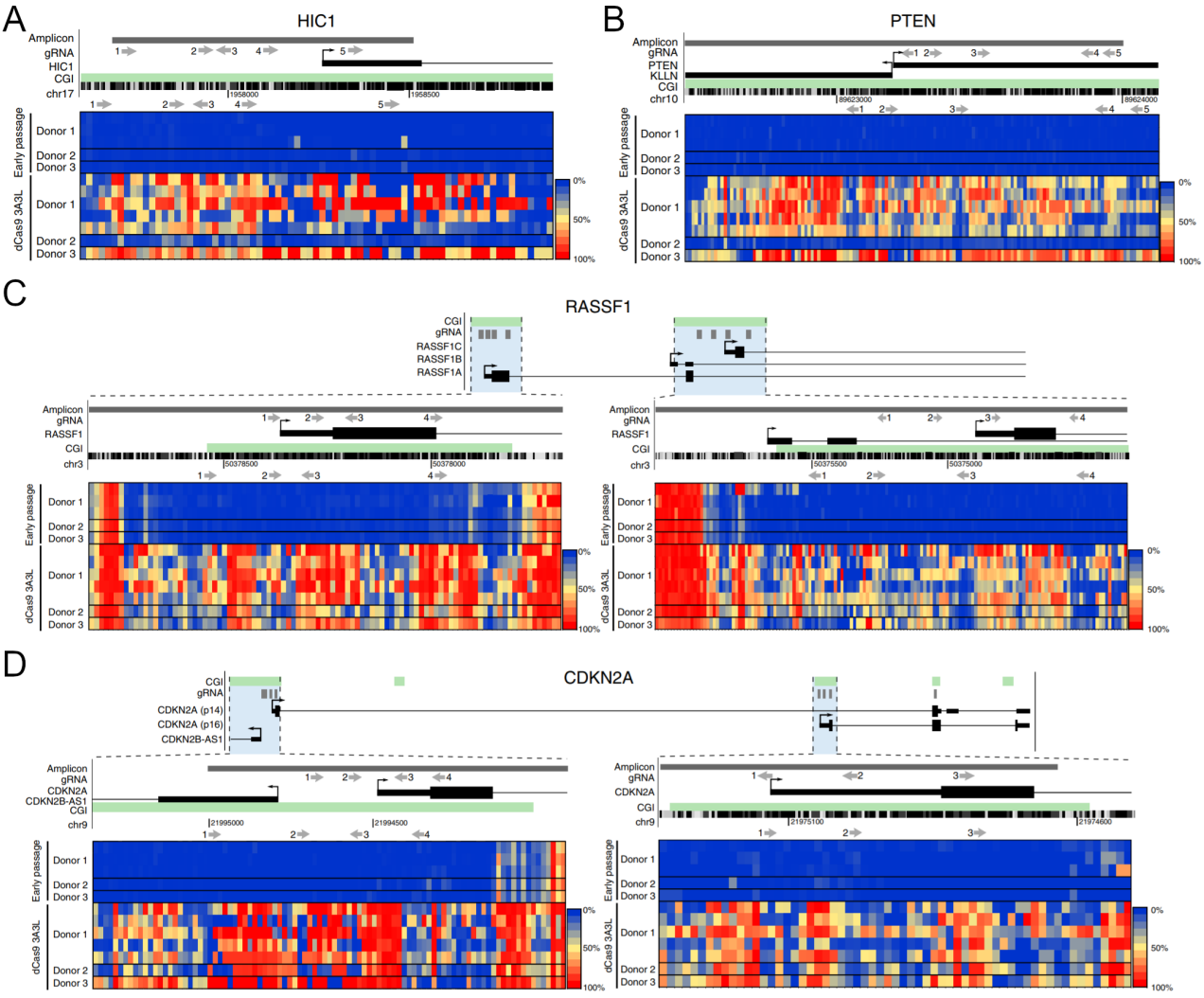

Figure 3.21: The introduced DNA methylation was maintained for at least 37 days (Saunderson et al., 2017). Bisulfite sequencing results at the promoters of **(A)** *HIC1*, **(B)** *PTEN*, ***(C)*** the three transcripts of *RASSF1*, and **(D)** the two transcripts *p14* and *p16* of *CDKN2A*. Like in figures 3.2 and 3.19, the sequenced locus with its relevant features is shown next to the heat map representing the DNA methylation state of each CpG site as shown in the color legend next to it. Each line represents data from a single replicate. The five lines at the top show results from early passage (passage 2) primary myoepithelial cells from donors 1, 2 and 3 while the bottom seven lines show data from primary cells 37–41 days after transfection with dCas9 3A3L targeting the depicted gene promoters.

stronger for the dCas9-Dnmt3a3L edited and proliferating cells. These cells were generally more similar to early passage myoepithelial cells. However, p14 associated targets like p53 were expressed at similar levels for both dCas9-Dnmt3a3L or dCas9-Dnmt3a3LΔ transfected cells. Together with previous studies (e.g. Foster et al., 1998; Lowe et al., 2015), this hinted at a driving role of the p16 transcript. Therefore, they tested whether targeting DNA methylation of this locus alone is sufficient to drive cell proliferation and prevent senescence entry. Interestingly, methylating only p16 with three gRNAs was indeed the only necessary stimulus to prevent senescence entry and led to a proliferating phenotype, whereas targeting the other genes individually did not. Nevertheless, targeting both p14 and p16 achieved twice as much cell growth compared to targeting only p16 with a single gRNA (data not shown here, but available in the publication Saunderson et al. 2017).

Taken together, these results demonstrated that hypermethylation of the CDKN2A and even the p16 promoter alone was sufficient to prevent senescence entry in late passage myoepithelial breast cells from healthy donors. The targeted DNA methylation was maintained for at least 37 days after the transient transfection and led to the continued proliferation of the edited cells as a first epigenetics-driven step in cell alteration in cancer initiation. This study resulted in the publication Saunderson et al. (2017).

3.5.2 Transdifferentiation of murine pancreatic α-cells to insulin-producing cells via targeted DNA methylation

In collaboration with the group of Dr. Melita Vidaković from the University of Belgrade we are investigating the transdifferentiation of murine pancreatic α-cells to insulin-producing cells via epigenome editing. In the experiments we use targeted repression and DNA methylation of the promoter of the master regulator *Arx* to thereby activate the expression and secretion of insulin. This is an example for a potential therapy based on epigenetic editing with CRISPR-Cas9.

The experiments for our part of the collaboration were performed together with Vivien Meyer, a student I supervised for her BSc thesis. We tested multiple transfection reagents with the fluorescent reporter plasmid mVenus to select the method with the highest efficacy for αTC1.9-cells. Out of PEI, FugeneHD and Lipofectamin 3000, FugeneHD at a ratio of 2.5:1 to DNA resulted in both the highest fraction of transfected and viable cells (data not shown). Figure 3.22A displays representative bright field and fluorescence microscopy pictures of transfected α-cells. However, with a transfection efficiency between 8% and 16% as assayed by FACS (see figure 3.22B) it was still necessary to select the transfected cells. But because the α-cells transfected with the eGFP-Puro construct died even at puromycin concentrations as low as 0.75 μg/ml, we had to rely on the enrichment of fluorescent cells with FACS.

In a first experiment, 1.6 million murine αTC1.9-cells were seeded in T25 flasks and transfected using FugeneHD with no antibiotics present in the cell culture medium during

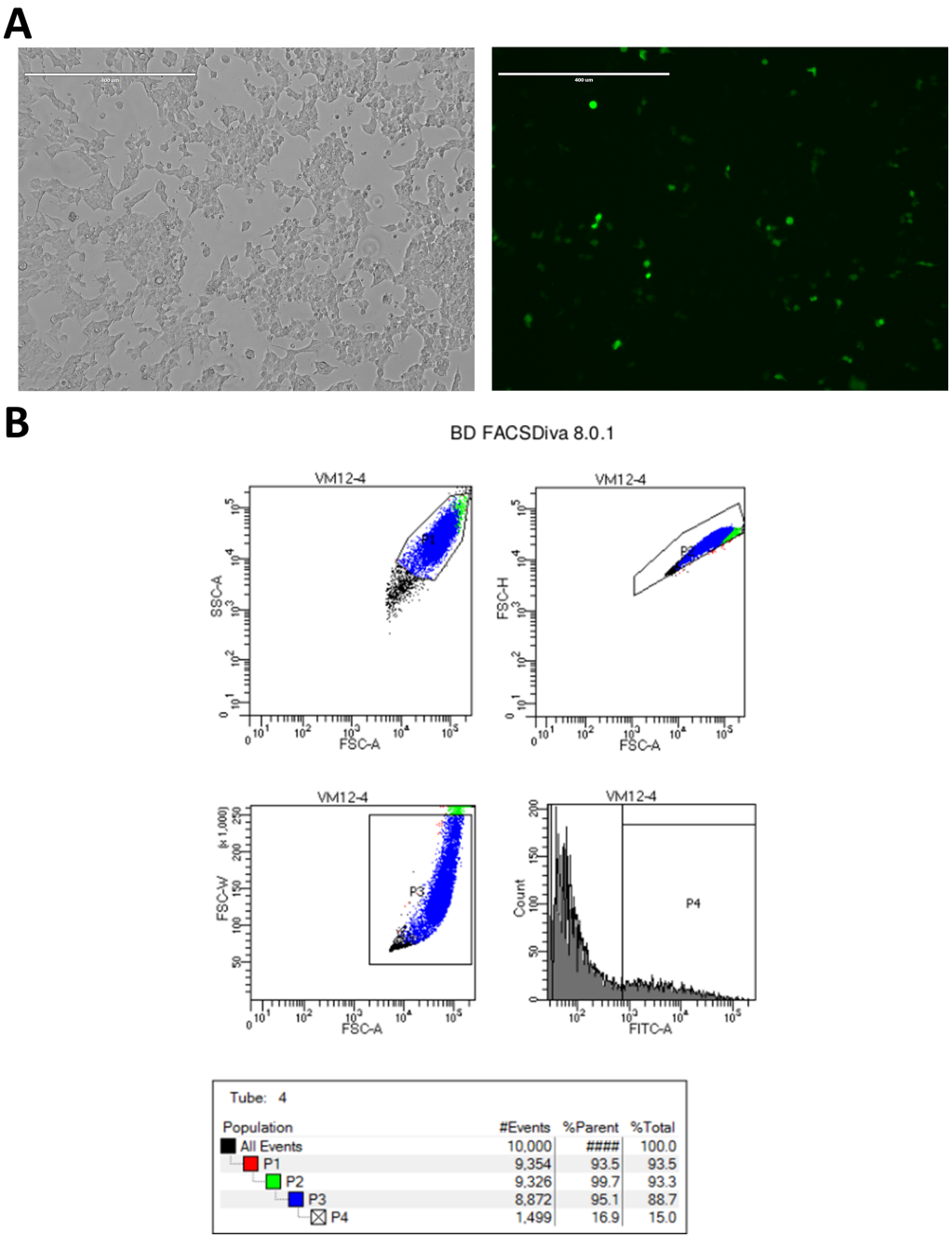

Figure 3.22: Transfection and FACS of α-cells. (A) Bright field (left) and fluorescence microscopy (right) pictures of transfected αTC1.9-cells and **(B)** diagnostic results from sorting the cells with a FACSAria III.

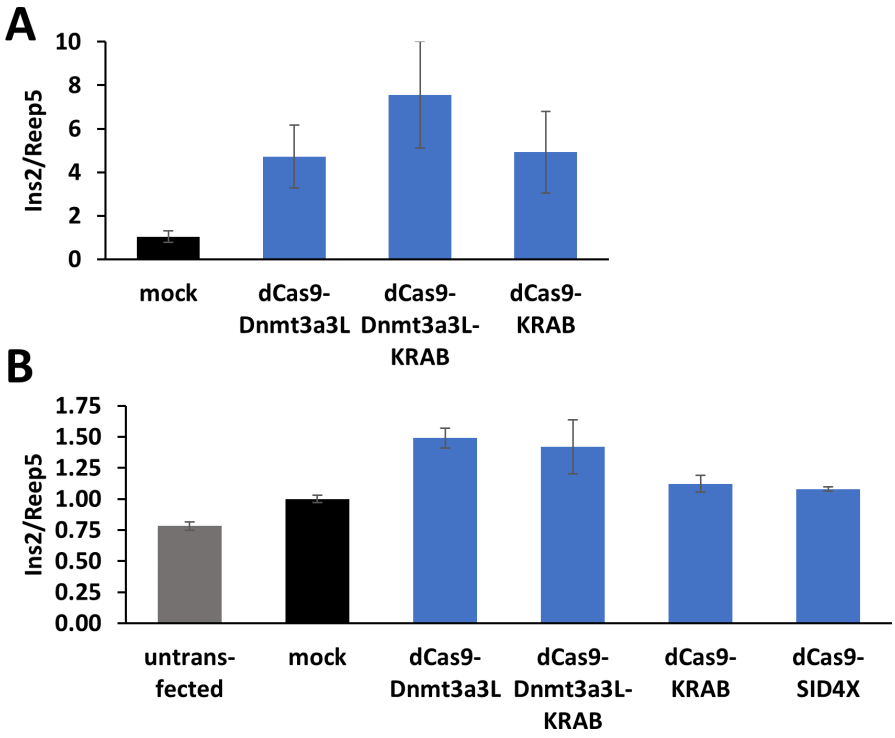

Figure 3.23: Targeted DNA methylation of *Arx* leads to expression of insulin in αTC1.9-cells. (A) In the first described experiment the transfected cells where FACSed 4 dpt and then assayed immediately. **(B)** In contrast to this, for the second described experiment the cells were FACSed 2 dpt, reseeded and eventually collected and assayed 11 dpt, resulting in much lower expression of *Ins2*. Error bars denote the standard deviation of the 6 technical replicates of the single biological replicate.

the transfection. The cells were transfected with mVenus (5%), the dCas9-ED construct (20%) and a mix of four gRNAs targeting the promoter region of Arx (75% of total DNA). Four days past transfection, fluorescent cells were selected with a BD FACSAria III (see figure 3.22B) and frozen as pellets. RT-qPCR was performed after RNA isolation with BOMB protocol 8.1 (see section 3.3). This revealed that all three tested effector domain fusions, dCas9-Dnmt3a3L, dCas9-Dnmt3a3L-KRAB and dCas9-KRAB achieved an upregulation of the *Ins2* gene (figure 3.23A) compared to the mock transfected cells (pUC19 and empty gRNA vector), reaching 7.6±2.5-fold activation with dCas9-Dnmt3a3L-KRAB. The analysis of Arx repression was not possible because of the low RNA yield, as too few cells were recovered after FACS.

We tried to increase the number of enriched transfected cells by performing the FACS enrichment 2 dpt and then continuing to culture the selected cells to achieve higher cell number yields. However, the cells proliferated slower than expected and had to be culture for 5 days in 24-well plates until the cells had multiplied enough to transfer them into 6-well plates, where they grew for another 4 days until they were finally harvested 11 dpt. At this time point, RT-qPCR showed some residual activation of *Ins2* with dCas9-Dnmt3a3L and dCas9-Dnmt3a3L-KRAB as shown in figure 3.23B, but only at low levels.

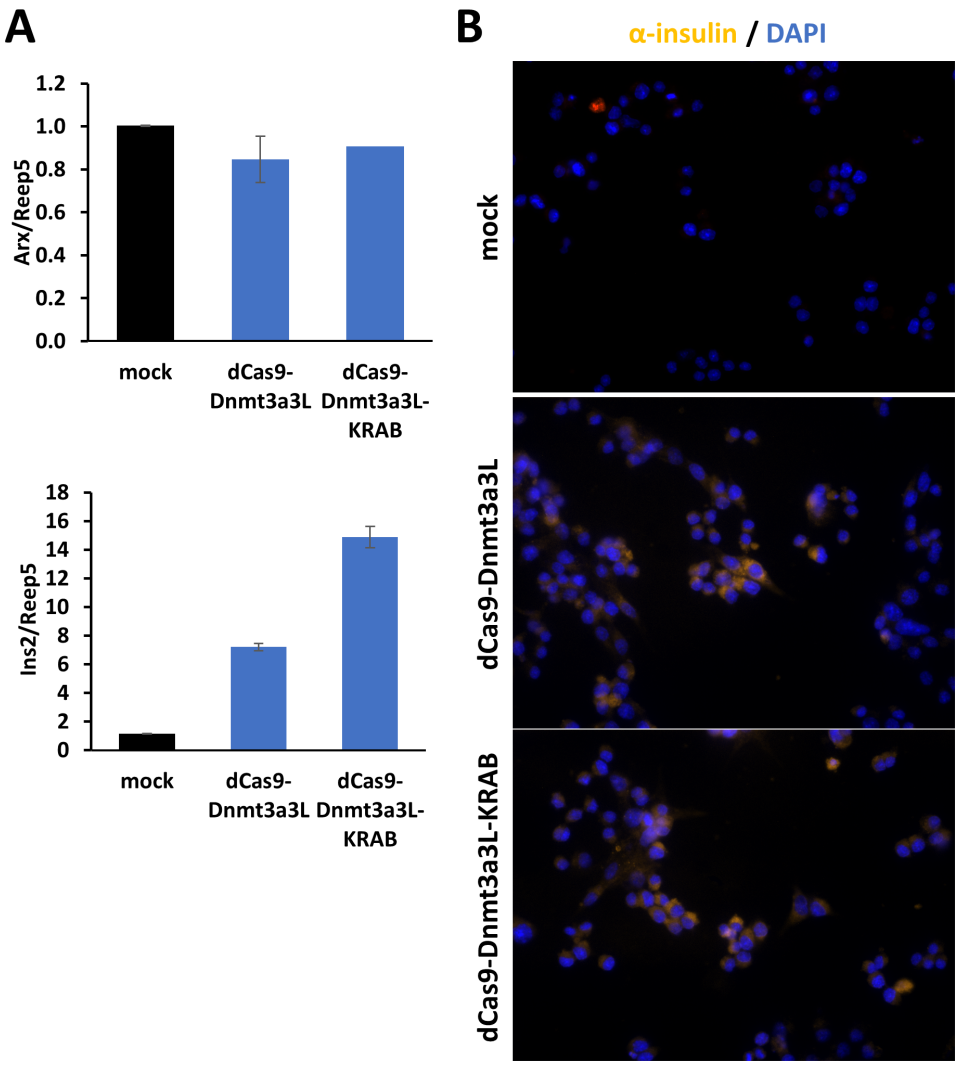

Figure 3.24: Epigenetically edited αTC1.6-cells produce insulin as demonstrated by experiments from our collaborators Marija Đordjević supervised by Dr. Melita Vidaković. **(A)** RT-qPCR analysis 9 dpt showed that *Ins2* expression is upregulated when targeting dCas9-Dnmt3a3L or dCas9-Dnmt3a3L-KRAB to the *Arx* promoter. However, while *Arx* was repressed with both effector domains 5 dpt, expression levels were back at the level of the mock-transfected sample for the shown time point 9 dpt. n=2, except for the dCas9-Dnmt3a3L-KRAB sample for the Arx expression analysis, which comprises only one biological sample. Error bars denote SEM. **(B)** Expression of insulin can be confirmed by staining transfected and edited α-cells with the α-insulin antibody (H-86) (sc-9168) and Texas Red conjugated α-rabbit antibody (sc-3841) at 9 dpt. Additionally, cells were stained with DAPI (blue). The microscopy images were taken at 40-fold magnification with an Axiocam digital camera on the Axio Observer Z1 microscope (Carl Zeiss Microscopy).

Our collaborators at the University of Belgrade additionally stained transfected α-cells using antibodies against insulin (see figure 3.24B). However, for their studies they used the slightly different αTC1.6 cell line instead of αTC1.9-cells. The following experiment was performed by Marija Đordjević supervised by Dr. Melita Vidaković.

Marija Đordjević transfected αTC1.6-cells as described above but used PEI instead of FugeneHD. Cell staining for insulin and imaging was then performed as described in section 2.2.2. The produced insulin in the cells was clearly visible for the samples transfected with dCas9-Dnmt3a3L and dCas9-Dnmt3a3L-KRAB targeted to the *Arx* promoter but was absent for mock transfected cells.

The expression analysis via qPCR confirmed this result. For this, Marija Đordjević isolated RNA from cells enriched for transfected cells via FACS 9 dpt and isolated RNA with the ZR-Duet DNA/RNA MiniPrep kit (Zymo Research). cDNA was prepared from 150 ng DNase I-digested RNA with the RevertAidTM First Strand cDNA Synthesis Kit (Thermo Scientific) using a mix of oligo-(dT) and random hexamer primers. The Maxima SYBR Green/ROX qPCR Master Mix (Thermo Scientific) was used for qPCR analysis with a QuantStudio 3 Real-Time PCR system (Applied Biosystems) in technical duplicates from two biological repeats. Her qPCR primers were different from ours and are listed in appendix table 5.1 with the suffix "Em". The expression of the *Ins2* gene was increased by 7.2$\pm$0.3-fold and 14.9$\pm$0.8-fold with dCas9-Dnmt3a3L and dCas9-Dnmt3a3L-KRAB, respectively (see figure 3.24A). Interestingly, while they did observe repression of the *Arx* gene 5 days past transfection (data not shown), the repression was already gone at the time point 9 dpt shown here. Consequently, expression of *Ins2* also decreased over time and was reduced to the level of mock-transfected cells 20 dpt (data not shown).

The results from our two groups demonstrated that cell transdifferentiation via targeted epigenome editing, especially of master regulators like the *Arx* gene, can work well. However, the lack of long-term stability of the introduced changes in cell type and gene expression show that a lot of improvements are still necessary before potential therapeutic approaches using this method could find their way into the clinic, even when disregarding safety concerns and the difficulty of delivering the ribonucleoproteins.

3.6 Targeted gene activation via novel engineered DNA demethylases

I had successfully demonstrated efficient targeted DNA methylation with CRISPR-Cas9 and the chimeric Dnmt3a3L fusion in section 3.1 and the resulting publication (Stepper et al., 2017). The obvious question was whether it was also possible to remove DNA methylation by targeting epigenetic effector domains. The first choice for this task were the TET enzymes, which can oxidize the methyl group of 5mC in a step wise process, resulting in the removal of the methylated base via TDG and the base excision pathway. However, many of the targeted DNA demethylation studies until then had not achieved strong effects, both on the DNA methylation level for multiple CpG sites and on the expression of the targeted genes (e.g. Chen et al., 2014; Choudhury et al., 2016), while others were more effective (e.g. Morita et al., 2016, using a SunTag-based approach).

We were therefore pursuing multiple possible avenues to boost the effect size: a) the addition of ascorbate to the cell culture medium to enhance TET activity via Fe(III) to Fe(II) reduction; b) the use of novel effector domains like a TET variant from the heterolobosean amoeboflagellate *Naegleria gruberi* (nTet), which exhibits much higher oxidation activity *in vitro* (Hashimoto et al., 2014); c) using many gRNAs to tile and completely edit the whole

promoter CpG island and d) mutating, engineering and combining effector domains for higher efficacy.

The following experiments were performed in HEK293 cells as they are very easy to work with and transfect. Additionally, the *EPCAM* gene promoter, which was used in section 3.1 for targeted DNA methylation and repression in SKOV-3 cells, is methylated and repressed in HEK293 cells, making it an ideal target for activation studies.

3.6.1 Targeting TET enzymes leads to DNA demethylation and gene expression

Targeting TET enzymes leads to DNA demethylation

Preliminary experiments did not show an effect on *EPCAM* expression when targeting nTet to its promoter (data not shown). Therefore, I constructed a longer Gly-Ser-based linker to increase the linker length between dCas9 and the effector domain from 28 amino acids to 48 amino acids to ensure that no sterical clashes or difficulties reaching the DNA could be the problem. This longer linker was later used for all activating effector domains, which is why the constructs with the shorter linker are designated "sl".

In the experiment shown here, several effector domains were targeted with a mix of all 13 EPCAM gRNAs to the methylated *EPCAM* promoter in HEK293 cells. In addition to the catalytic domains of the murine Tet1 and Tet2 proteins, the highly active, full-length nTet protein from *Naegleria gruberi* was transfected with both the shorter and the longer linker. Furthermore, the murine activation induced cytidine deaminase (AICDA/AID, amino acids 1-169) had been implicated in DNA demethylation (Santos et al., 2013) and was tested for potential expression activating effects, which might also result from the DNA repair response to its deaminase activity. As a positive control for gene activation, the VP64 and VPR effectors were included. These activators fused to dCas9 were transfected in a set with ascorbate included in the culture medium and a set without its addition. This allowed us to check whether the increased activity of Fe(II)-dependent oxygenases like TETs, which could potentially be recruited to activated genes, would increase the effect size. 100 µM Na-ascorbate was added while changing media after the transfection with PEI and on days 3 and 4 past transfection to all other samples, except the untreated HEK293 used for normalization.

The samples were harvested 5 dpt and RNA was isolated using the RNeasy Mini Plus kit (Qiagen). RT-qPCR of *EPCAM* expression relative to *ACTB* showed a weak activation of 2.1±0.2-fold when dCas9-AID was targeted to the promoter, and a 3.5±0.3-fold increase with dCas9-sl-mTet1 (see figure 3.25A). On the other hand, dCas9-sl-mTet2 and both version of dCas9-nTet did not lead to an increase in expression above 1.6-fold. Targeting dCas9-VP64 without and with the addition of ascorbate led to an increase in *EPCAM* expression of 4.3±0.4 and 3.9±0.5-fold, respectively, while dCas9-VPR achieved the strongest effect of 23.5±7.2 without ascorbate and 9.9±0.6 with ascorbate.

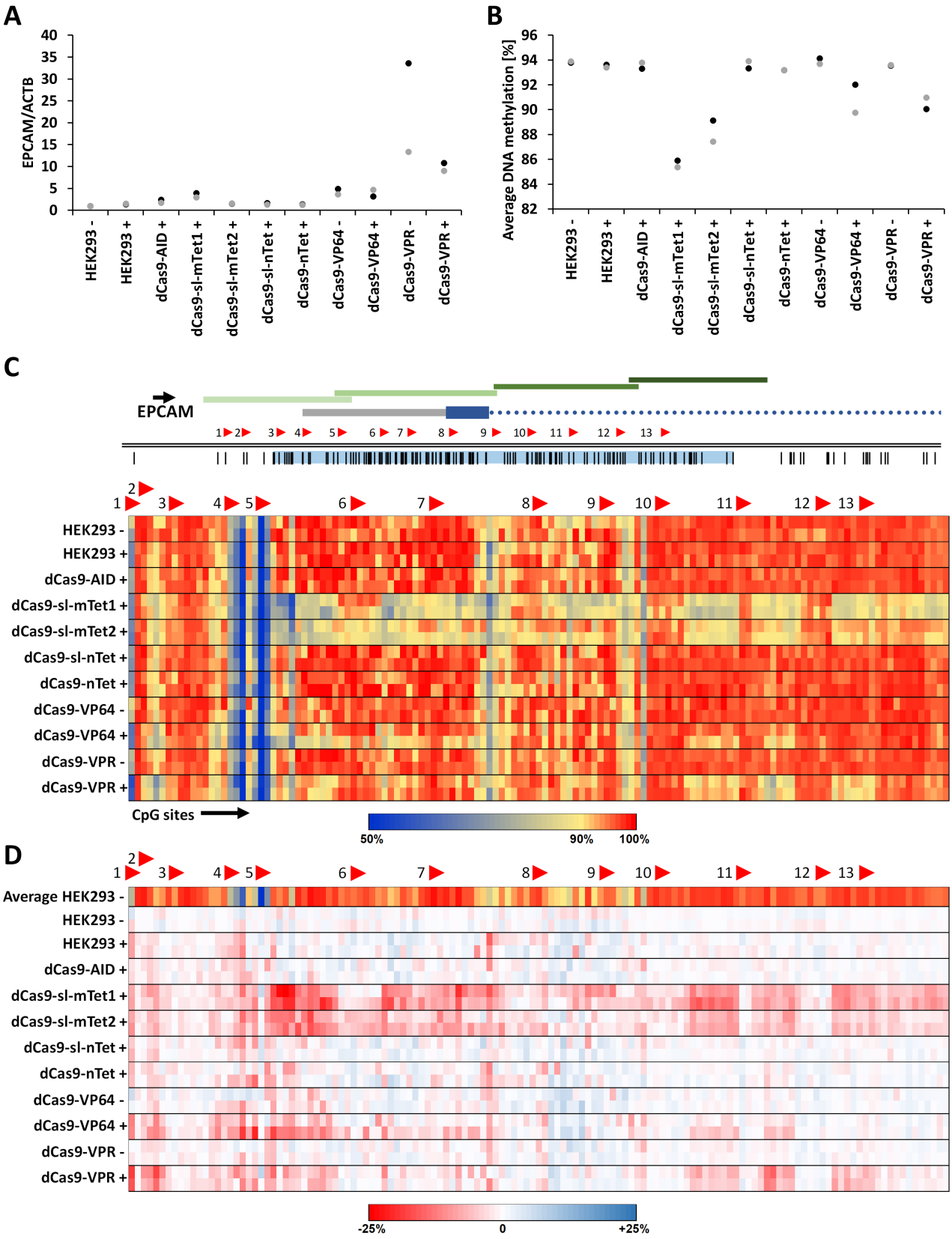

Figure 3.25: Targeting activating domains to the *EPCAM* promoter in HEK293 cells. **(A)** For some of these effector domains, this resulted in an upregulation of *EPCAM* as assayed by RT-qPCR 5 dpt. The + and - signs after the sample name denote the addition of 100 µM Na-ascorbate to the cell culture medium and its omission, respectively. The black and grey dots represent the average of the technical repeats in qPCR for the two biological replicates for each sample. **(B)** Average DNA methylation at the *EPCAM* promoter via NGS bisulfite amplicon sequencing. **(C)** Schematic of the *EPCAM* locus and heat map representation of the bisulfite sequencing results. See figure 3.2 for a more detailed description. **(D)** Difference in percentage points to the average of the two replicates of untransfected HEK293 cells without the addition of ascorbate. For reference, the average of the two HEK293 - replicates is shown at the top using the color scale from panel B.

Additionally, I isolated genomic DNA from the harvested cells and analyzed changes in the amount of DNA methylation present at the *EPCAM* promoter via bisulfite conversion, amplification and sequencing of the locus. The procedure was performed as in section 3.1 but with updated bisulfite primers (see appendix table 5.6). The *EPCAM* promoter was almost completely methylated in untreated HEK293 cells (93.84%±0.03%), which barely changed with the addition of ascorbate (93.52%±0.07%, see figure 3.25B). While targeting dCas9-AID to the promoter did not lead to changes in the amount of DNA methylation (93.55%±0.17%), both dCas9-sl-mTet1 and -mTet2 lead to a substantial loss of DNA methylation to 85.64%±0.19% and 88.28%±0.60%, respectively. This means that targeting dCas9-sl-mTet1 to the *EPCAM* promoter more than doubled the percentage of unmethylated CpG sites. This correlates to the increase in *EPCAM* expression as measured by qPCR. However, the dCas9-sl-mTet2 targeted cells do not show increased *EPCAM* expression despite exhibiting a loss of DNA methylation at the promoter. Of the remaining samples, the two nTet versions and the VP64 and VPR samples without the addition of ascorbate had no meaningful reduction of DNA methylation content. Strikingly, this was different for the dCas9-VP64 and dCas9-VPR samples with ascorbate in the media, which had DNA methylation levels of 90.90%±0.79% and 90.48%±0.33%, respectively. Despite this, these samples did not produce more *EPCAM* mRNA than the samples without ascorbate. Taken together, these results suggested that the amount of DNA demethylation is not the most important factor governing gene expression in this specific setting.

A heat map with the average methylation level for each CpG site is shown in figure 3.25C. Additionally, the change in DNA methylation in percentage points per site relative to the average of both biological replicates of the "HEK293 -" cells is presented as a separate panel in figure 3.25D. For the samples targeted with dCas9-sl-mTet1 and -mTet2, the methylation loss was evenly spread over the whole locus. However, especially in panel D a few hot spots were apparent, especially downstream of sgRNAs 6, 7 and 11. Interestingly, at least some of the gRNA binding sites protected the covered CpG sites from demethylation, as also observed when targeting DNA methylation with dCas9-Dnmt3a3L (see figures 3.2 and 3.4). Examples for this observation are sgRNAs 6 (which covers 6 CpGs), 11 and 12. Notably, the samples with dCas9-VP64 and VPR with the addition of ascorbate also showed clustered DNA demethylation hot spots close to the gRNA binding sites, albeit at a lower level, except for sgRNA 3.

In conclusion, both the dCas9-mTet fusions and dCas9-VP64 and dCas9-VPR lead to the removal of DNA methylation at the targeted promoter. However, the demethylation did not necessarily translate to an increase in *EPCAM* transcription.

Target gene activation is delayed for TET enzymes

Subsequently, we wanted to investigate the kinetics of targeted gene activation with the different effector domains. Therefore, I conducted an experiment together with Liv Paul, a student I supervised for her BSc thesis. In this experiment, we continually harvested the

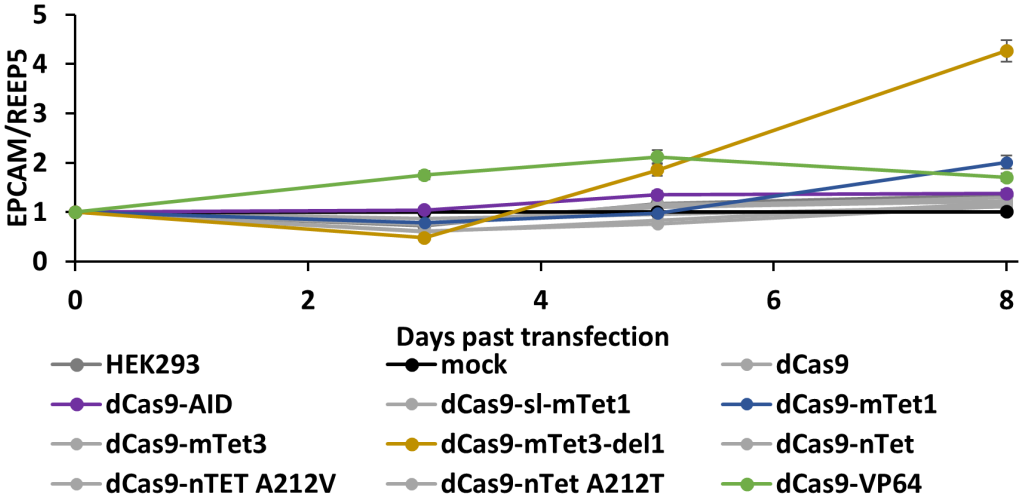

Figure 3.26: TET enzymes lead to a delayed increase in target gene expression. This time course experiment targeting *EPCAM* in HEK293 cells revealed that the engineered catalytic domain of mTet3, dCas9-mTet3-del1, can activate *EPCAM* expression starting 5 dpt and further increasing at 8 dpt. By then it strongly surpassed the activation achieved with dCas9-VP64, which achieved only transient target gene activation. Most other samples are greyed out for better visibility of the samples showing an effect. The experiment was performed with a single biological replicate, error bars denote technical standard deviation.

samples while culturing them for 8 days after transfection. Because we had noticed that adding Na-ascorbate to the cells immediately after passaging was toxic for them, we did not add any ascorbate to the cells at all for this experiment. Only a single biological replicate was used for this experiment for faster data generation, meaning these results should rather be regarded as exploratory.

In addition to the effectors used in the experiment above, we had generated a new version of mTet1 where the catalytic domain was fused to dCas9 with the longer linker initially developed for nTet. Furthermore, we transfected both the wild-type catalytic domain of murine Tet3 fused to dCas9 and a variant engineered for higher stability and activity. This version with an internal deletion was developed by Dr. Mirunalini Ravichandran and designated mTet3-del1 (Ravichandran, 2017). As the TET enzyme from *Naegleria gruberi* did not lead to either DNA demethylation or target gene activation, we additionally tested mutants of nTet. The A212V mutant drastically reduced the oxidation rate of 5hmC, leading to its accumulation (Hashimoto et al., 2014), which might boost target gene activation.

The analysis via RT-qPCR revealed that 3 days past transfection only dCas9-VP64 increased *EPCAM* expression, while dCas9-mTet3-del1 even repressed its expression (see figure 3.26). However, already on day 5 the sample transfected with dCas9-mTet3-del1 exhibited upregulated *EPCAM* expression (1.9-fold), almost as strong as the dCsa9-VP64 sample (2.1-fold). Strikingly, the target gene expression increased even further to 4.3-fold with the new dCas9-mTet3-del1 construct. Similarly, the dCas9-mTet1 sample also showed elevated *EPCAM* expression on this day (2.0-fold). Interestingly, this was the case even though the transfected cells were already barely fluorescing at that time point, indicating that both the fluorescent reporter plasmid mVenus and the dCas9-ED and gRNA plasmids

were either lost or silenced by then.

While target gene expression increased for the targeted TET enzymes despite the strong reduction of expression of the transfected plasmids, *EPCAM* expression was reduced from 2.1-fold to 1.7-fold activation for dCas9-VP64. This observed loss would be the expected result if target gene activation was depending on the continued presence of the effector. However, for dCas9-mTet1 and dCas9-mTet3-del1, target gene activation was decoupled from effector presence after the initial stimulus.

Furthermore, as both the wild-type version of mTet3 and the mTet1 construct with the original, shorter linker, cannot increase target gene expression in this experiment, this also highlights the importance of rational designing and optimizing effector domains.

Similarly to the experiment above, dCas9-AID lead only to a modest 1.4-fold increased expression of EPCAM on days 5 and 8. Finally, both the wild-type and mutated dCas9-nTet constructs failed to show an effect on target gene expression, hinting at potential problems of this effector domain in mammalian cells. Both dCas9-AID and nTet were not used in subsequent experiments.

3.6.2 Targeting TET enzymes can achieve stable gene activation

After the successful experiments in the previous section 3.6.1, which demonstrated target gene activation and promoter demethylation, Liv Paul and I continued investigating the intriguing result that target gene activation was increasing over time, even after the dCas9-TET fusion protein should have mostly stopped being expressed.

Therefore, we transfected HEK293 cells with the mix of 13 gRNAs targeting the *EPCAM* promoter and the effector domains showing the strongest effects so far: dCas9-mTet1, dCas9-mTet3-del1 and dCas9-VPR. A sample containing dCas9 without any fused effector domain and untreated HEK293 cells served as negative controls and all results were normalized to mock transfected samples. For this experiment, each sample was performed once with and without the addition of 25 μM Na-ascorbate, which was always added at least 6 h after harvesting or passaging the cells to minimize toxic effects.

RT-qPCR analysis of *EPCAM* expression for the first time point 2 days past transfection showed the expected strong activation for the dCas9-VPR samples (see figure 3.27). Again, the ascorbate-treated sample exhibited the stronger activation (10.9±0.4 vs. 8.7±0.3; average of two biological replicates±SEM). No other effector domain achieved target gene activation already at this time point. However, 4 dpt transfection the activation effect with VPR was already declining (9.2±1.6 and 6.7±0.9 with and without ascorbate, respectively).

Conversely, the ascorbate-treated samples of dCas9-mTet1 and dCas9-mTet3-del1 began to induce activation of *EPCAM*, reaching 1.6±0.2 and 3.3±0.7-fold activation, respectively. The next time point 7 dpt was the first on which the TET samples surpassed the dCas9-VPR transfected samples, which had decreased dramatically to 2.3±0.3 and 1.4±0.0, respectively. Furthermore, the dCas9-mTet3-del1 sample without ascorbate increased *EPCAM* expression

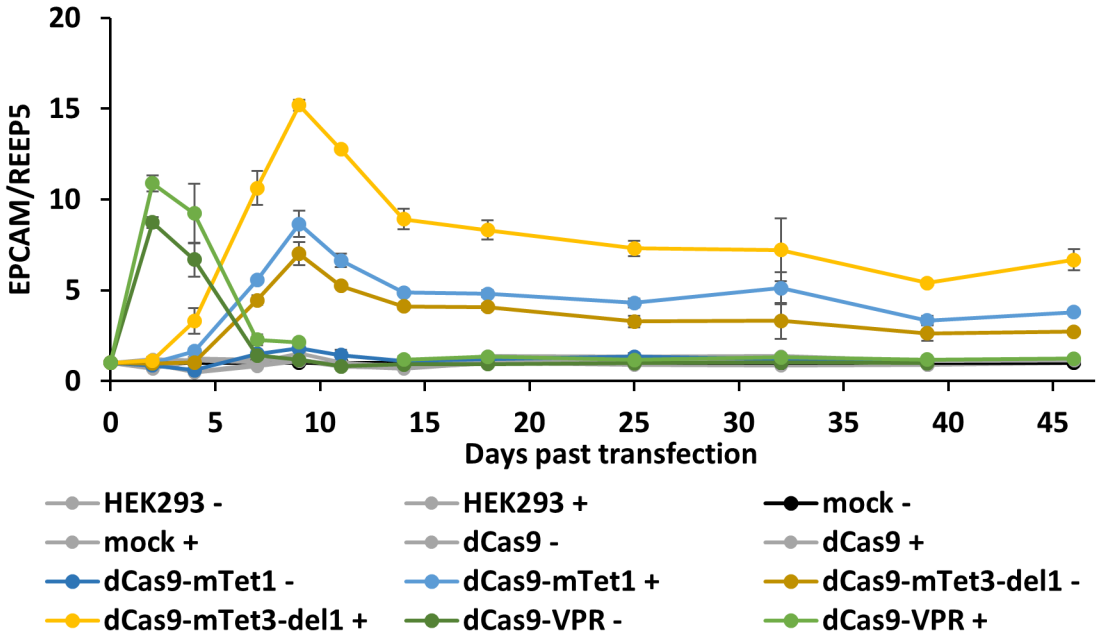

Figure 3.27: Stable activation of the target gene *EPCAM* via targeted DNA demethylation in HEK293 cells. The +/- behind sample names denote addition and omission of 25 µM ascorbate. Whereas dCas9-VPR achieved a strong initial increase in *EPCAM* expression, the effect declined quickly at 7-11 dpt. The TET enzymes, on the other hand, activate EPCAM only starting 4-9 dpt, with dCas9-mTet3-del with added ascorbate even surpassing the initial peak of dCas9-VPR at 9 dpt. Even though the effect size decreases for the next time points the activation remains mostly constant until 46 dpt, demonstrating stable gene activation by "hit-and-run" epigenetic editing. Most other samples are greyed out for better visibility of the samples showing an effect. Error bars denote SEM of the two biological replicates.

4.5±0.3-fold at this time point, demonstrating an even more delayed effect than the sample with ascorbate, but showing almost the same effect size as dCas9-mTet1 with ascorbate.

TET-induced activation reached its peak 9 dpt. At this time point, the dCas9-mTet3-del1 sample with added ascorbate (15.2±0.3-fold activation) even surpassed the strongest activation of dCas9-VPR with ascorbate observed 3 dpt. Similarly, dCas9-mTet1 without added ascorbate exhibited its peak at this time point at only 1.8±0.1-fold activation, while the sample with ascorbate reached 8.6±0.7-fold activation, higher than the dCas9-mTet3-del1 sample without ascorbate (7.0±0.6-fold activation).

After the peak 9 dpt, TET-induced activation of *EPCAM* dropped sharply until 14 dpt but, remarkably, remained mostly stable for the rest of the experiment. At the last time point 46 dpt, dCas9-mTet3-del1 with ascorbate still showed 6.7±0.6-fold increased *EPCAM* expression, followed by dCas9-mTet1 with ascorbate and dCas9-mTet3-del1 without ascorbate (3.8±0.1 and 2.7±0.1-fold activation, respectively). At these late time points, no other sample showed increased *EPCAM* expression, with the dCas9-VPR samples having completely dropped back to the levels of the control samples starting 15 dpt.

The stable activation for these samples persisted 46 days after transfection, long after the plasmids where cleared from the cells, hinting at a stable switch of the epigenetic state of the promoter and gene expression.

3.6.3 Synergistic activation by combining VPR and TET enzymes

We hypothesized that, similar to the successful fusion of KRAB to dCas9-Dnmt3a3L in section 3.2.2, a fusion of the fast but transient activator VPR to the engineered mTet3 DNA demethylases might achieve stronger effects by combining the two distinct mechanisms of action. Therefore, I cloned and tested multiple permutations to select the most active combination. VPR was combined with both mTet3-del1 and another improved variant with a C-terminal truncation mTet3-del1ΔC, respectively. mTet3-del1ΔC was designed and characterized by Dr. Mirunalini Ravichandran and was shown to be even more active than mTet3-del1 (Ravichandran et al., manuscript in preparation). This protein is abbreviated as 'mTet3h' for hyperactive. Initially, VPR was cloned to the C-terminal end of the dCas9-mTet3-del1 construct, as this was the most active configuration for the dCas9-Dnmt3a3L-KRAB fusion. However, this combination exhibited less than 10% of the effect size of dCas9-VPR alone (data not shown), hinting at folding or sterical problems and was therefore abandoned. Instead, I tested fusing VPR to the N-terminus of the dCas9-mTet3 constructs and cloning mTet3-del1 and mTet3h to the C-terminus of dCas9-VPR. Unfortunately, the combination dCas9-VPR-mTet3h was not completed in time for the experiment discussed in this section.

To test the new fusion proteins, I performed an experiment in HEK293 cells targeting either *EPCAM* with a mix of all 13 gRNAs or a new target gene, *LIN28A*, with a mix of LIN28A sgRNAs 1-6. Cells were transfected in single biological replicates and harvested and passaged on days 3, 5, 7 and 9 past transfection. To boost the demethylation activity, I added 25 μM ascorbate to all samples starting 1 dpt. However, to check the influence of ascorbate addition on the effect size, I included an additional sample targeting dCas9-mTet3-del1 to the *EPCAM* promoter to which I added ascorbate only on the first and third days past transfection, but not on the later time points.

As target gene activation via the TET proteins was strongest 9 dpt in previous experiments, I analyzed the samples for that time point (see figure 3.28). As expected for *EPCAM* at this time point, the cells transfected with dCas9-VPR showed almost no increase in target gene expression relative to the mock transfected sample. Of the samples transfected with only a single effector domain, dCas9-mTet3h resulted in the strongest increase in *EPCAM* expression (11.8±0.7-fold), followed by the two dCas9-mTet3-del1 samples. Interestingly, the sample that had ascorbate added only on days one and three past transfection, but not afterwards, exhibited stronger target gene activation than the sample with continually added ascorbate. However, I nevertheless added ascorbate to the cells for all time points in subsequent experiments, as this experiment was carried out with only a single biological replicate to quickly check the activity of the new fusion proteins. Even so, this result suggested firstly, that the difference in activity between mTet3-del1 and mTet3h might not be as drastic in cells as *in vitro*; and secondly that the catalytic activity of the TET proteins might only be important during the first few days after transfection, even though strong effects on target gene expression appear only later.

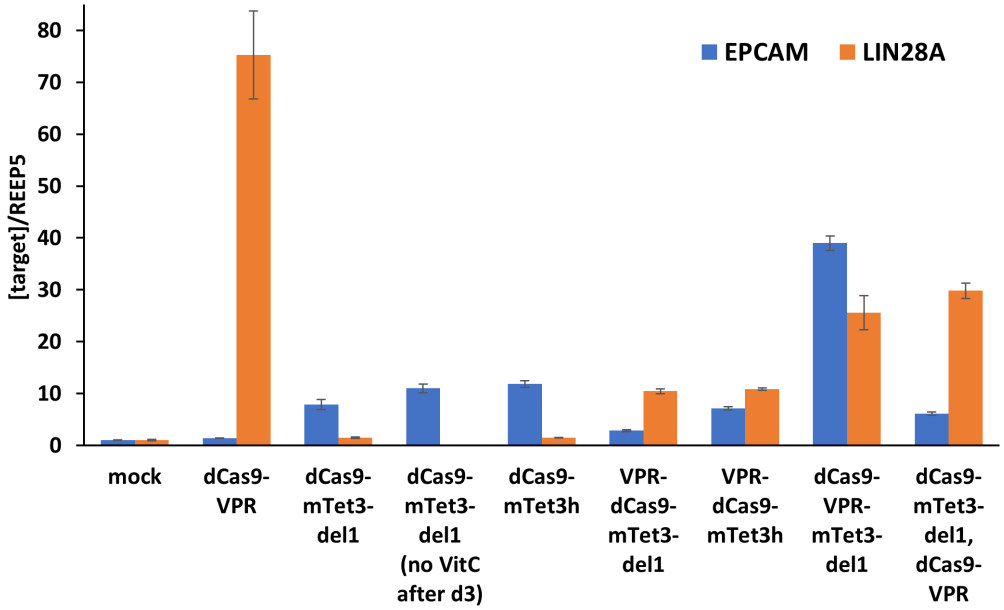

Figure 3.28: Testing new effector domain fusions for targeted gene activation. Effects on target gene expression 9 dpt when targeting *EPCAM* or *LIN28A* in HEK293 cells. At this time point, *EPCAM* was more susceptible to TET enzymes than dCas9-VPR and exhibited the strongest activation when targeted by the new fusion protein dCas9-VPR-mTet3h, which demonstrated synergistic effects compared to both the single effector domains on their own or when co-targeted. *LIN28A*, on the other hand, did not show increased expression when targeted by TET-enzymes, but was still activated by dCas9-VPR on this time point. n=1, error bars denote standard deviation of technical replicates.

By far the strongest activation of *EPCAM* expression was achieved with the fusion of mTet3-del1 via a 35 amino acids long linker to the C-terminus of VPR protein: dCas9-VPR-mTet3-del1. The 39.0±1.4-fold activation was much stronger than the summed effects of the two single effector domains, which suggests synergistic effects of the two mechanisms of action. Interestingly and unexpectedly, when co-transfecting the constructs for dCas9-VPR and dCas9-mTet3-del1, the resulting effect was much smaller than with the fusion protein, even when accounting for the fact that only half the amount of protein will be produced as the DNA contingent for dCas9-ED plasmids was split between the two plasmids.

Arranging the same effector domains in a different order (VPR-dCas9-mTet3-del1 and VPR-dCas9-mTet3h) lead to much weaker effects than with mTet3-del1 fused to the C-terminus of dCas9-VPR. This points to the importance of linker length, the position of the N- and C-termini and steric clashes.

The efficacies of the individual effector domains were almost reversed for the other target gene, *LIN28A* (see figure 3.28B). This gene barely responded to the targeting of TET enzymes, but was upregulated strongly by dCas9-VPR even 9 dpt. The best fusion protein here was again dCas9-VPR-mTet3-del1, as for *EPCAM*. In contrast to *EPCAM*, co-targeting dCas9-VPR and dCas9-mTet3-del1 to LIN28A resulted in at least the same increase in target gene expression for LIN28A as with the best fusion of these two effector domains. This result might be connected to the fact that *EPCAM* responds better to targeting TET enzymes, whereas *LIN28A* is much more sensitive to VPR.

3.6.4 Success of targeting gene activation depends on the right combination of effectors and individual epigenetic state

The previous time course experiment in section 3.6.2 revealed stable activation of *EPCAM* expression in HEK293 cells sustained for more than five weeks by targeting dCas9-mTet3-del1 to the *EPCAM* promoter. Afterwards, we had generated and validated both an engineered, hyperactive version of this effector domain, dCas9-mTet3h, as well as a fusion protein of the TET catalytic domain to the fast-acting but only transiently activating domain dCas9-VPR in section 3.6.3. This fusion exhibited strong synergistic effects on gene activation. Furthermore, I had introduced three new target genes in addition to *EPCAM*: *MAGEB2*, *LIN28A* and *ASCL1*. The following experiment was designed to answer the open questions of i) the behavior of these new target genes; ii) the kinetics and effect size of the dCas9-VPR-mTet3h fusion; and iii) the mechanism of both the strong and stable activation via the TET enzymes and its delayed onset.

Therefore, HEK293 cells were seeded in 6-well plates and transfected in triplicates. I included the samples shown on the left side in figure 3.29. As for the experiment in section 3.4.2, 5% of the transfected DNA was the eGFP-Puro plasmid for the enrichment of transfected cells via puromycin selection. However, in contrast to the other experiment, the four genes were not co-targeted. Because each promoter was targeted with 5 to 13 gRNAs, they were transfected as separate samples to ensure that enough gRNAs were available to achieve epigenetic editing of all target sites. The gRNA mixes contained EPCAM sgRNAs 1-13; MAGEB2 sgRNAs 1-5; LIN28A sgRNAs 1-6; and ASCL1 sgRNAs 1-5, respectively. The effector domains used in this experiment were dCas9-VPR, dCas9-mTet3h, its catalytically dead double mutant dCas9-mTet3h DM and the fusion dCas9-VPR-mTet3h. Furthermore, I wanted to test whether I could achieve stronger effects by combining VPR with untargeted, but simply overexpressed full length mTet 1,2 and 3 (mTet1-3 FL).

Like in section 3.4.2, 20% of the transfected DNA was used for the effector domain and 75% for the gRNA mixes, except for the samples with mTet1-3 FL, in which 15% of the gRNA DNA was exchanged for the mTet1-3 FL plasmids (5% each of total DNA). Puromycin at a concentration of 2.5 μM was added on the first and second days after the transfection to the transfected samples.

To boost the activity of TET enzymes, I again added 25 μM Na-ascorbate to the cell culture medium as in section 3.6.2. Control samples were cultured both with and without added ascorbate. Additionally, all samples transfected with a dCas9-ED included ascorbate in the cell culture medium. I collected cells of every sample every three days. Jule Focken, a student I supervised for her BSc thesis, isolated RNA according to BOMB protocol 8.1 and performed RT-qPCR to assess the expression of the target genes relative to *REEP5* for the time points 3 to 18 dpt. The results are presented as a heat map in figure 3.29.

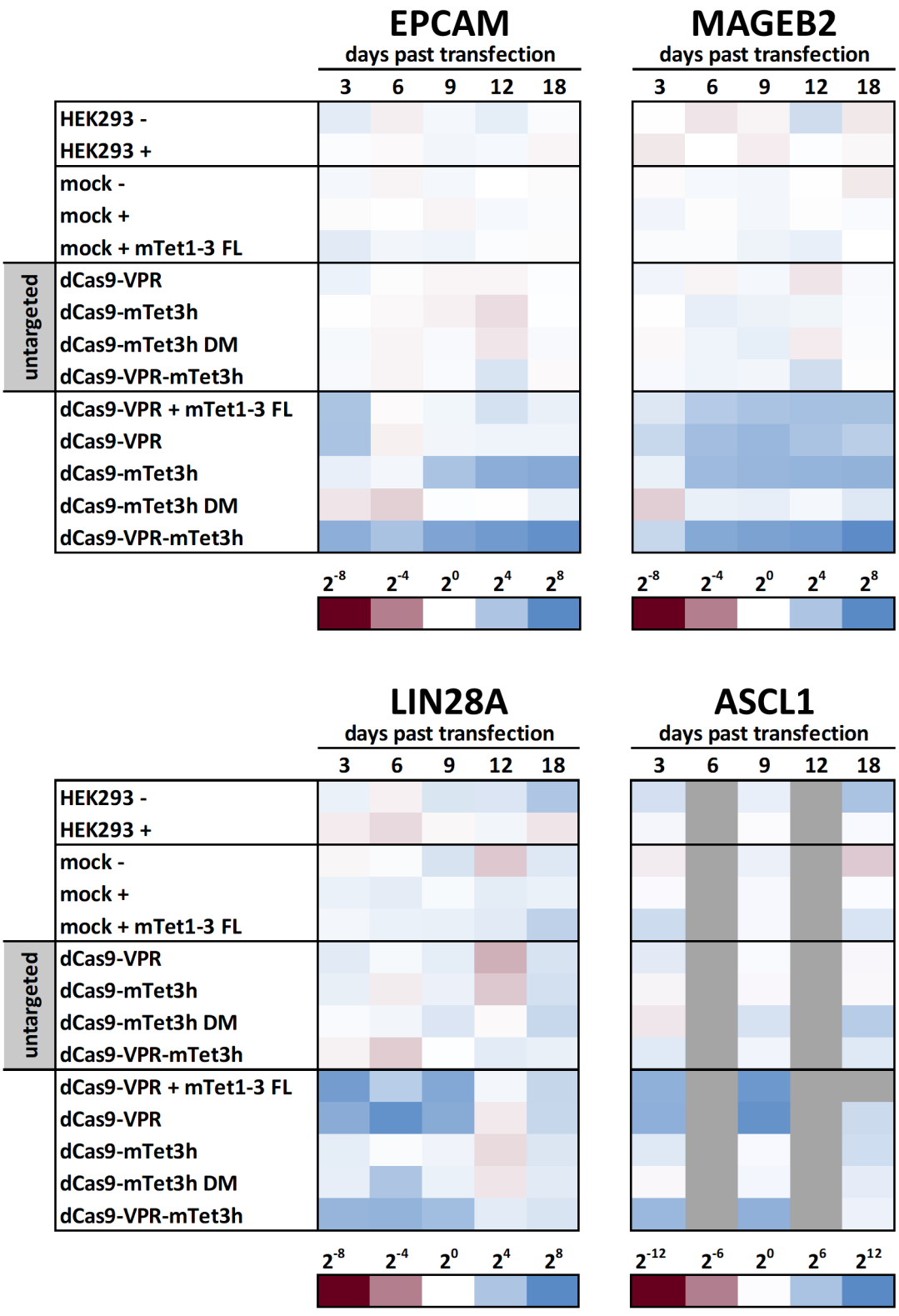

Figure 3.29: Heat map representation showing the RT-qPCR results of targeting four target genes with multiple effector domains. The table on the left lists the different dCas9-EDs transfected into HEK293 cells targeting the promoters of *EPCAM*, *MAGEB2*, *LIN28A* and *ASCL1*. For the untransfected and mock transfected samples, a "+" indicates that ascorbate was added to the cell culture medium, whereas a "-" signifies its omission. For all other samples, ascorbate was added as well. The color of each heat map cell denotes the average expression of this target gene for the time point indicated at the top with the color legend located at the bottom of the figure. Grey cells mean that no data was available for these samples. To better show smaller changes in gene expression, the expression values were log2 transformed. The target gene activation was normalized to the average Cq of both the HEK293 + and mock + samples. Generally, the average expression was taken from three biological replicates except for a few time points for which one replicate was removed for technical reasons

EPCAM

The target gene *EPCAM* responded similarly to previous experiments. On the first time point 3 dpt the dCas9-VPR samples, with and without the co-transfection of mTet1-3 FL, increased *EPCAM* expression to 17.0±3.1-fold and 17.8±3.6-fold activation, respectively. As seen before, the dCas9-mTet3h effector domain did not lead to a strong target gene activation on the first time point with just 2.2±0.4-fold expression compared to the control samples. Notably, the fusion dCas9-VPR-mTet3h achieved 48.5±2.7-fold activation of *EPCAM* expression, much more than the sum of the effects of dCas9-VPR and dCas9-mTet3h on their own. This was similar to the effect seen in section 3.6.3 9 dpt. But it is unclear, how mTet3h was able to boost the activation of VPR when mTet3h alone did not yet lead to increased EPCAM transcription at 3 dpt. And while the transcriptional activation of EPCAM declined on the next time point for both the dCas9-VPR samples as well as dCas9-VPR-mTet3h, *EPCAM* expression increased again on day 9 for the fusion protein while dCas9-VPR did not activate transcription on the following time points. This is shown again in more detail as line charts in figures 3.30A and B.

Just like the dCas9-VPR-mTet3h samples, the dCas9-mTet3h effector alone also led to increased *EPCAM* expression over time, rising from 18.2±3.4-fold and 1.5±0.1-fold expression 6 dpt to 193.9±13.7-fold and 59.6±6.6-fold activation, respectively. On the other hand, this demonstrated that targeting dCas9-VPR-mTet3h or dCas9-mTet3h again achieved stable activation of *EPCAM*, but at much higher levels than seen before in section 3.6.2. Additionally, target gene expression even continued to increase until at least 18 dpt in the case of dCas9-VPR-mTet3h. The catalytically dead mutant dCas9-mTet3h DM did not lead to changes in *EPCAM* expression, proving that the activating effect must be due to the catalytic properties of the effector domain. Furthermore, the transient expression peak achieved with dCas9-VPR confirmed the results of section 3.6.2. Increasing the general expression level of TET enzymes in the cell by co-transfecting full-length mTet1-3 did not boost or prolong the activating effects of VPR on *EPCAM* expression (see figure 3.30A).

MAGEB2

The target gene *MAGEB2* reacted differently to the effector domains compared to *EPCAM*. There was no strong initial peak of activation for dCas9-VPR or the fusion protein dCas9-VPR-mTet3h. They reached only 6.5±1.0-fold and 7.0±1.5-fold activation on the first time point 3 dpt. However, expression increased for these samples on the following days, just like for dCas9-mTet3h, instead of declining again already on day 6 at *EPCAM*. The strongest activation achieved by dCas9-VPR was 30.9±11.5 on day 9, which then declined again to 10.5±3.7 on the last time point 18 dpt. In contrast, the dCas9-VPR sample with the full-length mTet1-3 proteins showed a slower, but steady increase in *MAGEB2* expression, reaching 20.5±3.1-fold on day 18 (see figure 3.31A).

Similarly, dCas9-mTet3h and the fusion dCas9-VPR-mTet3h both achieved higher upregulation of *MAGEB2* for every subsequent time point, culminating in 40.2-fold and even

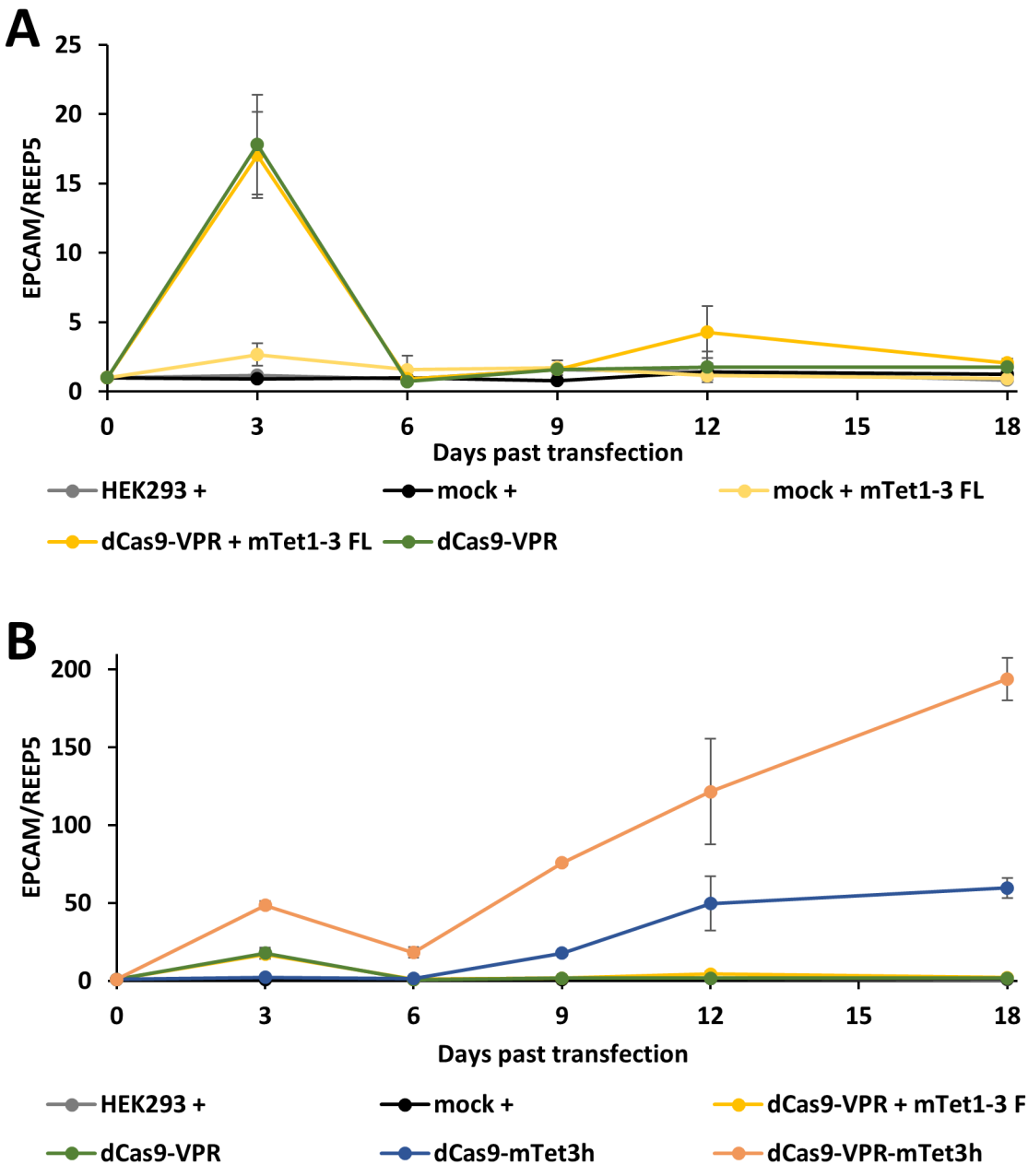

Figure 3.30: Kinetics of *EPCAM* expression when targeted with different effector domains. **(A)** Line chart showing only the dCas9-VPR samples, while **(B)** compares also the dCas9-mTet3h and dCas9-VPR-mTet3h fusions. Generally, the average expression was taken from three biological replicates except for a few time points for which one replicate was removed for technical reasons. Error bars denote SEM.

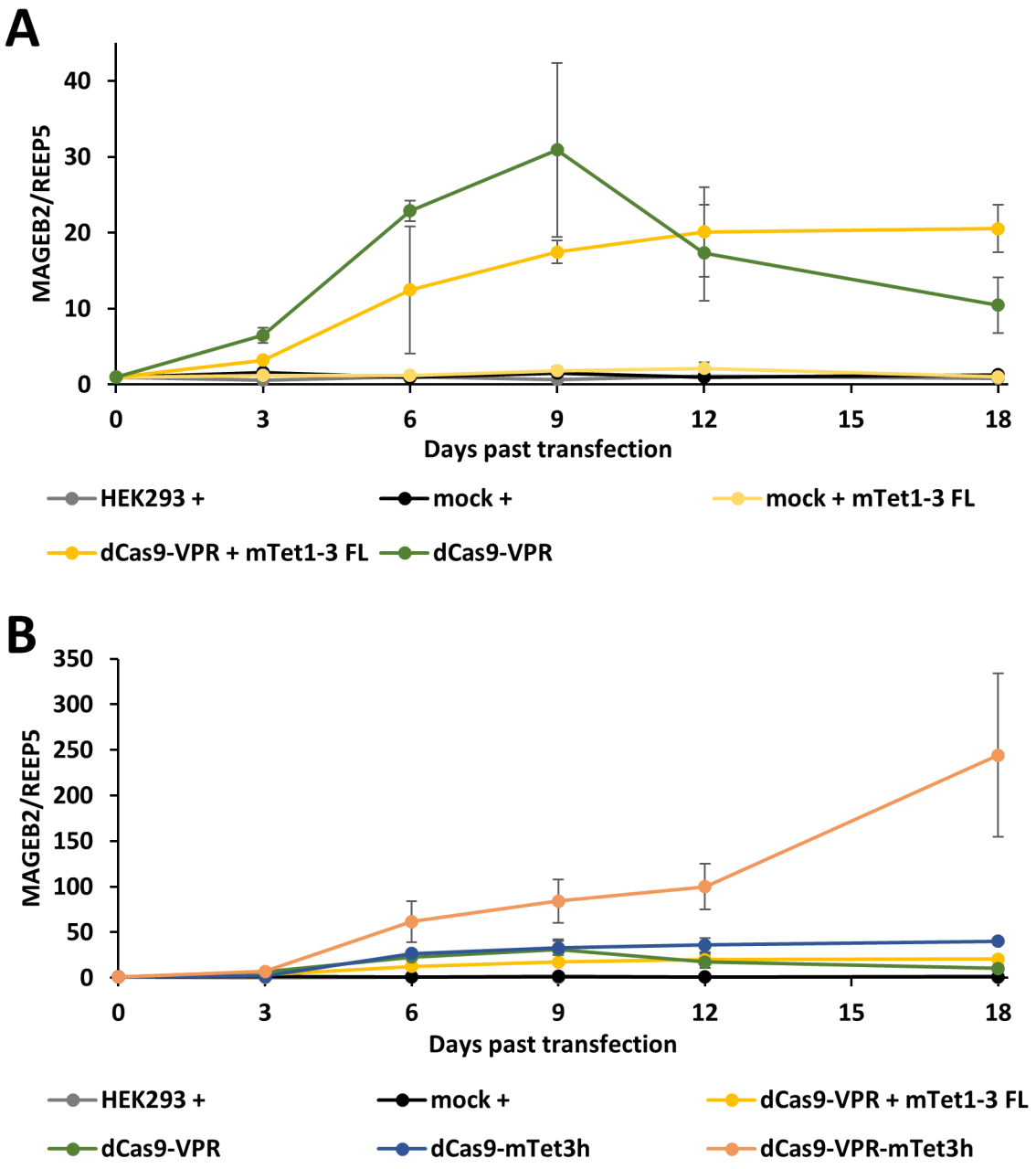

Figure 3.31: Kinetics of *MAGEB2* expression when targeted with different effector domains. **(A)** Line chart showing only the dCas9-VPR samples, while **(B)** compares also the dCas9-mTet3h and dCas9-VPR-mTet3h fusions. Generally, the average expression was taken from three biological replicates except for a few time points for which one replicate was removed for technical reasons. Error bars denote SEM.

244.2-fold activation on the last time point, respectively (see figure 3.31B). Again, the new fusion dCas9-VPR-mTet3h achieved by far the strongest and most stable increase in target gene expression. The activation with dCas9-VPR was slower, but more persistent than at the *EPCAM* promoter. However, the effect still declined over time. Interestingly, co-transfecting the full-length mTet proteins might have achieved stable *MAGEB2* activation in conjunction with dCas9-VPR, albeit at a lower level or in fewer cells than with the dCas9-VPR-mTet3h fusion.

LIN28A and ASCL1

Variability was higher for the other two target genes, *LIN28A* and especially *ASCL1*. This was the case partially because these genes have a very low basal expression in HEK293 cells, making RT-qPCR analysis difficult and increasing the variability in case of imperfect RNA isolation yields. Nevertheless, the results allowed us to discern interesting trends: Mainly, that both *LIN28A* and *ASCL1* barely responded to targeted DNA demethylation via TET enzymes, something previous experiments had already shown for *LIN28A* (see section 3.6.3) and *ASCL1* (data not shown).

The maximum activation achieved with dCas9-mTet3h was only 3.3±1.0-fold for *LIN28A* on day 18 and 14.0±8.6-fold for *ASCL1* 6 dpt. And while the fusion dCas-VPR-mTet3h did lead to stronger activation (37.4±15.7-fold for *LIN28A* and 2337.4±245.2-fold for *ASCL1*, both times on day 6), that was still less than with dCas9-VPR. For LIN28A, the sample dCas9-VPR + mTet1-3 FL reached its peak activation already 3 dpt with 109.1±19.7, while dCas9-VPR alone achieved 190.1±106.3-fold activation on day 6. However, *LIN28A* expression declined again for both dCas9-VPR samples to just 7.1±3.6-fold and 6.9±3.0-fold activation on day 18. The results where similar for *ASCL1*, except that the achieved activation was much higher, reaching up to 3249.8±1109.1-fold activation with dCas9-VPR on day 6. This might be because the gene was barely expressed in untreated HEK293 cells, thereby allowing a stronger effect. However, already for these selected results the errors were big due to the mentioned technical difficulties. Therefore, these experiments should be repeated with higher cell numbers to yield more RNA for a higher quality of the results.

Just like in section 3.6.3, these results again suggested that the VPR component in the fusion protein dCas9-VPR-mTet3h cannot function just as well as without the fused TET catalytic domain, making the stronger activation at the *EPCAM* and *MAGEB2* promoters even more striking. Furthermore, the different responses of the four targeted genes to the precise epigenome perturbations highlight the importance of the epigenetic state of the target for the success of epigenetic editing.

Kinetics of dCas9-ED plasmid expression

As seen before in section 3.6.2, the transcriptional activation achieved with targeted TET enzymes started relatively late between 3 and 9 dpt. While the loss of fluorescence from the eGFP-Puro plasmid around 5 dpt already showed that the expression of the transfected plasmids was strongly reduced by that time point, I wanted to check this for the dCas9-ED plasmids as well.

Therefore, I designed two pairs of primers amplifying either the dCas9-ED plasmid or only the dCas9 gene located on the plasmid (Q-pC-Cas9 and Q-SPn-Cas9, respectively, in appendix table 5.1). By subtracting the expression of the plasmid from the dCas9-mRNA, I could calculate the expression of only the mRNA. By comparing this signal to the *REEP5* expression, as for the target gene expression, I tracked the decline of dCas9-ED expression over time. The resulting chart is displayed in figure 3.32.

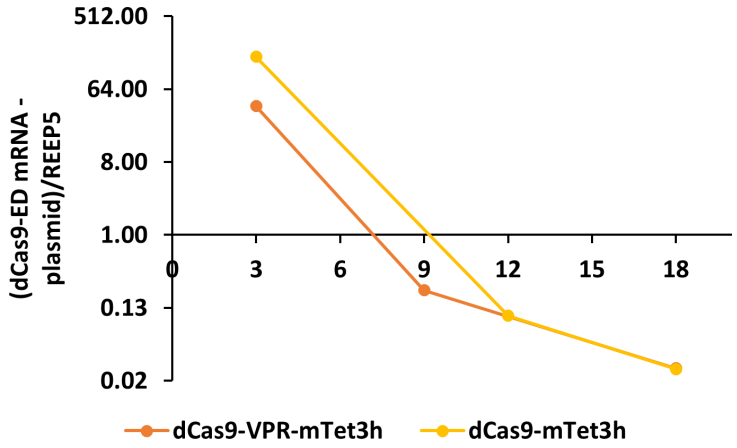

Figure 3.32: Loss of expression of the transfected plasmids over time. By subtracting the signal from an amplicon in the dCas9 promoter from the expression of an amplicon in the dCas9 coding region and normalizing to *REEP5* expression, the kinetics of dCas9-ED expression could be tracked. This showed that 18 dpt they exhibited only 0.01% (dCas9-mTet3h) and 0.06% (dCas9-VPR-mTet3h) of the expression measured on the first time point.

Whereas dCas9-mTet3h and dCas9-VPR-mTet3h were initially expressed 161-fold and 39-fold stronger than the gene *REEP5* used for normalization of expression data, this dropped sharply at later time points. Nine days past transfection, dCas9-VPR-mTet3h was expressed only 0.2-times as strongly as *REEP5*, and for dCas9-mTet3h this value had dropped to 0.1-times 12 dpt. Comparing the expression of the dCas9-ED proteins on the last time point 18 dpt to 3 dpt, they exhibited only 0.01% (dCas9-mTet3h) and 0.06% (dCas9-VPR-mTet3h) of the expression observed on the first time point. This enormous drop was expected, but confirmed that very little dCas9-mTet3h was expressed anymore when its activating effect on transcription started. Nevertheless, the activation has to arise from earlier catalytic activity, because the dead mutant of dCas9-mTet3h did not lead to target gene activation.

Summary

Targeting the catalytic domain of TET enzymes to the DNA methylated *EPCAM* promoter in HEK293 led to the stable activation of that gene. However, compared to earlier experiments, both the exact kinetics as well as the achieved effect size were different. The combination of puromycin selection of transfected cells and the use of the hyperactive mTet3h domain allowed us to eliminate the drop of target gene expression after day 9 observed in section 3.6.2 and to boost the maximal *EPCAM* activation from 15.2$\pm$0.3-fold to 59.6$\pm$6.6-fold. I was able to show that the events leading to the activation of *EPCAM* expression had to occur due to the catalytic activity of the effector domain, because a catalytically dead mutant did not lead to target gene expression. Furthermore, by the time the target gene activation was measurable, the expression of dCas9-mTet3h had already declined dramatically, raising the question of the molecular basis for this delayed effect. The fusion protein of mTet3h with VPR was very successful, combining the effects and kinetics of both its constituting domains, and even surpassed their individual effect sizes by far at the *EPCAM* promoter. However, how these synergistic effects come to pass has yet to be investigated.

Interestingly, the four target genes reacted differently to the effector domains: *MAGEB2* behaved related similar to *EPCAM*, with the exceptions that the overall response was slower. Furthermore, activation via VPR was more stable than for *EPCAM*. The other two target genes LIN28A and ASCL1 could not be activated at all with targeted TET enzymes, only with dCas9-VPR. Therefore, the exact epigenetic state, probably especially the amount of DNA methylation for CGI promoters, and the basal rate of transcription determine the possible effect size and the best (combination of) effector domains.

4 Discussion

Despite tremendous progress in uncovering epigenetic phenomena and mapping epigenetic marks to chromatin states and transcriptional profiles, descriptive approaches have failed to demonstrate the causative role of epigenetic marks and to elucidate their individual importance in the epigenetic network. The ability to precisely perturb the epigenetic state of a promoter or enhancer via epigenome editing is a very powerful method to address these issues (Thakore et al., 2016). Analyzing the response of the epigenetic network to a targeted perturbation can allow us to understand the function of epigenetic modifications and how all of them together shape the chromatin state of a given locus. In my doctoral project I developed a toolbox for epigenome editing based on the revolutionary CRISPR-Cas9 system. This toolbox focused on the modulation of DNA methylation at gene promoters using DNA methyltransferases and demethylases, but I also modified histone marks at the same time and locus to enhance the effects and their stability.

4.1 Targeted DNA methylation for gene repression

4.1.1 Efficient and widespread DNA methylation introduced by dCas9-Dnmt3a3L

I first used the catalytically dead dCas9 protein to target the single-chain fusion protein Dnmt3a3L to the unmethylated *EPCAM* promoter in the human ovarian cancer cell line SKOV-3. For this, I transfected the cells transiently with plasmids coding for dCas9-Dnmt3a3L, one or a mix of multiple gRNA plasmids targeting the *EPCAM* promoter, the Venus fluorescent reporter and the pLNGFR plasmid to enrich transfected cells via MACS. This resulted in the widespread deposition of DNA methylation at the targeted promoter 5 dpt, even when using only one gRNA. Interestingly, the best gRNAs were just as effective on their own at targeting DNA methylation as a mix of four of the best gRNAs together, reaching 35.8%$\pm$2.3% methylated CpGs over the whole analyzed region (see figures 3.2 and 3.3A). Furthermore, the binding sites of dCas9 itself were protected from DNA methylation, suggesting very stable binding of the protein to the gRNA target sites. This is supported by studies showing that Cas9 remains tightly bound to even cleaved DNA, requiring high molarity urea to disrupt the complex (Sternberg et al., 2014).

Comparing the gRNAs tiling the promoter region, the gRNAs binding upstream or in the first two thirds of the annotated CpG island were more effective at inducing DNA methylation in the analyzed regions. In contrast to this, the two gRNAs furthest into the

gene body introduced almost no additional DNA methylation. However, that region was also the only part of the analyzed section where untreated SKOV-3 cells already contained any substantial DNA methylation. Additionally, CpG methylation in the first parts of the gene body was also increased for the sample with untargeted dCas9-Dnmt3a3L without any gRNAs, suggesting that this region could be primed for DNA methylation without the need to target it specifically. This might be due to the presence of H3K36me3, which would be expected to be present in the gene bodies of actively transcribed genes.

EPCAM expression was reduced for the samples targeted with dCas9-Dnmt3a3L, as shown in figure 3.3B. However, there was no clear correlation between the average DNA methylation over the region and *EPCAM* repression. One possible explanation would be that the methylation of specific CpG sites, for example in transcription factor binding motives, determines the expression state instead of the average methylation over a larger stretch of DNA. But as the sample with the catalytically inactive Dnmt3a3L mutant C706A also showed significantly reduced *EPCAM* expression, the more likely explanation is that the mere presence of dCas9 bound to critical parts of the promoter leads to gene repression by blocking transcription initiation or elongation. This effect was observed in multiple studies (e.g. Gilbert et al., 2013; Qi et al., 2013).

In contrast to *EPCAM* in SKOV-3 cells, targeting Dnmt3a3L C706A to the promoters of *TFRC* and *CXCR4* in HEK293 cells did not lead to gene silencing. Apart from this difference, the results for *TFRC* and *CXCR4* were very similar those observed at the *EPCAM* promoter and showed that targeting DNA methylation works on multiple target genes and cell lines. But once again, DNA methylation did not automatically cause gene repression, as seen for TFRC gRNAs 4 and 5. This demonstrates the importance of testing multiple gRNAs per target gene and the effects of the methylation of individual CpG sites.

The analysis in figure 3.6 of the amount of introduced DNA methylation relative to the gRNA binding sequence in experiments with only a single gRNA showed that the strongest increase of DNA methylation was observed around 20 to 30 bp downstream of the PAM site with an average increase of ~45 percentage points. Further downstream of the binding site, ΔCpG methylation initially decreased again but showed minor peaks around 200 bp and 350 bp downstream of the PAM site. These values are close to the 146 bp of a nucleosome when including the linker region, which suggests that the positioning of nucleosomes hinders the introduction of DNA methylation as was shown previously (Felle et al., 2011). Upstream of the PAM site, the first 30 bp are protected from DNA methylation by dCas9 occupancy. The peak around 40 bp upstream of the PAM site suggested that the DNMT fused to dCas9 can also reach around the DNA binding protein to methylate DNA there, even though the peptide linking the proteins originates at the downstream site of dCas9. Further upstream, DNA was introduced with similar efficiencies as downstream of the PAM site with broad peaks at around 200 and 360 bp, again fitting to potential nucleosome positions.

The cause for differing amounts of introduced DNA methylation between the gRNAs could be differences in the accessibility of target site. It was shown that dCas9 binds less

strongly to DNA bound to histones (Horlbeck et al., 2016), which would also reduce the amount of introduced DNA methylation. Furthermore, because the DNA methyltransferase is presented mainly in the direction of the PAM site of the gRNA binding site, the dCas9 binding orientation determines the direction of fiber formation. This could explain why EPCAM sgRNAs 11 and 12, and TFRC sgRNAs 1 and 10 failed to introduce strong DNA methylation in the analyzed regions, as they are either oriented towards the end of the analyzed regions or bind further upstream of it.

As shown in figure 3.7, the single chain fusion protein Dnmt3a3L was between 3.8-fold and 4.9-fold more efficient at targeting DNA methylation than Dnmt3a, depending on the target gene. This difference is smaller than observed *in vitro* (~7.5 fold), but larger than the ~2-fold difference in the targeted DNA methylation experiment in cells via zinc-fingers reported in the same study by Siddique et al. (2013). However, even though our system of dCas9-Dnmt3a3L seemed to be highly efficient at targeting widespread DNA methylation to gene promoters, the highest amount of DNA methylation reached was 38.5%$\pm$2.2% with *TFRC* gRNA 3, leading to a 2-fold reduction in gene expression. Most likely, this is because there was never a uniform cell population and only a fraction of all analyzed cells were actually epigenome edited: for HEK293 cells, only around 80% of all cells showed fluorescence of the Venus reporter, and even after enriching for transfected cells via MACS only 88% of SKOV-3 cells were fluorescent. Furthermore, a cell must receive enough of each plasmid necessary for targeted DNA methylation: the dCas9-ED plasmid and the gRNA plasmid(s). Additionally, the plasmids must be neither lost via cell division nor silenced within the first few days, as the transfections were only transient. All these factors reduce the fraction of cells with successful epigenome editing, decreasing the possible effect size. This is especially problematic for targeted gene repression experiments, as the untreated cell population skews the average expression stronger compared to activation experiments. Stringent selection for cells which received all components would alleviate this problem.

When we investigated methylation changes at predicted off-target binding sites of dCas9-Dnmt3a3L, the two analyzed sites for TFRC sgRNA 2 exhibited around 10 percentage points increased DNA methylation (see figure 3.10). In contrast to this, no increase in DNA methylation was detected at four mostly unmethylated CpG islands without predicted off-target binding sites, showing that the increased amount of DNA methyltransferases in the transfected cells did not lead to unspecific DNA methylation. Effects at predicted gRNA off-targets could be mitigated using dCas9 systems with higher specificity, for example by engineering the Cas9 protein (e.g. Slaymaker et al., 2016; Lee et al., 2018), using shorter gRNAs (Fu et al., 2014), and/or titering the plasmid amounts or expression strengths of the system components.

When I had started this project, no publication had been available demonstrating the use of the novel CRISPR-Cas9 system for direct epigenome editing. However, during the course of the experiments, two papers used dCas9 fusion proteins to edit histone modifications (Hilton et al., 2015; Kearns et al., 2015). Then, while writing the manuscript for our findings,

two other groups published targeted DNA methylation studies using the CRISPR-Cas9 system and Dnmt3a. Vojta et al. (2016) targeted two genes in HEK293 cells and observed less spreading of DNA methylation but focused on shorter stretches of DNA with fewer CpGs. In their studies, multiple gRNAs targeting one region led to more DNA methylation compared to a single gRNA. Furthermore, their dCas9-Dnmt3a construct introduced less DNA compared to the dCas9-Dnmt3a3L in my studies, even though they used puromycin to enrich for transfected cells. They also found that the dCas9 binding sites were protected from DNA methylation and that an inactive mutant can lead to significant repression for one of the targeted genes. McDonald et al. (2016) targeted three genes in HEK293T cells, among them the tumor suppressor *CDKN2A*. They achieved no target repression, except when tiling and targeting the whole CpG island with 8 or more gRNAs. Taken together, these two first publications confirm my findings that targeted DNA methylation with the CRISPR-Cas9 system can be effective but highlight potential pitfalls of not achieving sufficient DNA methylation levels across the targeted promoter. Comparing the results of my project to these two studies, my added value was, firstly, the use of the hyperactive single chain fusion protein Dnmt3a3L, which lead to widespread DNA methylation and gene silencing with even a single gRNA. Secondly, the site-specific bisulfite analysis of more than 1 kb per gene to catch this phenomenon, and thirdly, the use of mutant protein to elucidate the mechanism of DNA methylation spreading, which I will discuss in the next section.

Taken together, in this initial part of my thesis I established and characterized a highly efficient tool for targeted DNA methylation using the novel CRISPR-Cas9 system, which I validated using two human cell lines and three target genes.

4.1.2 Elucidation of Dnmt3 fiber formation

The finding that targeting dCas9-Dnmt3a3L with a single gRNA led to the widespread DNA methylation of the whole promoter, covering more than 1 kb, was astounding and I set out to understand the mechanism. Previous reports had indicated that DNMT3A can bind and methylate DNA in a cooperative manner (Rajavelu et al., 2012; Emperle et al., 2014). As shown in scanning force microscopy (SFM) images in figure 4.1A, WT Dnmt3a can completely cover DNA fragments as a fiber, which also increases its catalytic activity (Emperle et al., 2014). In contrast to this, disrupting the lateral multimerization interface with the R832E mutant yields a protein which cannot form fibers and is visible only as single distributed spots on the DNA.

Targeting this mutant instead of the catalytically active Dnmt3a3L indeed led to only small spots of DNA methylation close to the dCas9 binding site and at a few other loci, but not to the widespread DNA methylation observed with the WT protein (see figure 3.9). The results of these experiments suggested a model as in figure 4.1C. Widespread DNA methylation, reaching up to 1 kb from the target site, is only possible due to the ability of Dnmt3a(3L) to multimerize as a fiber along the DNA. This is possible both *in cis* and *in trans* relative to the

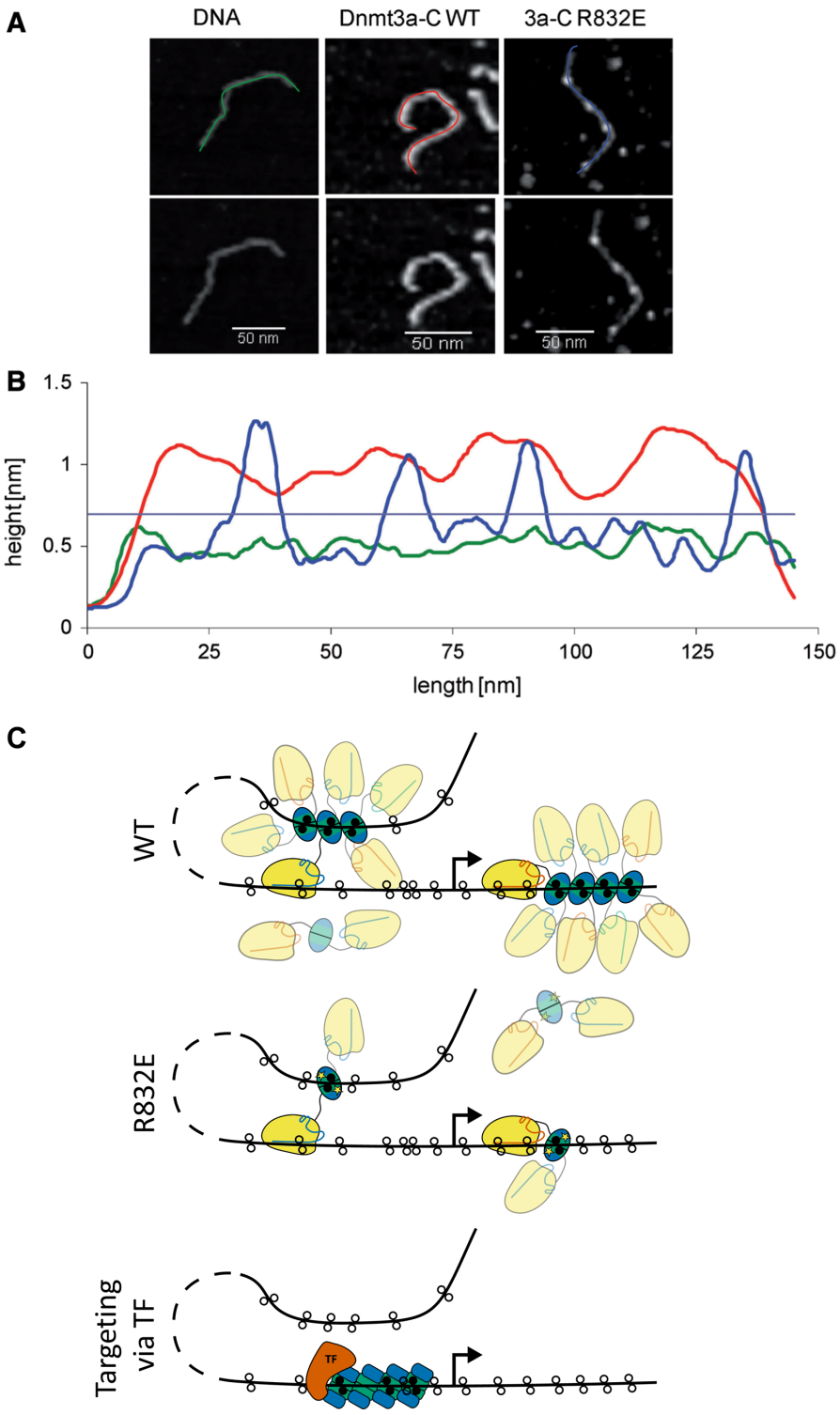

Figure 4.1: Fiber formation of Dnmt3a and its impact on DNA methylation. **(A)** SFM images of DNA alone, DNA with the catalytic domain of WT Dnmt3a and of the R832E mutant with disrupted multimerization interface. **(B)** Height analysis of along the covered lines from (A) reveal binding spots of the wt Dnmt3a fiber (red) and Dnmt3a R832E (blue) on DNA (green). Panels (A) and (B) were taken from Rajavelu et al. (2012). **(C)** Model explaining the spreading of DNA methylation via Dnmt3a3L multimerization (Stepper et al., 2017). When targeting dCas9-Dnmt3a3L, the effector domain can bind either the same or a different DNA strand (*in cis* or *in trans* relative to the targeted DNA strand). This results in DNA methylation (indicated with black circles) either directly next to the gRNA binding site, or on DNA which looped into proximity. For the unmutated protein (top), the Dnmt3a3L dimers can multimerize as a fiber along the DNA. However, the R832E mutant disrupts this binding interface, leading to only localized DNA methylation either next to the binding site or via DNA looping (middle). In a native setting (bottom), endogenous Dnmt3a (green) and Dnmt3L (blue) can be recruited by a transcription factor (brown) to start the formation of a fiber for cooperative introduction of DNA methylation along a stretch of DNA.

DNA strand bound by dCas9. These initial contact sites can then serve as nucleation points for the multimerization of Dnmt3a3L dimers, leading to widespread DNA methylation. If the binding interface is disrupted by the R832E mutant, only these initial contact points or their close vicinity become methylated. This mode of fiber formation could happen both in targeted epigenetic editing events with CRISPR-Cas9 or a similar targeting mechanism, or with a transcription factor recruiting native Dnmt3a(3L) proteins.

4.1.3 Stability and maintenance of targeted DNA methylation

Loss of targeted DNA methylation

The time course experiment to test the stability of the introduced DNA methylation after hit-and-run epigenome editing via dCas9-Dnmt3a3L at the *TFRC* promoter in HEK293 cells revealed that it was lost quickly. After a peak of DNA methylation 5 dpt, its content dropped from 16.7%±1.7% to 6.5%±0.2% within just three days and continued to decline to the initial level. Interestingly, this kinetic was the same when inhibiting TET enzymes with the dioxygenase inhibitor DMOG (see figure 3.11). As this indicated that active DNA demethylation via TET-dependent oxidation was not responsible for the loss of introduced DNA methylation, the other likely explanation was passive dilution and loss during cell division. This meant that the DNA methylation maintenance machinery of DNMT1 targeted by UHRF1 must be either not present at this edited locus or be inhibited.

The loss of targeted epigenetic modifications was observed in several studies (e.g. Kungulovski et al., 2015; Vojta et al., 2016; McDonald et al., 2016), but its cause and dependency on the epigenetic state of the target gene was rarely investigated in detail (Cano-Rodriguez et al., 2016). It seems that changing one modification and one layer of epigenetic control alone is generally not sufficient for stable epigenome editing, as the remaining and probably unchanged epigenetic modifications would inhibit maintenance of the introduced mark and 'mutated' chromatin state.

Thus, to increase target gene repression and edit the epigenetic state of the target genes on multiple layers to facilitate a complete switch, we investigated the fusion of both Dnmt3a3L and the repression domain KRAB to dCas9. As KRAB leads to H3K9 methylation and histone deacetylation (Thakore et al., 2016), we reasoned that changing this layer of epigenetic control together with DNA methylation would function synergistically. We tested the novel fusion proteins in a time course experiment, in which we passaged the samples while harvesting them every two to three days to keep the cells in a continued growth phase. To discern the effect of cell division on the stability of targeted DNA methylation, one sample was not passaged until the final time point 10 dpt, which should cause the cells to stop dividing due to contact inhibition. Yet another sample was additionally cultivated in serum-free conditions to stop cell growth. Indeed, both of these samples retained far more DNA methylation 10 dpt than the continually passaged samples as shown in figure 3.12. Together with the lack of an effect of DMOG addition, these results point to the passive loss of DNA methylation

during cell division due to missing maintenance of the introduced epigenetic mark. Another possible explanation that could contribute to this result is that the growth-inhibited cells retained more plasmids, giving dCas9-Dnmt3a3L-KRAB more time to methylate the *TFRC* promoter.

Additionally, the fusion protein dCas9-Dnmt3a3L-KRAB strongly outperformed dCas9-Dnmt3a3L alone, leading to 11.3% compared to 6.9% DNA methylation 5 dpt. As it is unlikely that KRAB directly increased the catalytic activity of Dnmt3a3L, the most likely explanation is that recruitment of histone deacetylases and H3K9 methyltransferases by KRAB already stimulated maintenance of the introduced DNA methylation during the first cell divisions.

To summarize, I observed quick loss of the introduced DNA methylation after the expression of the effector domains likely ended. The kinetics of DNA methylation loss seemed independent of TET enzyme activity, but the loss was slower when cell division was inhibited, pointing towards passive DNA methylation loss during cell division due to impaired maintenance as the mechanism. Furthermore, I introduced the dCas9-Dnmt3a3L-KRAB fusion protein as a more potent repressor by modulating multiple levels of epigenetic control simultaneously.

Achieving a complete transition of the epigenetic state of targeted genes and investigating the cause of DNA methylation loss

In order to activate the maintenance of the introduced DNA methylation through the endogenous methylation machinery, I attempted to achieve a complete switch of the epigenetic state of the targeted genes. I reasoned that by presenting marks recruiting UHRF1 and DNMT1 and removing marks which inhibit them or recruiting activating factors, I could achieve maintenance of the novel repressed state. To this end, I continued using the fusion protein composed of Dnmt3a3L and KRAB to introduce DNA methylation, H3K9 methylation and histone de-acetylation. Additionally, I investigated the non-targeted, sequence unspecific reduction of H3K4 methylation by inhibiting MLL1 via a small molecule inhibitor and reducing H3K79 methylation via DOT1L inhibition. Furthermore, I again used the dioxygenase inhibitor DMOG to distinguish between active DNA methylation through TET enzymes and passive DNA methylation via cell division and lack of maintenance.

To reduce experimental variability and the influence of untransfected cells on the results, I performed an experiment with puromycin selection of transfected cells. Additionally, to reduce the number of samples necessary to test the conditions listed above, I also investigated the feasibility of co-targeting *CXCR4*, *TFRC* and the new target gene *IL6ST* simultaneously in HEK293 cells. For this, targeted each gene with 2-3 gRNAs in the same samples. Both of these measures were successful: the removal of untransfected cells with puromycin increased the observed target gene repression compared to previous experiments by selecting only cells which received the transfected plasmids for epigenome editing. As all of the genes targeted for repression are overexpressed in various cancers (see section 1.4.1) and *TFRC* was found to be an essential gene in multiple screens (Chen et al., 2017), untransfected and non-edited

cells might otherwise outgrow cells with repressed target genes. Furthermore, targeting dCas9-Dnmt3a3L-KRAB to three different promoters did not diminish its repressive effects.

The long term experiment with multiplexing, puromycin selection, and inhibitors against MLL1, DOT1L and TET enzymes in section 3.4.2 was so far only investigated regarding changes in transcription, while the effects on DNA methylation and histone modifications have yet to be analyzed. However, what I found via qPCR was very promising: again, the fusion protein dCas9-Dnmt3a3L-KRAB outperformed all other effector domains tested in this study and led to persistent silencing of the *TFRC* target gene. In contrast to this, the effector domains KRAB and SID4X only repressed the target genes for the first 3-6 dpt in the absence of any inhibitors. dCas9-Dnmt3a3L consistently induced less target gene repression than Dnmt3a3L-KRAB, but nevertheless caused a reduction in target expression for 12-21 days, depending on the target gene. The different kinetics observed here fit to previously published studies, where KRAB and HDAC4 led to more short lived effects than DNA methyltransferases (Bintu et al., 2016). On the other hand, DNA methylation is often not sufficient for gene silencing, as some genes are simply regulated differently (Ford et al., 2017). That's another reason, in addition to its superior silencing capabilities, why the dCas9-Dnmt3a3L-KRAB fusion protein can work as a universal effector for targeted repression of gene expression. Several other groups confirmed the synergistic effects of targeting KRAB and a DNMT, either via co-targeting (Amabile et al., 2016; O'Geen et al., 2019) or with a fusion protein as I also did in this thesis (Mlambo et al., 2018, using TALEs). Similarly, other fusion proteins also show promising results, such as the fusion of KRAB and MECP2 generated by Yeo et al. (2018), confirming the importance of editing the different layers of the epigenome simultaneously.

Interestingly, inhibiting MLL1 strengthened and prolonged the silencing effects for all tested effector domains and target genes (see figure 3.16). This could mean that reducing the re-establishment of this important activating mark at loci where repressing marks are being artificially introduced, might enable more cells to complete a switch of the transcriptional state. However, a more precise way of achieving the removal of H3K4 methylation compared to the systemic approach with a small molecule inhibitor would be preferable, for example by co-targeting an H3K4 demethylase such as the JARID1 family. H3K4 methylation is an important inhibitor of both DNMT3s and UHRF1, directly linking this modification to the maintenance of the introduced DNA methylation via the endogenous DNA methylation machinery. Similarly, Ford et al. (2017) reported in their preprint that targeted DNA methylation at promoters containing H3K4me3 was lost more quickly. Inhibiting DOT1L, on the other hand, improved the silencing effects only slightly for KRAB and SID4X and did not increase long-term silencing by Dnmt3a3L and Dnmt3a3L-KRAB. Therefore, an involvement of H3K79 methylation in the regulation of DNA methylation seems unlikely, although the efficacies of the small molecule inhibitors in this experiment have not yet been validated. More likely, the small effects seen for KRAB and SID4X stem from the missing activating effects of H3K79 methylation (Cano-Rodriguez et al., 2016).

Similarly, and as seen before, DMOG showed no clear effect on the stability of targeted gene repression by Dnmt3a3L and Dnmt3a3L-KRAB, which again points to the passive loss of DNA methylation via impaired maintenance after cell division instead of active DNA demethylation by TET enzymes. However, the actual changes in DNA methylation have yet to be analyzed at the time of writing. These results will reveal the kinetics of DNA methylation and their dependence on the effector domain (combination), target gene, and inhibitors, as well as how the DNA methylation level correlates with the transcriptional state of the target. Furthermore, analyzing the initial chromatin state and how it changed after targeting the different effector domains via chromatin immunoprecipitation (ChIP) would complete the picture and could identify regulators of DNA methylation maintenance.

Unfortunately, the time point 18 dpt for the *IL6ST* target gene did not show any changes in gene expression, even though the time points before and after showed the expected results. Potentially, technical problems with the RNA extraction or some perturbation to the control samples caused this outlier time point, but repeating the qPCRs did not change the results. Similarly, the data quality achieved for the *CXCR4* expression analysis was worse than expected and observed in the previous multiplexing test experiment, leading to suboptimal variability in the qPCR results. Interestingly, the long-term kinetics observed in the multiplexing test experiment (section 3.4.1) and the inhibitor experiment (section 3.4.2) for *TFRC* and *CXCR4* were almost contrary: while *CXCR4* stayed repressed for all of the analyzed 9 dpt in the multiplexing experiment, expression of *TFRC* rebounded towards the initial level between days 3 and 9 past transfection. But in the inhibitor experiment, Dnmt3a3L-KRAB retained its silencing effect on *TFRC* expression for the analyzed period, while *CXCR4* increased again in expression between days 3 and 9. The cause for this discrepancy is unknown, but might be caused by problems with RNA isolation.

Taken together, the results from this experiment provide a provisional explanation for the generally observed loss of introduced DNA methylation. I again showed that removal of targeted DNA methylation is not dependent on TETs, but is rather caused by a lack of DNA methylation maintenance. The repressed state can be prolonged by editing the epigenome on multiple layers by additionally recruiting H3K9 methylation and histone deacetylation with KRAB. Additionally, reducing the re-introduction of H3K4 methylation by inhibiting MLL1 further increases the propagation of the repressed states over multiple cell divisions, probably by allowing the maintenance machinery of UHRF1 and DNMTs to be active at the edited locus.

A possible way to further investigate the importance of DNA methylation maintenance could be the analysis of changes in UHRF1 and/or DNMT1 occupancy via ChIP after targeted DNA methylation. Additionally, the hairpin bisulfite method developed by Laird et al. (2004) to compare the methylation states of both DNA strands after cell divisions could reveal that the parental strand is indeed more strongly methylated than the daughter strand. This would indicate that indeed the lack of maintenance causes demethylation over time, instead of active demethylation, which would lead to similarly reduced DNA methylation levels for

both strands. Selected samples from this experiment were prepared for sequencing using this method.

Furthermore, I performed targeted DNA methylation in TET-triple-knock-out murine embryonic stem cells. By comparing the DNA methylation kinetics obtained in this cell line with the results in normal murine E14 cells, it will be possible to discern the influence of TET enzymes on the rate of DNA methylation loss. I expect to see no difference in DNA methylation kinetics, as my results obtained so far point to the paramount importance of DNA methylation maintenance, while no TET dependency was observed. However, the results from this experiment are not yet available.

4.1.4 Applications of targeted DNA methylation

Studying tumor initiation in primary breast cells using targeted DNA methylation

In cooperation with the group of Dr. Gabriella Ficz, we studied the influence of promoter methylation of tumor suppressor genes on the transformation of primary myoepithelial breast cells to apply the epigenetic editing toolbox to other biological questions than basic epigenetic research. We found that targeting dCas9-Dnmt3a3L by transient transfection to the promoters of *HIC1*, *PTEN*, *RASSF1* and *CDKN2A* led to strong DNA methylation of the promoters 10 dpt and repressed the expression of the *CDKN2A* and *RASSF1* transcripts. Importantly, the edited cells did not become senescent but continued proliferating, which is a first step towards tumor formation. Interestingly, targeting the p16 promoter of *CDKN2A* alone was already sufficient to induce the senescence escape phenotype, but it was weaker than when additionally targeting p14 or all four target genes at once.

In contrast to the experiments discussed in the previous sections, just targeting dCas9-Dnmt3a3L without KRAB or inhibitors led to persistent DNA methylation and target gene repression. However, this was the case for only two of the three investigated patient derived cells, which highlights again the importance of the precise local chromatin state and points to a stochastic nature of transcriptional and epigenetic switching events. Furthermore, it is conceivable that successfully edited cells with silenced tumor suppressor genes and a proliferative phenotype outperformed both weakly and non-edited cells, which would soon enter senescence. Thus, edited cells would be selected, which explains the difference to my previous results, where I instead targeted cells which are generally upregulated in cancer and not silenced as in this experiment.

This study demonstrated the potential of epigenome editing to dissect changes in chromatin and transcriptional states in the development of cancer and other diseases and is part of a growing field of studies using this innovative approach (Gun et al., 2013; Cui et al., 2015; Boulay et al., 2018).

Inducing insulin production in murine pancreatic α-cells via targeted DNA methylation

The reconstitution of insulin production in diabetes patients is an important and active field of research (e.g. Shapiro et al., 2006; Pagliuca et al., 2014; Li et al., 2017). In cooperation with the group of Dr. Melita Vidaković, we achieved the induction of insulin production by targeting DNA methylation with dCas9-Dnmt3a3L and dCas9-Dnmt3a3L-KRAB to the *Arx* promoter in murine pancreatic α-cells. By repressing the master regulator *Arx*, which is also silenced in the insulin producing β-cells, the cells changed their transcriptional programs and started producing insulin. However, the changes were relatively small on a population level with a maximum upregulation of 7.6±2.5-fold activation in our hands and 14.9±0.8-fold by our collaborators. A potential explanation for the observed differences between the results obtained by our and our collaborator's group are different cell lines: we both used α-cell cancer cell lines (αTC1.6 and αTC1.9), but they show small differences in glucagon secretion and closeness to healthy α-cells.

In contrast to the activation achieved here, insulin constitutes around 40-50% of all mRNA in β-cells (Nica et al., 2013). Additionally, insulin production was only transient, falling back to initial levels after less than 20 days. Potentially, only a small percentage of all sorted cells were actually successfully edited and started producing insulin, because only a fraction of cells stained strongly for insulin. This would be encouraging, because it would show that a sufficiently complete switch to insulin production had actually taken place, just not in enough cells. However, more stringent selection of transfected cells or single cell analysis would be necessary to confirm this hypothesis.

Both the fraction of edited cells and the strength of insulin induction could be increased by, firstly, more stringent selection such as MACS or a combination with FACS, secondly, the activation of *Pax4* expression in addition to *Arx* repression via orthogonal multiplexing and thirdly, the targeted removal of H3K4 methylation at the *Arx* promoter.

Recently, other groups also attempted to induce insulin production in pancreatic alpha-cells. For example, one study claimed to achieve regeneration of β-cell mass from α-cells via artemisinins (Li et al., 2017, but could not be reproduced by Meulen et al., 2018). Others repressed *Arx* expression via shRNA and observed a 12.8-fold upregulation of insulin expression (Friedman-Mazursky, Elkon, and Efrat, 2016), similar to what we achieved.

4.2 Targeted DNA demethylation for gene activation

To complete the epigenome editing toolbox with effector domains for targeted DNA demethylation and transcriptional activation, I developed dCas9 fusion proteins with engineered, hyperactive TET enzymes. This culminated in the dCas9-VPR-mTet3h effector, which achieved synergistic and stable target gene activation. In contrast to this, most studies employing

targeted DNA demethylation until then had found only small changes in DNA methylation and expression levels (e.g. Chen et al., 2014; Xu et al., 2016).

4.2.1 Delayed but stable gene activation

My initial experiments testing the effects on *EPCAM* expression in HEK293 cells for a variety of effector domains 5 dpt showed only minor transcriptional activation for dCas9-mTet1 and no activation with dCas9-mTet2. In contrast to this, both of these effectors induced DNA demethylation at the promoter as shown in figure 3.25. Interestingly, both dCas9-VP64 and dCas9-VPR achieved DNA demethylation at similar positions, albeit at lower levels and only if ascorbate was added to the cell media. This suggests that also the DNA demethylation after targeting VP64 and VPR was dependent on endogenous TET activity, as ascorbate increases TET activity by reducing Fe(III) to Fe(II) (Hore et al., 2016). Generally, stronger DNA demethylation did not automatically lead to higher *EPCAM* expression (compare mTet1 and mTet2; and VP64/VPR with and without ascorbate), suggesting a more complex control of transcription activation, potentially depending on the methylation of only a few important CpG sites. Furthermore, most gRNA binding sites were protected from DNA demethylation, just like for dCas9-Dnmt3a3L with regard to the introduction of DNA demethylation.

The slight activation of *EPCAM* expression by dCas9-AID was probably caused by increased accessibility of the local chromatin caused by DNA repair machinery induced by the mutations introduced by AID, a feature which is harnessed in the CRISPR-X method (Hess et al., 2016). However, this did not lead to a decrease in DNA methylation content, even though AID was implicated in DNA demethylation (Santos et al., 2013).

Surprisingly, the subsequent time course experiment revealed that the activation seen with mTet1 5 dpt was just the first sign of a delayed target gene activation, as *EPCAM* expression increased further from 5 to 8 dpt (see figure 3.26). The fact that the achieved activation of *EPCAM* expression was lower than in the previous experiment, as seen also for VP64, was probably caused by a combination of inferior transfection efficiency and lack of ascorbate addition in this experiment (for TET enzymes). Interestingly, the dCas9-mTet1 version with a longer linker (48 amino acids compared to 28) outperformed the previous version, suggesting that the catalytic domains of TET enzymes need more space and flexibility to contact the DNA compared to e.g. Dnmt3a3L. Furthermore, the engineered mTet3 version with a removed unstructured and not conserved region (mTet3-del1; Ravichandran, 2017) resulted in an even stronger but also delayed upregulation on day 8. Surprisingly, the activation of *EPCAM* expression was higher than achieved with VP64.

The lack of any effect of the TET protein from *Naegleria gruberi* (nTet) was puzzling, as the enzyme exhibited higher conversion activity to 5fC and 5caC than mTet3-de1 *in vitro* (Ravichandran, 2017). One explanation would be that nTet is less active in a mammalian nucleus, although the mechanism for this would be unclear. It is more likely that the

construct contained an undiscovered mistake or the fusion to dCas9 somehow abolished its catalytic activity.

When we continued culturing the transfected cells past the 8 days assayed so far for the next experiment, we saw that TET induced *EPCAM* expression peaked 9 dpt, but remained strongly elevated for at least 46 dpt (see figure 3.27). Furthermore, peak activation with dCas9-mTet3-del1 was higher than the strongest activation achieved with VPR, which had peaked 3 dpt. This striking result posed a lot of questions: why was the activation kinetic delayed and why was the peak achieved after the plasmids should already have been mostly lost from the cells? Why did *EPCAM* expression decrease after the peak 9 dpt but then level off? As the activation achieved with VPR probably mirrors the expression of the dCas9-ED construct, which would also fit to the expression of the fluorescence reporter, activation via targeted TET enzymes only started to have strong effects after the effector was already gone. I will try to answer these questions in the discussion of the subsequent experiments. Nevertheless, this experiment demonstrated the importance to assay target gene expression after epigenome editing for multiple time points, and highlighted the beneficial effect of ascorbate on targeted activation studies, both with TET enzymes and other activators as VPR.

4.2.2 Engineered TET enzymes including fusions with activating domains

As fusing KRAB to dCas9-Dnmt3a3L improved the extent and stability of targeted gene repression dramatically, I tried to achieve a similar response for activation by combining the successful activators mTet3-del1 and VPR. Furthermore, as targeting VPR led to early but transient activation while mTet3-del1 achieved delayed but stable target gene expression, we reasoned that both effectors together should enable robust transcriptional activation starting shortly after transfection. The initial tests in figure 3.28 showed that the order in which the effectors were fused to dCas9 had strong effects on their ability to induce transcriptional activation. Furthermore, the results already hinted at the differences in target gene response observed in the final time course experiment, as dCas9-VPR led to sustained activation of *LIN28A* at the analyzed time point 9 dpt, in contrast to *EPCAM*. Additionally, the further engineered mTet3 catalytic domain del1-ΔC (mTet3h) with a shortened C-terminus surpassed the activation achieved with dCas9-mTet3-del1, yielding the most active TET construct that we have tested.

Interestingly, the fusion dCas9-VPR-mTet3h led to only around half the increase in *LIN28A* expression achieved with dCas9-VPR alone, similar to the sample where dCas9-VPR and dCas9-mTet3-del1 were co-targeted. This suggests that, on the one hand, VPR functioned less efficiently when TET is fused to it, potentially due to spatial restrains. On the other hand, activation with VPR seemed to be dose dependent even at the relatively high expression levels in the transient transfection system: the co-targeting sample was transfected with

half the amount of dCas9-VPR compared to the VPR only sample, leading to half the effect on gene expression. Surprisingly, co-targeting the two effector domains to *EPCAM* did not achieve any synergistic effects as with their fusion protein. Instead, *EPCAM* expression was slightly lower than for dCas9-mTet3-del1 alone, similar to dCas9-VPR for *LIN28A*. However, there seems to be no reason why VPR and mTet3-del1 need to be fused together to increase *EPCAM* expression, as targeting them to the same gRNA binding sites at the same promoter should suffice.

The result that fusing an additional activation domain to a TET enzyme for epigenome editing boosts its activation of target gene expression was recently confirmed by two other studies: Firstly, Baumann et al. (2019) combined VP64 and Tet1 and found that Tet1 increased the conversion rate to an active transcription state compared to VP64 alone. Secondly, Taghbalout et al. (2019) co-delivered either the DNA repair associated protein GADD45A or the DNA glycosylase/AP-lyase NEIL2 with TET1 to streamline the removal of oxidized 5mC and found that this strongly increased both DNA demethylation and target gene activation. This constitutes a very interesting approach to study and alleviate the potential bottleneck of removing the oxidized cytosines.

4.2.3 Differences in target gene response

After validating the new tools and targets, we combined the new dCas9 fusions with mTet3h and VPR-mTet3h with the new target genes *ASCL1*, *LIN28A* and *MAGEB2* in a detailed time course experiment (see figure 3.29). We found that the fusion of VPR and mTet3h yielded strong synergistic effects not only by boosting the activation via TET at later time points as observed in the previous experiment, but also at the first time point 3 dpt for *EPCAM*, at which the dCas9-mTet3h fusion protein alone did not yet increase *EPCAM* expression at all. This probably means that mTet3h starts oxidizing 5mC from the beginning, similarly to the fact that adding ascorbate only until the first time point did not change the achieved activation with dCas9-mTet3-del at day 9 as observed in the previous experiment. The 5mC oxidation and demethylation increases the activation achieved with VPR, but does not show an effect initially on its own. Potentially, other activating factors like histone modifications and transcription factors are recruited only slowly to the freshly demethylated DNA, leading to the delayed target gene activation. Furthermore, the opening of chromatin by VPR might aid the demethylation via TET. Another explanation would be that the TET enzyme introduces only very localized DNA demethylation which then spread over the rest of the locus over multiple cell divisions due to impaired DNA methylation maintenance. Furthermore, the DNA methylation maintenance machinery might be blocked more thoroughly by activating modifications recruited by VPR.

Whereas *EPCAM* and *MAGEB2* both responded strongly to activation by dCas9-mTet3h and dCas9-VPR-mTet3h and even continued to increase in expression over time, the other two target genes *ASCL1* and *LIN28A* showed almost the opposite response. For both genes,

mTet3h failed to increase gene expression and the VPR-mTet3h fusion was less effective than VPR alone. Apparently, these genes are not primarily regulated via DNA methylation. Interestingly, the activation with dCas9-VPR was stronger and more stable at these genes compared to *EPCAM*. However, the data quality problems with *ASCL1* due to its very low basal expression, which made detection via qPCR challenging, prohibit definitive answers in this regard. It might be necessary to repeat the experiment with higher cell numbers to increase the RNA yield to improve the quality of the results for *ASCL1*. Similarly, running another time course for *EPCAM* might show that the decrease in expression for dCas9-VPR-mTet3h on day 6 was an outlier and expression would actually increase more or less linearly, which would be more logical based on the theories outlined above. Importantly, the continued increase in *EPCAM* and *MAGEB2* expression, long after the constructs declined strongly in expression as shown by qPCR, might be caused by growth promoting effects of both of these genes (e.g. Yamashita et al., 2007; Nobeyama and Nakagawa, 2016). Even though the puromycin expression should ensure that only transfected cells survive, not all of them will achieve a successful switch of the chromatin and transcriptional state, potentially leading to a slow change of the population composition.

To explain the delayed target gene activation with TET enzymes and the differences in the response of the target genes to the various effector domains, the results of bisulfite sequencing and ChIP of the most important time points would be necessary. However, these results have yet to be finalized and analyzed. Studying the site-specific content of 5mC and 5hmC over time will show whether demethylation starts localized or distributed over the promoter, and whether 5mC levels correlate inversely with transcription or already drop dramatically on the first time points. Similarly, knowledge of the basal chromatin states, and the kinetics of loss of repressive histone marks and gain of activating modifications at the four target genes should allow us to explain their differing responses and kinetics of transcriptional activation.

Finally, I want to revisit the possible avenues to boost the demethylation activity in epigenome editing I had outlined at the beginning of section 3.6: the addition of ascorbate to the cell culture medium indeed increased target gene activation for both TET enzymes and VPR, as seen in figure 3.27. In contrast to this, the TET variant from *Naegleria gruberi* failed to achieve any effects on DNA methylation and transcription, despite promising *in vitro* results. The effects of using just one versus a mix of many gRNAs per promoter were not investigated, but based on the available bisulfite sequencing results in figure 3.25 it seems that demethylation was mostly induced close to the gRNA binding sites. This would mean that tiling the promoter with multiple gRNAs as I did in this thesis is more effective than using just a single gRNA. However, Xu et al. (2016) found that a single gRNA can be more effective, if targeted to a sensitive region. Lastly, both the engineered mTet3 variants and the fusion protein with VPR were complete successes, improving vastly on the initial results achieved with the dCas9-sl-mTet1 construct.

These improvements allowed me to surpass most of the targeted DNA demethylation

results achieved so far by other groups. For example, Chen et al. (2014) found only small differences in the methylation of single CpG sites and subsequently barely any effect on target gene transcription. In both regards, already my first experiment achieved stronger effects on both methylation and expression. The first studies using the CRISPR-Cas9 system for targeted DNA methylation were already more successful and achieved strong demethylation, especially when using systems recruiting multiple TET proteins like the MS2 coat protein (Xu et al., 2016) or SunTag (Morita et al., 2016). However, the detailed time course experiments for multiple effectors and target genes remain unprecedented.

4.3 Bio-On-Magnetic-Beads (BOMB)

Faced with the need to dramatically scale up the number of samples for my experiments but a lack of funds and researchers we turned to magnetic beads as a scalable and flexible way to isolate and handle nucleic acids. By synthesizing and coating the beads ourselves and developing and preparing our own methods and buffers, we were able to cut down the costs per sample by at least an order of magnitude compared to commercial kits. We did not only save money, but also a lot of time as processing 96 or even more samples simultaneously using single-tube spin columns is almost impossible.

However, this task is very easy with magnetic beads thanks to the use of 96-well plates and multichannel pipettes. These measures enabled me to perform and analyze the final time course experiments for both targeted gene activation and repression, which otherwise would have been prohibitively money and time intensive.

The platform bomb.bio, which we established after developing and optimizing the protocols out of necessity for our own use, aims to bring these benefits to all other scientists and will enable labs to perform many more and better experiments than before. As cost and speed are the two main advantages, not only budget-restricted labs will benefit, but also richer research groups which can enjoy the time-savings and flexibility of the platform. As shown in figure 4.2, the BOMB platform allows a wide variety of input samples and allows the manipulation of nucleic acids on the beads between washing steps during the isolation or clean-up, allowing many applications and downstream processes.

The feedback that we received since publishing the platform has been very encouraging as scientists from many institutions with varying levels of funding were inspired by our ideas, started using our protocols and even began supplying their own solutions on our homepage. I believe that the advantages of this platform can make it a powerful tool to enable a new kind of science for many researchers, benefiting all of us.

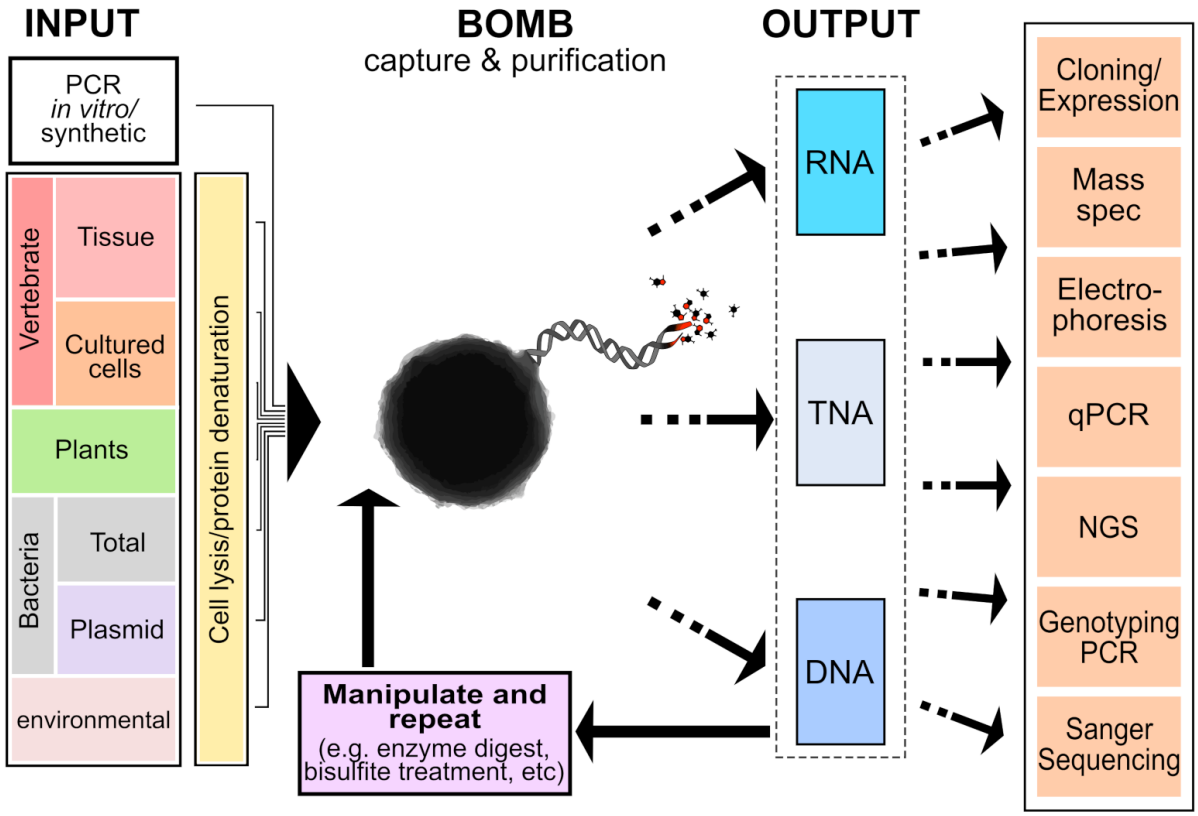

Figure 4.2: Overview of the input sources for BOMB and possible downstream processes (Oberacker et al., 2019). Nucleic acids can be purified from a variety of biological sources. The isolation, manipulation and purification steps are modular and based on homemade magnetic beads. Resulting nucleic acid isolates are of high quality and suitable for sensitive downstream processes.

4.4 Outlook

While this work already contains a lot of data and the time course expression data generated important insights, unfortunately, not all downstream analysis could be performed in time for this thesis. As already written above, puzzling questions remain regarding the stability of targeted DNA methylation, the influence of the local chromatin state on maintenance and the differing response of the target genes, and finally, the mechanism of delayed target gene activation with targeted TET enzymes. To answer them, it will be necessary to analyze the changes in DNA (hydroxy-)methylation and histone modifications. Furthermore, the hairpin bisulfite results for the targeted repression with inhibitors time course, which are in sequencing, will be an important step to find a definitive answer to the question whether active or passive demethylation is the cause for the loss of introduced DNA methylation. Similarly, the bisulfite sequencing for the comparison of demethylation kinetics after targeting DNA methylation between TET triple-KO and WT murine embryonic stem cells will likely prove the lack of importance of TET enzymes. To elucidate the mechanisms of regulation of DNA methylation maintenance, ChIP of active and repressive histone marks will be invaluable to discover the critical factors influencing the binding and activity of UHRF1 and DNMT1.

To further increase the efficacy of the dCas9-Dnmt3a3L-KRAB fusion protein, it should be beneficial to combine it with an H3K4 demethylase. On the other hand, combining dCas9-VPR-mTet3h with factors recruiting MLL complexes, and proteins of the DNA repair machinery as done by Baumann et al. (2019) would likely boost its efficacy even higher. Generally, the avenue of combining multiple factors to edit the epigenome on multiple layers as done in this thesis seems like a promising way to achieve a complete conversion of the target chromatin and transcriptional state. Importantly, the composition of the effector domains should be tailored to the target gene(s) for best effects.

Advances in sequencing technology, especially in regard to droplet based methods for singe-cell analysis (e.g. Datlinger et al., 2017; Lareau et al., 2019) will allow the field to capture the switching behavior of target genes and explain differences in responses between single cells due to stochastic variances in chromatin states and transcription. I expect that in future experiments, researches will edit many loci simultaneously, allowing them to drastically change the transcriptional profile and phenotype of the cells to study developmental and transformational events. The next major milestone for the field of epigenome editing will hopefully be the elucidation of the epigenetic code and network by analyzing the responses to perturbation in a high-throughput manner. Furthermore, research will likely continue to move towards *in vivo* models, both to better capture the environments in organisms and to get closer to a potential therapeutic application of epigenome editing. This would be in contrast to current epigenetic drugs, which affect the whole cell and organism instead of the actually necessary localized changes (Heerboth et al., 2014).

Regarding the high-throughput methods and online community bomb.bio that we developed, I hope that it will continue to thrive and inspire other scientists to start both using the published protocols and adapt and expand them for their own unique applications. The platform can develop into a hub of researchers who want to reduce their dependency of expensive commercial kits and instead spend their money on the people designing and performer better experiments. Furthermore, the flexibility of the magnetic bead system can allow us to expand into other areas using additional bead coatings, such as for protein isolation, handling and selection.

Bibliography

Abbate, J., J.C. Lacayo, M. Prichard, G. Pari, and M.A. McVoy (2001). "Bifunctional Protein Conferring Enhanced Green Fluorescence and Puromycin Resistance". In: *Biotechniques* 31.2, pp. 336–340.

Adams, Grace E, Aditya Chandru, and Shaun M Cowley (2018). "Co-repressor, co-activator and general transcription factor: the many faces of the Sin3 histone deacetylase (HDAC) complex." In: *Biochem. J.* 475.24, pp. 3921–3932.

Aggarwal, Anuj, Rahul Saxena, Bo Wang, and Gerard T. Caneba (1996). "Studies of the polymerization of methacrylic acid via free-radical retrograde precipitation polymerization process". In: *J. Appl. Polym. Sci.* 62.12, pp. 2039–2051.

Allen, Matthew (2015). "Compelled by the Diagram: Thinking through C. H. Waddington's Epigenetic Landscape". In: *Contemp. Hist. Presence Vis. Cult.* 4, pp. 119–142.

Amabile, Angelo, Alessandro Migliara, Paola Capasso, Mauro Biffi, Davide Cittaro, et al. (2016). "Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing". In: *Cell* 167.1, 219–232.e14.

Arand, Julia, David Spieler, Tommy Karius, Miguel R. Branco, Daniela Meilinger, et al. (2012). "In Vivo Control of CpG and Non-CpG DNA Methylation by DNA Methyltransferases". In: *PLoS Genet.* 8.6. Ed. by Dirk Schübeler, e1002750.

Atlasi, Yaser and Hendrik G. Stunnenberg (2017). "The interplay of epigenetic marks during stem cell differentiation and development". In: *Nat. Rev. Genet.* 18.11, pp. 643–658.

Baker, Darren J., Bennett G. Childs, Matej Durik, Melinde E. Wijers, Cynthia J. Sieben, et al. (2016). "Naturally occurring p16Ink4a-positive cells shorten healthy lifespan". In: *Nature* 530.7589, pp. 184–189.

Barrangou, Rodolphe and Luciano A. Marraffini (2014). "CRISPR-Cas Systems: Prokaryotes Upgrade to Adaptive Immunity". In: *Mol. Cell* 54.2, pp. 234–244.

Barski, Artem, Suresh Cuddapah, Kairong Cui, Tae-Young Roh, Dustin E. Schones, et al. (2007). "High-Resolution Profiling of Histone Methylations in the Human Genome". In: *Cell* 129.4, pp. 823–837.

Bashtrykov, Pavel, Gytis Jankevicius, Renata Z Jurkowska, Sergey Ragozin, and Albert Jeltsch (2014). "The UHRF1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase DNMT1 by an allosteric mechanism." In: *J. Biol. Chem.* 289.7, pp. 4106–15.

Baumann, Valentin, Maximilian Wiesbeck, Christopher T Breunig, Julia M Braun, Anna Köferle, et al. (2019). "Targeted removal of epigenetic barriers during transcriptional reprogramming." In: *Nat. Commun.* 10.1, p. 2119.

Ben-Porath, I and H Cedar (2001). "Epigenetic crosstalk." In: *Mol. Cell* 8.5, pp. 933–5.

Berdasco, María and Manel Esteller (2011). "DNA methylation in stem cell renewal and multipotency." In: *Stem Cell Res. Ther.* 2.5, p. 42.

Berger, S. L., T. Kouzarides, R. Shiekhattar, and A. Shilatifard (2009). "An operational definition of epigenetics". In: *Genes Dev.* 23.7, pp. 781–783.

Bernstein, Bradley E, John A Stamatoyannopoulos, Joseph F Costello, Bing Ren, Aleksandar Milosavljevic, et al. (2010). "The NIH Roadmap Epigenomics Mapping Consortium." In: *Nat. Biotechnol.* 28.10, pp. 1045–8.

Bhan, Arunoday, Paromita Deb, and Subhrangsu S. Mandal (2017). "Epigenetic Code". In: *Gene Regul. Epigenetics Horm. Signal.* Weinheim, Germany: Wiley-VCH, pp. 29–58.

Bheda, Poonam and Robert Schneider (2014). "Epigenetics reloaded: the single-cell revolution". In: *Trends Cell Biol.* 24.11, pp. 712–723.

Bintu, L., J. Yong, Y. E. Antebi, K. McCue, Y. Kazuki, et al. (2016). "Dynamics of epigenetic regulation at the single-cell level". In: *Science (80-.).* 351.6274, pp. 720–724.

Bird, A. (2002). "DNA methylation patterns and epigenetic memory". In: *Genes Dev.* 16.1, pp. 6–21.

Bird, Adrian (2007). "Perceptions of epigenetics". In: *Nature* 447.7143, pp. 396–398.

Bird, Adrian P. (1986). "CpG-rich islands and the function of DNA methylation". In: *Nature* 321.6067, pp. 209–213.

Boch, J., H. Scholze, S. Schornack, A. Landgraf, S. Hahn, et al. (2009). "Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors". In: *Science (80-.).* 326.5959, pp. 1509–1512.

Booth, Michael J, Tobias W B Ost, Dario Beraldi, Neil M Bell, Miguel R Branco, et al. (2013). "Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine". In: *Nat. Protoc.* 8.10, pp. 1841–1851.

Bostick, Magnolia, Jong Kyong Kim, Pierre-Olivier Estève, Amander Clark, Sriharsa Pradhan, and Steven E Jacobsen (2007). "UHRF1 plays a role in maintaining DNA methylation in mammalian cells." In: *Science* 317.5845, pp. 1760–4.

Boulay, Gaylor, Angela Volorio, Sowmya Iyer, Liliane C Broye, Ivan Stamenkovic, et al. (2018). "Epigenome editing of microsatellite repeats defines tumor-specific enhancer functions and dependencies." In: *Genes Dev.* 32.15-16, pp. 1008–1019.

Brien, Gerard L., Kimberly Stegmaier, and Scott A. Armstrong (2019). "Targeting chromatin complexes in fusion protein-driven malignancies". In: *Nat. Rev. Cancer* 19.5, pp. 255–269.

Bujold, David, David Anderson de Lima Morais, Carol Gauthier, Catherine Côté, Maxime Caron, et al. (2016). "The International Human Epigenome Consortium Data Portal." In: *Cell Syst.* 3.5, 496–499.e2.

Cabrero-de Las Heras, Sara and Eva Martínez-Balibrea (2018). "CXC family of chemokines as prognostic or predictive biomarkers and possible drug targets in colorectal cancer." In: *World J. Gastroenterol.* 24.42, pp. 4738–4749.

Campisi, Judith (2013). "Aging, Cellular Senescence, and Cancer". In: *Annu. Rev. Physiol.* 75.1, pp. 685–705.

Cano-Rodriguez, David, Rutger A F. Gjaltema, Laura J Jilderda, Pytrick Jellema, Jelleke Dokter-Fokkens, et al. (2016). "Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner". In: *Nat. Commun.* 7.1, p. 12284.

Caporaso, J Gregory, Justin Kuczynski, Jesse Stombaugh, Kyle Bittinger, Frederic D Bushman, et al. (2010). "QIIME allows analysis of high-throughput community sequencing data". In: *Nat. Methods* 7.5, pp. 335–336.

Chatterjee, Aniruddha, Euan J. Rodger, and Michael R. Eccles (2018). "Epigenetic drivers of tumourigenesis and cancer metastasis". In: *Semin. Cancer Biol.* 51, pp. 149–159.

Chavez, Alejandro, Jonathan Scheiman, Suhani Vora, Benjamin W Pruitt, Marcelle Tuttle, et al. (2015). "Highly efficient Cas9-mediated transcriptional programming". In: *Nat. Methods* 12.4, pp. 326–328.

Chavez, Alejandro, Marcelle Tuttle, Benjamin W Pruitt, Ben Ewen-Campen, Raj Chari, et al. (2016). "Comparison of Cas9 activators in multiple species". In: *Nat. Methods* 13.7, pp. 563–567.

Chen, Chun-Wei, Richard P Koche, Amit U Sinha, Aniruddha J Deshpande, Nan Zhu, et al. (2015). "DOT1L inhibits SIRT1-mediated epigenetic silencing to maintain leukemic gene expression in MLL-rearranged leukemia". In: *Nat. Med.* 21.4, pp. 335–343.

Chen, Hui, Hinke G Kazemier, Marloes L. de Groote, Marcel H. J. Ruiters, Guo-Liang Xu, and Marianne G. Rots (2014). "Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter". In: *Nucleic Acids Res.* 42.3, pp. 1563–1574.

Chen, Wei-Hua, Guanting Lu, Xiao Chen, Xing-Ming Zhao, and Peer Bork (2017). "OGEE v2: an update of the online gene essentiality database with special focus on differentially essential genes in human cancer cell lines". In: *Nucleic Acids Res.* 45.D1, pp. D940–D944.

Chen, Zhao-Xia, Jeffrey R. Mann, Chih-Lin Hsieh, Arthur D. Riggs, and Frédéric Chédin (2005). "Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family". In: *J. Cell. Biochem.* 95.5, pp. 902–917.

Cheng, Albert W, Nathaniel Jillette, Phoebe Lee, Dylan Plaskon, Yasuhiro Fujiwara, et al. (2016). "Casilio: a versatile CRISPR-Cas9-Pumilio hybrid for gene regulation and genomic labeling". In: *Cell Res.* 26.2, pp. 254–257.

Chomczynski, Piotr and Nicoletta Sacchi (1987). "Single-step method of RNA isolation by acid guanidinium extraction by acid guanidinium thiocyanate-phenol-chloroform extraction". In: *Anal. Biochem.* 162.1, pp. 156–159.

Chomczynski, Piotr and Nicoletta Sacchi (2006). "The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: Twenty-something years on". In: *Nat. Protoc.* 1.2, pp. 581–585.

Chory, Emma J., Joseph P. Calarco, Nathaniel A. Hathaway, Oliver Bell, Dana S. Neel, and Gerald R. Crabtree (2019). "Nucleosome Turnover Regulates Histone Methylation Patterns over the Genome". In: *Mol. Cell* 73.1, 61–72.e3.

Chouchane, Malek and Marcos R. Costa (2018). "Instructing neuronal identity during CNS development and astroglial-lineage reprogramming: Roles of NEUROG2 and ASCL1". In: *Brain Res.*

Choudhury, Samrat Roy, Yi Cui, Katarzyna Lubecka, Barbara Stefanska, and Joseph Irudayaraj (2016). "CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter." In: *Oncotarget* 7.29, pp. 46545–46556.

Ciccarone, Fabio, Elisabetta Valentini, Michele Zampieri, and Paola Caiafa (2015). "5mC-hydroxylase activity is influenced by the PARylation of TET1 enzyme." In: *Oncotarget* 6.27, pp. 24333–47.

Citterio, E., R. Papait, F. Nicassio, M. Vecchi, P. Gomiero, et al. (2004). "Np95 Is a Histone-Binding Protein Endowed with Ubiquitin Ligase Activity". In: *Mol. Cell. Biol.* 24.6, pp. 2526–2535.

Collombat, Patrick, Xiaobo Xu, Philippe Ravassard, Beatriz Sosa-Pineda, Sébastien Dussaud, et al. (2009). "The ectopic expression of Pax4 in the mouse pancreas converts progenitor cells into alpha and subsequently beta cells." In: *Cell* 138.3, pp. 449–62.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, et al. (2013). "Multiplex Genome Engineering Using CRISPR/Cas Systems". In: *Science (80-.).* 339.6121, pp. 819–823.

Costa, Yael, Junjun Ding, Thorold W. Theunissen, Francesco Faiola, Timothy A. Hore, et al. (2013). "NANOG-dependent function of TET1 and TET2 in establishment of pluripotency". In: *Nature* 495.7441, pp. 370–374.

Costello, Joseph F., Michael C. Frühwald, Dominic J. Smiraglia, Laura J. Rush, Gavin P. Robertson, et al. (2000). "Aberrant CpG-island methylation has non-random and tumour-type–specific patterns". In: *Nat. Genet.* 24.2, pp. 132–138.

Cubeñas-Potts, Caelin and Michael J. Matunis (2013). "SUMO: A Multifaceted Modifier of Chromatin Structure and Function". In: *Dev. Cell* 24.1, pp. 1–12.

Cui, Chenghua, Ying Gan, Liankun Gu, James Wilson, Zhaojun Liu, et al. (2015). "P16-specific DNA methylation by engineered zinc finger methyltransferase inactivates gene transcription and promotes cancer metastasis". In: *Genome Biol.* 16.1, p. 252.

D'Urso, Agustina and Jason H Brickner (2014). "Mechanisms of epigenetic memory." In: *Trends Genet.* 30.6, pp. 230–6.

Dang, L., K. Yen, and E. C. Attar (2016). "IDH mutations in cancer and progress toward development of targeted therapeutics". In: *Ann. Oncol.* 27.4, pp. 599–608.

Datlinger, Paul, André F Rendeiro, Christian Schmidl, Thomas Krausgruber, Peter Traxler, et al. (2017). "Pooled CRISPR screening with single-cell transcriptome readout". In: *Nat. Methods* 14.3, pp. 297–301.

Davis, Carrie A, Benjamin C Hitz, Cricket A Sloan, Esther T Chan, Jean M Davidson, et al. (2018). "The Encyclopedia of DNA elements (ENCODE): data portal update". In: *Nucleic Acids Res.* 46.D1, pp. D794–D801.

DeAngelis, Margaret M., David G. Wang, and Trevor L. Hawkins (1995). "Solid-phase reversible immobilization for the isolation of PCR products". In: *Nucleic Acids Res.* 23.22, pp. 4742–4743.

Deshpande, Aniruddha J., Anagha Deshpande, Amit U. Sinha, Liying Chen, Jenny Chang, et al. (2014). "AF10 Regulates Progressive H3K79 Methylation and HOX Gene Expression in Diverse AML Subtypes". In: *Cancer Cell* 26.6, pp. 896–908.

Deursen, Jan M. van (2019). "Senolytic therapies for healthy longevity". In: *Science (80-.).* 364.6441, pp. 636–637.

Dhawan, Sangeeta, Senta Georgia, Shuen-ing Tschen, Guoping Fan, and Anil Bhushan (2011). "Pancreatic β Cell Identity Is Maintained by DNA Methylation-Mediated Repression of Arx". In: *Dev. Cell* 20.4, pp. 419–429.

Dhayalan, Arunkumar, Arumugam Rajavelu, Philipp Rathert, Raluca Tamas, Renata Z. Jurkowska, et al. (2010). "The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation". In: *J. Biol. Chem.* 285.34, pp. 26114–26120.

Doudna, J. A. and E. Charpentier (2014). "The new frontier of genome engineering with CRISPR-Cas9". In: *Science (80-.).* 346.6213, pp. 1258096–1258096.

Du, Quan, Zhen Wang, and Vern L Schramm (2016). "Human DNMT1 transition state structure." In: *Proc. Natl. Acad. Sci. U. S. A.* 113.11, pp. 2916–21.

Easwaran, Hariharan, Hsing Chen Tsai, and Stephen B. Baylin (2014). "Cancer Epigenetics: Tumor Heterogeneity, Plasticity of Stem-like States, and Drug Resistance". In: *Mol. Cell* 54.5, pp. 716–727.

Eid, Ayman, Sahar Alshareef, and Magdy M Mahfouz (2018). "CRISPR base editors: genome editing without double-stranded breaks." In: *Biochem. J.* 475.11, pp. 1955–1964.

Eisenberg, Eli and Erez Y Levanon (2013). "Human housekeeping genes, revisited." In: *Trends Genet.* 29.10, pp. 569–74.

Emmerik, Clara L. van and Hugo van Ingen (2019). "Unspinning chromatin: Revealing the dynamic nucleosome landscape by NMR". In: *Prog. Nucl. Magn. Reson. Spectrosc.* 110, pp. 1–19.

Emperle, Max, Arumugam Rajavelu, Richard Reinhardt, Renata Z. Jurkowska, and Albert Jeltsch (2014). "Cooperative DNA Binding and Protein/DNA Fiber Formation Increases the Activity of the Dnmt3a DNA Methyltransferase". In: *J. Biol. Chem.* 289.43, pp. 29602–29613.

Ernst, Jason and Manolis Kellis (2010). "Discovery and characterization of chromatin states for systematic annotation of the human genome". In: *Nat. Biotechnol.* 28.8, pp. 817–825.

Esteller, Manel (2008). "Epigenetics in Cancer". In: *N. Engl. J. Med.* 358.11, pp. 1148–1159.

Esvelt, Kevin M, Prashant Mali, Jonathan L Braff, Mark Moosburner, Stephanie J Yaung, and George M Church (2013). "Orthogonal Cas9 proteins for RNA-guided gene regulation and editing". In: *Nat. Methods* 10.11, pp. 1116–1121.

Faircloth, B and T Glenn (2011). *Serapure*. URL: `https://ethanomics.files.wordpress.com/2012/08/serapure_v2-2.pdf` (visited on 03/12/2019).

Felle, Max, Helen Hoffmeister, Julia Rothammer, Andreas Fuchs, Josef H Exler, and Gernot Längst (2011). "Nucleosomes protect DNA from DNA methylation in vivo and in vitro." In: *Nucleic Acids Res.* 39.16, pp. 6956–69.

Feng, Y, C C Broder, P E Kennedy, and E A Berger (1996). "HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor." In: *Science* 272.5263, pp. 872–7.

Fernández, José María, Victor de la Torre, David Richardson, Romina Royo, Montserrat Puiggròs, et al. (2016). "The BLUEPRINT Data Analysis Portal." In: *Cell Syst.* 3.5, 491–495.e5.

Fisher, Sheila, Andrew Barry, Justin Abreu, Brian Minie, Jillian Nolan, et al. (2011). "A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries". In: *Genome Biol.* 12.1, R1.

Flavahan, William A., Elizabeth Gaskell, and Bradley E. Bernstein (2017). "Epigenetic plasticity and the hallmarks of cancer". In: *Science (80-.).* 357.6348, eaal2380.

Fomenkov, Alexey, Tamas Vincze, Sergey K. Degtyarev, and Richard J. Roberts (2017). "Complete Genome Sequence and Methylome Analysis of Acinetobacter calcoaceticus 65". In: *Genome Announc.* 5.12.

Ford, Ethan, Matthew R. Grimmer, Sabine Stolzenburg, Ozren Bogdanovic, Alex de Mendoza, et al. (2017). "Frequent lack of repressive capacity of promoter DNA methylation identified through genome-wide epigenomic manipulation". In: *bioRxiv*, p. 170506.

Foster, S A, D J Wong, M T Barrett, and D A Galloway (1998). "Inactivation of p16 in human mammary epithelial cells by CpG island methylation." In: *Mol. Cell. Biol.* 18.4, pp. 1793–801.

Friedman-Mazursky, Orr, Ran Elkon, and Shimon Efrat (2016). "Redifferentiation of expanded human islet β cells by inhibition of ARX". In: *Sci. Rep.* 6.1, p. 20698.

Fu, Yanfang, Jeffry D Sander, Deepak Reyon, Vincent M Cascio, and J Keith Joung (2014). "Improving CRISPR-Cas nuclease specificity using truncated guide RNAs". In: *Nat. Biotechnol.* 32.3, pp. 279–284.

Gaj, Thomas, Charles A. Gersbach, and Carlos F. Barbas (2013). "ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering". In: *Trends Biotechnol.* 31.7, pp. 397–405.

Gantz, Valentino M., Nijole Jasinskiene, Olga Tatarenkova, Aniko Fazekas, Vanessa M. Macias, et al. (2015). "Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi". In: *Proc. Natl. Acad. Sci.* 112.49, E6736–E6743.

Gilan, Omer, Enid Y N Lam, Isabelle Becher, Dave Lugo, Ester Cannizzaro, et al. (2016). "Functional interdependence of BRD4 and DOT1L in MLL leukemia". In: *Nat. Struct. Mol. Biol.* 23.7, pp. 673–681.

Gilbert, Luke A., Matthew H. Larson, Leonardo Morsut, Zairan Liu, Gloria A. Brar, et al. (2013). "CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes". In: *Cell* 154.2, pp. 442–451.

Goldberg, Aaron D., C. David Allis, and Emily Bernstein (2007). "Epigenetics: A Landscape Takes Shape". In: *Cell* 128.4, pp. 635–638.

Grebien, Florian, Masoud Vedadi, Matthäus Getlik, Roberto Giambruno, Amit Grover, et al. (2015). "Pharmacological targeting of the Wdr5-MLL interaction in C/EBPα N-terminal leukemia". In: *Nat. Chem. Biol.* 11.8, pp. 571–578.

Groote, Marloes L de, Pernette J Verschure, and Marianne G Rots (2012). "Epigenetic Editing: targeted rewriting of epigenetic marks to modulate expression of selected target genes." In: *Nucleic Acids Res.* 40.21, pp. 10596–613.

Gun, B T F van der, C Huisman, S Stolzenburg, H G Kazemier, M H J Ruiters, et al. (2013). "Bidirectional modulation of endogenous EpCAM expression to unravel its function in ovarian cancer". In: *Br. J. Cancer* 108.4, pp. 881–886.

Gun, Bernardina T.F. van der, Lieuwe J. Melchers, Marcel H.J. Ruiters, Lou F.M.H. de Leij, Pamela M.J. McLaughlin, and Marianne G. Rots (2010). "EpCAM in carcinogenesis: The good, the bad or the ugly". In: *Carcinogenesis* 31.11, pp. 1913–1921.

Guo, Xue, Ling Wang, Jie Li, Zhanyu Ding, Jianxiong Xiao, et al. (2015). "Structural insight into autoinhibition and histone H3-induced activation of DNMT3A". In: *Nature* 517.7536, pp. 640–644.

Haeussler, Maximilian, Kai Schönig, Hélène Eckert, Alexis Eschstruth, Joffrey Mianné, et al. (2016). "Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR". In: *Genome Biol.* 17.1, p. 148.

Hammond, Andrew, Roberto Galizi, Kyros Kyrou, Alekos Simoni, Carla Siniscalchi, et al. (2016). "A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae". In: *Nat. Biotechnol.* 34.1, pp. 78–83.

Hanly, David J, Manel Esteller, and María Berdasco (2018). "Interplay between long non-coding RNAs and epigenetic machinery: emerging targets in cancer?" In: *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 373.1748.

Hanson, Eric S, Mindy L Rawlins, and Elizabeth A Leibold (2003). "Oxygen and iron regulation of iron regulatory protein 2." In: *J. Biol. Chem.* 278.41, pp. 40337–42.

Hashimoto, Hideharu, June E. Pais, Xing Zhang, Lana Saleh, Zheng-Qing Fu, et al. (2014). "Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA". In: *Nature* 506.7488, pp. 391–395.

Hathaway, Nathaniel A, Oliver Bell, Courtney Hodges, Erik L Miller, Dana S Neel, and Gerald R Crabtree (2012). "Dynamics and memory of heterochromatin in living cells." In: *Cell* 149.7, pp. 1447–60.

Hawkins, Trevor L., Tarra O'Connor-Morin, Aparna Roy, and Cynthia Santillan (1994). "DNA purification and isolation using a solid-phase". In: *Nucleic Acids Res.* 22.21, pp. 4543–4544.

He, Y.-F., B.-Z. Li, Z. Li, P. Liu, Y. Wang, et al. (2011). "Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA". In: *Science (80-.).* 333.6047, pp. 1303–1307.

Heerboth, Sarah, Karolina Lapinska, Nicole Snyder, Meghan Leary, Sarah Rollinson, and Sibaji Sarkar (2014). "Use of epigenetic drugs in disease: an overview." In: *Genet. Epigenet.* 6, pp. 9–19.

Heigwer, Florian, Grainne Kerr, and Michael Boutros (2014). "E-CRISP: fast CRISPR target site identification". In: *Nat. Methods* 11.2, pp. 122–123.

Hess, Gaelen T, Laure Frésard, Kyuho Han, Cameron H Lee, Amy Li, et al. (2016). "Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells". In: *Nat. Methods* 13.12, pp. 1036–1042.

Hilton, Isaac B, Anthony M D'Ippolito, Christopher M Vockley, Pratiksha I Thakore, Gregory E Crawford, et al. (2015). "Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers". In: *Nat. Biotechnol.* 33.5, pp. 510–517.

Holoch, Daniel and Raphaël Margueron (2017). "Mechanisms Regulating PRC2 Recruitment and Enzymatic Activity." In: *Trends Biochem. Sci.* 42.7, pp. 531–542.

Hore, Timothy Alexander, Ferdinand von Meyenn, Mirunalini Ravichandran, Martin Bachman, Gabriella Ficz, et al. (2016). "Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naïve pluripotency by complementary mechanisms." In: *Proc. Natl. Acad. Sci. U. S. A.* 113.43, pp. 12202–12207.

Horlbeck, Max A, Lea B Witkowsky, Benjamin Guglielmi, Joseph M Replogle, Luke A Gilbert, et al. (2016). "Nucleosomes impede Cas9 access to DNA in vivo and in vitro". In: *Elife* 5.

Horsthemke, Bernhard (2018). "A critical view on transgenerational epigenetic inheritance in humans". In: *Nat. Commun.* 9.1, p. 2973.

Horvath, Steve (2013). "DNA methylation age of human tissues and cell types". In: *Genome Biol.* 14.10, R115.

Horvath, Steve and Kenneth Raj (2018). "DNA methylation-based biomarkers and the epigenetic clock theory of ageing". In: *Nat. Rev. Genet.* 19.6, pp. 371–384.

Hsu, Patrick D., Eric S. Lander, and Feng Zhang (2014). "Development and Applications of CRISPR-Cas9 for Genome Engineering". In: *Cell* 157.6, pp. 1262–1278.

Hu, Lulu, Ze Li, Ping Wang, Yan Lin, and Yanhui Xu (2011). "Crystal structure of PHD domain of UHRF1 and insights into recognition of unmodified histone H3 arginine residue 2". In: *Cell Res.* 21.9, p. 1374.

Hu, Lulu, Junyan Lu, Jingdong Cheng, Qinhui Rao, Ze Li, et al. (2015). "Structural insight into substrate preference for TET-mediated oxidation". In: *Nature* 527.7576, pp. 118–122.

Hyun, Kwangbeom, Jongcheol Jeon, Kihyun Park, and Jaehoon Kim (2017). "Writing, erasing and reading histone lysine methylations". In: *Exp. Mol. Med.* 49.4, e324–e324.

Iida, Tetsuo, Isao Suetake, Shoji Tajima, Hiroshi Morioka, Satoshi Ohta, et al. (2002). "PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA." In: *Genes Cells* 7.10, pp. 997–1007.

Jeltsch, Albert and Renata Z. Jurkowska (2014). "New concepts in DNA methylation". In: *Trends Biochem. Sci.* 39.7, pp. 310–318.

Jeltsch, Albert and Renata Z. Jurkowska (2016). "Allosteric control of mammalian DNA methyltransferases – a new regulatory paradigm". In: *Nucleic Acids Res.* 44.18, pp. 8556–8575.

Jerram, Samuel T, Mary N Dang, and R David Leslie (2017). "The Role of Epigenetics in Type 1 Diabetes." In: *Curr. Diab. Rep.* 17.10, p. 89.

Jia, Da, Renata Z. Jurkowska, Xing Zhang, Albert Jeltsch, and Xiaodong Cheng (2007). "Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation". In: *Nature* 449.7159, pp. 248–251.

Jin, Seung-Gi, Zhi-Min Zhang, Thomas L. Dunwell, Matthew R. Harter, Xiwei Wu, et al. (2016). "Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration". In: *Cell Rep.* 14.3, pp. 493–505.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier (2012). "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity". In: *Science (80-.).* 337.6096, pp. 816–821.

Jolivet, Philippe and Joseph W Foley (2015). *Solutions for purifying nucleic acids by solidphase reversible immobilization (SPRI).*

Jones, Peter A. (2012). "Functions of DNA methylation: islands, start sites, gene bodies and beyond". In: *Nat. Rev. Genet.* 13.7, pp. 484–492.

Jones, Peter A. and Stephen B. Baylin (2002). "The fundamental role of epigenetic events in cancer". In: *Nat. Rev. Genet.* 3.6, pp. 415–428.

Joung, J Keith and Jeffry D Sander (2013). "TALENs: a widely applicable technology for targeted genome editing." In: *Nat. Rev. Mol. Cell Biol.* 14.1, pp. 49–55.

Joung, Julia, Silvana Konermann, Jonathan S Gootenberg, Omar O Abudayyeh, Randall J Platt, et al. (2017). "Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening". In: *Nat. Protoc.* 12.4, pp. 828–863.

Jurkowska, Renata Z., Tomasz P. Jurkowski, and Albert Jeltsch (2011). "Structure and Function of Mammalian DNA Methyltransferases". In: *ChemBioChem* 12.2, pp. 206–222.

Jurkowski, Tomasz P, Mirunalini Ravichandran, and Peter Stepper (2015). "Synthetic epigenetics—towards intelligent control of epigenetic states and cell identity". In: *Clin. Epigenetics* 7.1, p. 18.

Kang, Joo-Young, Ji-Young Kim, Kee-Beom Kim, Jin Woo Park, Hana Cho, et al. (2018). "KDM2B is a histone H3K79 demethylase and induces transcriptional repression via sirtuin-1–mediated chromatin silencing". In: *FASEB J.* 32.10, pp. 5737–5750.

Kearns, N. A., R. M. J. Genga, M. S. Enuameh, M. Garber, S. A. Wolfe, and R. Maehr (2013). "Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells". In: *Development* 141.1, pp. 219–223.

Kearns, Nicola A, Hannah Pham, Barbara Tabak, Ryan M Genga, Noah J Silverstein, et al. (2015). "Functional annotation of native enhancers with a Cas9–histone demethylase fusion". In: *Nat. Methods* 12.5, pp. 401–403.

Kent, W. James, Charles W Sugnet, Terrence S Furey, Krishna M Roskin, Tom H Pringle, et al. (2002). "The human genome browser at UCSC". In: *Genome Res.* 12.6, pp. 996–1006. arXiv: 0109172 [cond-mat].

Khorasanizadeh, Sepideh (2004). "The nucleosome: from genomic organization to genomic regulation." In: *Cell* 116.2, pp. 259–72.

Kim, J. W., L. U. Kim, and C. K. Kim (2007). "Size Control of Silica Nanoparticles and Their Surface Treatment for Fabrication of Dental Nanocomposites". In: *Biomacromolecules* 8.1, pp. 215–222.

Kim, Jin-Soo, Jaesang Kim, Karyn L. Cepek, Phillip A. Sharp, and Carl O. Pabo (1997). "Design of TATA box-binding protein/zinc finger fusions for targeted regulation of gene expression". In: *Proc. Natl. Acad. Sci. U. S. A.* 94.8, p. 3616.

Kim, Y G, J Cha, and S Chandrasegaran (1996). "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain." In: *Proc. Natl. Acad. Sci. U. S. A.* 93.3, pp. 1156–60.

Kirkland, James L., Tamara Tchkonia, Yi Zhu, Laura J. Niedernhofer, and Paul D. Robbins (2017). "The Clinical Potential of Senolytic Drugs". In: *J. Am. Geriatr. Soc.* 65.10, pp. 2297–2301.

Klann, Tyler S, Joshua B Black, Malathi Chellappan, Alexias Safi, Lingyun Song, et al. (2017). "CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome." In: *Nat. Biotechnol.* 35.6, pp. 561–568.

Konermann, Silvana, Mark D. Brigham, Alexandro E. Trevino, Patrick D. Hsu, Matthias Heidenreich, et al. (2013). "Optical control of mammalian endogenous transcription and epigenetic states". In: *Nature* 500.7463, pp. 472–476.

Konermann, Silvana, Mark D. Brigham, Alexandro E. Trevino, Julia Joung, Omar O. Abudayyeh, et al. (2015). "Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex". In: *Nature* 517.7536, pp. 583–588.

Kong, Xiangqian, Jie Chen, Wenbing Xie, Stephen M. Brown, Yi Cai, et al. (2019). "Defining UHRF1 Domains that Support Maintenance of Human Colon Cancer DNA Methylation and Oncogenic Properties". In: *Cancer Cell* 35.4, 633–648.e7.

Kooistra, Susanne Marije and Kristian Helin (2012). "Molecular mechanisms and potential functions of histone demethylases". In: *Nat. Rev. Mol. Cell Biol.* 13.5, pp. 297–311.

Kungulovski, Goran and Albert Jeltsch (2016). "Epigenome Editing: State of the Art, Concepts, and Perspectives". In: *Trends Genet.* 32.2, pp. 101–113.

Kungulovski, Goran, Suneetha Nunna, Maria Thomas, Ulrich M Zanger, Richard Reinhardt, and Albert Jeltsch (2015). "Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained". In: *Epigenetics Chromatin* 8.1, p. 12.

Laird, C. D., N. D. Pleasant, A. D. Clark, J. L. Sneeden, K. M. A. Hassan, et al. (2004). "Hairpin-bisulfite PCR: Assessing epigenetic methylation patterns on complementary strands of individual DNA molecules". In: *Proc. Natl. Acad. Sci.* 101.1, pp. 204–209.

Lareau, Caleb A., Fabiana M. Duarte, Jennifer G. Chew, Vinay K. Kartha, Zach D. Burkett, et al. (2019). "Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility". In: *Nat. Biotechnol.* P. 1.

Lee, Jungjoon K., Euihwan Jeong, Joonsun Lee, Minhee Jung, Eunji Shin, et al. (2018). "Directed evolution of CRISPR-Cas9 to increase its specificity". In: *Nat. Commun.* 9.1, p. 3048.

Li, Jin, Tamara Casteels, Thomas Frogne, Camilla Ingvorsen, Christian Honoré, et al. (2017). "Artemisinins Target GABAA Receptor Signaling and Impair α Cell Identity." In: *Cell* 168.1-2, 86–100.e15.

Li, Wei, Fei Teng, Tianda Li, and Qi Zhou (2013). "Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems". In: *Nat. Biotechnol.* 31.8, pp. 684–686.

Liang, Puping, Yanwen Xu, Xiya Zhang, Chenhui Ding, Rui Huang, et al. (2015). "CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes". In: *Protein Cell* 6.5, pp. 363–372.

Ling, Charlotte and Tina Rönn (2019). "Epigenetics in Human Obesity and Type 2 Diabetes". In: *Cell Metab.* 29.5, pp. 1028–1044.

Lombardi, Laura Marie, Steven Andrew Baker, and Huda Yahya Zoghbi (2015). "MECP2 disorders: from the clinic to mice and back". In: *J. Clin. Invest.* 125.8, pp. 2914–2923.

Lowe, Robert, Marita G. Overhoff, Sreeram V. Ramagopalan, James C. Garbe, James Koh, et al. (2015). "The senescent methylome and its relationship with cancer, ageing and germline genetic variation in humans". In: *Genome Biol.* 16.1, p. 194.

Luger, Karolin, Armin W. Mäder, Robin K. Richmond, David F. Sargent, and Timothy J. Richmond (1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution". In: *Nature* 389.6648, pp. 251–260.

Lukashevich, Olga V., Natalia A. Cherepanova, Renata Z. Jurkovska, Albert Jeltsch, and Elizaveta S. Gromova (2016). "Conserved motif VIII of murine DNA methyltransferase Dnmt3a is essential for methylation activity". In: *BMC Biochem.* 17.1, p. 7.

Lungu, Cristiana-Elena (2018). "Regulation and readout of mammalian DNA methylation". PhD thesis. University of Stuttgart, p. 218.

Lyst, Matthew J. and Adrian Bird (2015). "Rett syndrome: a complex disorder with simple roots". In: *Nat. Rev. Genet.* 16.5, pp. 261–275.

Ma, Hanhui, Li-Chun Tu, Ardalan Naseri, Maximiliaan Huisman, Shaojie Zhang, et al. (2016). "Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow". In: *Nat. Biotechnol.* 34.5, pp. 528–530.

Maeder, Morgan L, Samantha J Linder, Vincent M Cascio, Yanfang Fu, Quan H Ho, and J Keith Joung (2013). "CRISPR RNA–guided activation of endogenous human genes". In: *Nat. Methods* 10.10, pp. 977–979.

Majidi, Sima, Fatemeh Zeinali Sehrig, Samad Mussa Farkhani, Mehdi Soleymani Goloujeh, and Abolfazl Akbarzadeh (2016). "Current methods for synthesis of magnetic nanoparticles". In: *Artif. Cells, Nanomedicine, Biotechnol.* 44.2, pp. 722–734.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, et al. (2013). "RNA-Guided Human Genome Engineering via Cas9". In: *Science (80-.).* 339.6121, pp. 823–826.

Mali, Prashant, Kevin M Esvelt, and George M Church (2013). "Cas9 as a versatile tool for engineering biology". In: *Nat. Methods* 10.10, pp. 957–963.

Marchbank, Tania, Asif Mahmood, Sarah Harten, Patrick H Maxwell, and Raymond J Playford (2011). "Dimethyloxalyglycine stimulates the early stages of gastrointestinal repair processes through VEGF-dependent mechanisms". In: *Lab. Investig.* 91.12, pp. 1684–1694.

Mayer, Wolfgang, Alain Niveleau, Jörn Walter, Reinald Fundele, and Thomas Haaf (2000). "Demethylation of the zygotic paternal genome". In: *Nature* 403.6769, pp. 501–502.

McDonald, James I., Hamza Celik, Lisa E. Rois, Gregory Fishberger, Tolison Fowler, et al. (2016). "Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation". In: *Biol. Open* 5.6, pp. 866–874.

Meas, Rithy and Peng Mao (2015). "Histone ubiquitylation and its roles in transcription and DNA damage response". In: *DNA Repair (Amst).* 36, pp. 36–42.

Meilinger, Daniela, Karin Fellinger, Sebastian Bultmann, Ulrich Rothbauer, Ian Marc Bonapace, et al. (2009). "Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells". In: *EMBO Rep.* 10.11, pp. 1259–1264.

Melamed, Philippa, Yahav Yosefzon, Cfir David, Anna Tsukerman, and Lilach Pnueli (2018). "Tet Enzymes, Variants, and Differential Effects on Function." In: *Front. cell Dev. Biol.* 6, p. 22.

Meulen, Talitha van der, Sharon Lee, Els Noordeloos, Cynthia J Donaldson, Michael W Adams, et al. (2018). "Artemether Does Not Turn α Cells into β Cells." In: *Cell Metab.* 27.1, 218–225.e4.

Milutinovic, Snezana, Ana C. D'Alessio, Nancy Detich, and Moshe Szyf (2007). "Valproate induces widespread epigenetic reprogramming which involves demethylation of specific genes". In: *Carcinogenesis* 28.3, pp. 560–571.

Mlambo, Tafadzwa, Sandra Nitsch, Markus Hildenbeutel, Marianna Romito, Maximilian Müller, et al. (2018). "Designer epigenome modifiers enable robust and sustained gene silencing in clinically relevant human cells". In: *Nucleic Acids Res.* 46.9, pp. 4456–4468.

Morita, Sumiyo, Hirofumi Noguchi, Takuro Horii, Kazuhiko Nakabayashi, Mika Kimura, et al. (2016). "Targeted DNA demethylation in vivo using dCas9–peptide repeat and scFv–TET1 catalytic domain fusions". In: *Nat. Biotechnol.* 34.10, pp. 1060–1065.

Morselli, Marco, William A Pastor, Barbara Montanini, Kevin Nee, Roberto Ferrari, et al. (2015). "In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse". In: *Elife* 4, e06205.

Moscou, M. J. and A. J. Bogdanove (2009). "A Simple Cipher Governs DNA Recognition by TAL Effectors". In: *Science (80-.).* 326.5959, pp. 1501–1501.

Münz, Markus, Cuong Kieu, Brigitte Mack, Bärbel Schmitt, Reinhard Zeidler, and Olivier Gires (2004). "The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation". In: *Oncogene* 23.34, pp. 5748–5758.

Musselman, Catherine A, Marie-Eve Lalonde, Jacques Côté, and Tatiana G Kutateladze (2012). "Perceiving the epigenetic landscape through histone readers". In: *Nat. Struct. Mol. Biol.* 19.12, pp. 1218–1227.

Myers, Samuel A., Jason Wright, Ryan Peckner, Brian T. Kalish, Feng Zhang, and Steven A. Carr (2018). "Discovery of proteins associated with a predefined genomic locus via dCas9–APEX-mediated proximity labeling". In: *Nat. Methods* 15.6, pp. 437–439.

Nakayama, J.-i., J C Rice, B D Strahl, C D Allis, and S I Grewal (2001). "Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly". In: *Science (80-.).* 292.5514, pp. 110–113.

Neri, Francesco, Danny Incarnato, Anna Krepelova, Stefania Rapelli, Andrea Pagnani, et al. (2013). "Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells". In: *Genome Biol.* 14.8, R91.

Neri, Francesco, Stefania Rapelli, Anna Krepelova, Danny Incarnato, Caterina Parlato, et al. (2017). "Intragenic DNA methylation prevents spurious transcription initiation". In: *Nature* 543.7643, pp. 72–77.

Nguyen, Carvell T, Daniel J Weisenberger, Mihaela Velicescu, Felicidad A Gonzales, Joy C Y Lin, et al. (2002). "Histone H3-lysine 9 methylation is associated with aberrant gene silencing in cancer cells and is rapidly reversed by 5-aza-2'-deoxycytidine." In: *Cancer Res.* 62.22, pp. 6456–61.

Ni, Peng, Neng Huang, Zhi Zhang, De-Peng Wang, Fan Liang, et al. (2019). "DeepSignal: detecting DNA methylation state from Nanopore sequencing reads using deep-learning". In: *Bioinformatics*. Ed. by Arne Elofsson.

Nica, Alexandra C, Halit Ongen, Jean-Claude Irminger, Domenico Bosco, Thierry Berney, et al. (2013). "Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome." In: *Genome Res.* 23.9, pp. 1554–62.

Nishiyama, Atsuya, Luna Yamaguchi, Jafar Sharif, Yoshikazu Johmura, Takeshi Kawamura, et al. (2013). "Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication". In: *Nature* 502.7470, pp. 249–253.

Nobeyama, Yoshimasa and Hidemi Nakagawa (2016). "Aberrant demethylation and expression of MAGEB2 in a subset of malignant peripheral nerve sheath tumors from neurofibromatosis type 1". In: *J. Dermatol. Sci.* 81.2, pp. 118–123.

Normile, Dennis (2018). "Shock greets claim of CRISPR-edited babies." In: *Science (80-.).* 362.6418, pp. 978–979.

Nuland, R. van, A. H. Smits, P. Pallaki, P. W. T. C. Jansen, M. Vermeulen, and H. T. M. Timmers (2013). "Quantitative Dissection and Stoichiometry Determination of the Human SET1/MLL Histone Methyltransferase Complexes". In: *Mol. Cell. Biol.* 33.10, pp. 2067–2077.

Nunna, Suneetha, Richard Reinhardt, Sergey Ragozin, and Albert Jeltsch (2014). "Targeted Methylation of the Epithelial Cell Adhesion Molecule (EpCAM) Promoter to Silence Its Expression in Ovarian Cancer Cells". In: *PLoS One* 9.1. Ed. by Jorg Tost, e87703.

O'Geen, Henriette, Sofie L. Bates, Sakereh S. Carter, Karly A. Nisson, Julian Halmai, 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.

Oberacker, Phil, Peter Stepper, Donna M. Bond, Sven Höhn, Jule Focken, et al. (2019). "Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation". In: *PLOS Biol.* 17.1, e3000107.

Osta, Walid A., Yian Chen, Kaidi Mikhitarian, Michael Mitas, Mohamed Salem, et al. (2004). "EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy". In: *Cancer Res.* 64.16, pp. 5818–5824.

Pabo, Carl O., Ezra Peisach, and Robert A. Grant (2001). "Design and Selection of Novel Cys2His2 Zinc Finger Proteins". In: *Annu. Rev. Biochem.* 70.1, pp. 313–340.

Pagliuca, Felicia W., Jeffrey R. Millman, Mads Gürtler, Michael Segel, Alana Van Dervort, et al. (2014). "Generation of Functional Human Pancreatic β Cells In Vitro". In: *Cell* 159.2, pp. 428–439.

Pal, Sangita and Jessica K Tyler (2016). "Epigenetics and aging." In: *Sci. Adv.* 2.7, e1600584.

Pastor, William A, L Aravind, and Anjana Rao (2013). "TETonic shift: biological roles of TET proteins in DNA demethylation and transcription." In: *Nat. Rev. Mol. Cell Biol.* 14.6, pp. 341–56.

Peche, Leticia Y, María F Ladelfa, María F Toledo, Miguel Mano, Julieta E Laiseca, et al. (2015). "Human MageB2 Protein Expression Enhances E2F Transcriptional Activity, Cell Proliferation, and Resistance to Ribotoxic Stress." In: *J. Biol. Chem.* 290.49, pp. 29652–62.

Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR". In: *Nucleic Acids Res.* 29.9.

Pickar-Oliver, Adrian and Charles A. Gersbach (2019). "The next generation of CRISPR–Cas technologies and applications". In: *Nat. Rev. Mol. Cell Biol.* P. 1.

Qi, Lei S, Matthew H Larson, Luke A Gilbert, Jennifer A Doudna, Jonathan S Weissman, et al. (2013). "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." In: *Cell* 152.5, pp. 1173–83.

Qin, Weihua, Heinrich Leonhardt, and Fabio Spada (2011). "Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1". In: *J. Cell. Biochem.* 112.2, pp. 439–444.

Rajavelu, Arumugam, Renata Z. Jurkowska, Jürgen Fritz, and Albert Jeltsch (2012). "Function and disruption of DNA Methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation". In: *Nucleic Acids Res.* 40.2, pp. 569–580.

Rajavelu, Arumugam, Cristiana Lungu, Max Emperle, Michael Dukatz, Alexander Bröhm, et al. (2018). "Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2". In: *Nucleic Acids Res.* 46.17, pp. 9044–9056.

Ravichandran, M., R. Z. Jurkowska, and T. P. Jurkowski (2018). "Target specificity of mammalian DNA methylation and demethylation machinery". In: *Org. Biomol. Chem.* 16.9, pp. 1419–1435.

Ravichandran, Mirunalini (2017). "Biochemical characterisation of TET DNA hydroxylases". Dissertation. University of Stuttgart, p. 153.

Reither, Sabine, Fuyang Li, Humaira Gowher, and Albert Jeltsch (2003). "Catalytic Mechanism of DNA-(cytosine-C5)-methyltransferases Revisited: Covalent Intermediate Formation is not Essential for Methyl Group Transfer by the Murine Dnmt3a Enzyme". In: *J. Mol. Biol.* 329.4, pp. 675–684.

Reyon, Deepak, Shengdar Q Tsai, Cyd Khayter, Jennifer A Foden, Jeffry D Sander, and J Keith Joung (2012). "FLASH assembly of TALENs for high-throughput genome editing". In: *Nat. Biotechnol.* 30.5, pp. 460–465.

Robinson, Philip JJ and Daniela Rhodes (2006). "Structure of the '30nm' chromatin fibre: A key role for the linker histone". In: *Curr. Opin. Struct. Biol.* 16.3, pp. 336–343.

Rohland, Nadin and David Reich (2012). "Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture." In: *Genome Res.* 22.5, pp. 939–46.

Rondelet, Grégoire, Thomas Dal Maso, Luc Willems, and Johan Wouters (2016). "Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B". In: *J. Struct. Biol.* 194.3, pp. 357–367.

Rothbart, S. B., B. M. Dickson, M. S. Ong, K. Krajewski, S. Houliston, et al. (2013). "Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation". In: *Genes Dev.* 27.11, pp. 1288–1298.

Rowley, M. Jordan and Victor G. Corces (2018). "Organizational principles of 3D genome architecture". In: *Nat. Rev. Genet.* 19.12, pp. 789–800.

Roy, Siddhartha and Tapas K. Kundu (2014). "Gene regulatory networks and epigenetic modifications in cell differentiation". In: *IUBMB Life* 66.2, pp. 100–109.

Santos, Fátima, Julian Peat, Heather Burgess, Cristina Rada, Wolf Reik, and Wendy Dean (2013). "Active demethylation in mouse zygotes involves cytosine deamination and base excision repair". In: *Epigenetics Chromatin* 6.1, p. 39.

Saunderson, Emily A., Peter Stepper, Jennifer J. Gomm, Lily Hoa, Adrienne Morgan, et al. (2017). "Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors". In: *Nat. Commun.* 8.1, p. 1450.

Schmitt, Anthony D., Ming Hu, and Bing Ren (2016). "Genome-wide mapping and analysis of chromosome architecture". In: *Nat. Rev. Mol. Cell Biol.* 17.12, pp. 743–755.

Schoenfelder, Stefan and Peter Fraser (2019). "Long-range enhancer–promoter contacts in gene expression control". In: *Nat. Rev. Genet.*

Schultz, D. C., J R Friedman, and F J Rauscher (2001). "Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD". In: *Genes Dev.* 15.4, pp. 428–443.

Schultz, David C., Kasirajan Ayyanathan, Dmitri Negorev, Gerd G. Maul, and Frank J. Rauscher (2002). "SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins". In: *Genes Dev.* 16.8, pp. 919–932.

Sdelci, Sara, André F. Rendeiro, Philipp Rathert, Wanhui You, Jung-Ming G. Lin, et al. (2019). "MTHFD1 interaction with BRD4 links folate metabolism to transcriptional regulation". In: *Nat. Genet.* 51.6, pp. 990–998.

Sen, Payel, Parisha P Shah, Raffaella Nativio, and Shelley L Berger (2016). "Epigenetic Mechanisms of Longevity and Aging." In: *Cell* 166.4, pp. 822–839.

Shalem, Ophir, Neville E. Sanjana, Ella Hartenian, Xi Shi, David A. Scott, et al. (2014). "Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells". In: *Science (80-.).* 343.6166, pp. 84–87.

Shapiro, A.M. James, Camillo Ricordi, Bernhard J. Hering, Hugh Auchincloss, Robert Lindblad, et al. (2006). "International Trial of the Edmonton Protocol for Islet Transplantation". In: *N. Engl. J. Med.* 355.13, pp. 1318–1330.

Sharif, Jafar, Masahiro Muto, Shin-ichiro Takebayashi, Isao Suetake, Akihiro Iwamatsu, et al. (2007). "The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA". In: *Nature* 450.7171, pp. 908–912.

Shen, Ying, Xin Li, Dandan Dong, Bin Zhang, Yanru Xue, and Peng Shang (2018). "Transferrin receptor 1 in cancer: a new sight for cancer therapy." In: *Am. J. Cancer Res.* 8.6, pp. 916–931.

Shiraishi, M. and Hikoya Hayatsu (2004). "High-Speed Conversion of Cytosine to Uracil in Bisulfite Genomic Sequencing Analysis of DNA Methylation". In: *DNA Res.* 11.6, pp. 409–415.

Siddique, Abu Nasar, Suneetha Nunna, Arumugam Rajavelu, Yingying Zhang, Renata Z. Jurkowska, 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: *J. Mol. Biol.* 425.3, pp. 479–491.

Singh, Dhirender, JoEllyn M McMillan, Xin-Ming Liu, Hemant M Vishwasrao, Alexander V Kabanov, et al. (2014). "Formulation design facilitates magnetic nanoparticle delivery to diseased cells and tissues". In: *Nanomedicine* 9.3, pp. 469–485.

Slaymaker, Ian M, Linyi Gao, Bernd Zetsche, David A Scott, Winston X Yan, and Feng Zhang (2016). "Rationally engineered Cas9 nucleases with improved specificity." In: *Science (80-.).* 351.6268, pp. 84–8.

Song, Chun-Xiao, Keith E. Szulwach, Qing Dai, Ye Fu, Shi-Qing Mao, et al. (2013). "Genome-wide Profiling of 5-Formylcytosine Reveals Its Roles in Epigenetic Priming". In: *Cell* 153.3, pp. 678–691.

Stepper, Peter (2015). "Synthetic epigenetics: Targeted methylation of EpCAM promoter by CRISPR-Cas9/Dnmt fusions". MSc-thesis. Universität Stuttgart.

Stepper, Peter, Goran Kungulovski, Renata Z. Jurkowska, Tamir Chandra, Felix Krueger, et al. (2017). "Efficient targeted DNA methylation with chimeric dCas9– Dnmt3a– Dnmt3L methyltransferase". In: *Nucleic Acids Res.* 45.4, pp. 1703–1713.

Sternberg, Samuel H, Sy Redding, Martin Jinek, Eric C Greene, and Jennifer A Doudna (2014). "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9." In: *Nature* 507.7490, pp. 62–7.

Suelves, Mònica, Elvira Carrió, Yaiza Núñez-Álvarez, and Miguel A. Peinado (2016). "DNA methylation dynamics in cellular commitment and differentiation". In: *Brief. Funct. Genomics* 15.6, elw017.

Sun, Xueqing, Guangcun Cheng, Mingang Hao, Jianghua Zheng, Xiaoming Zhou, et al. (2010). "CXCL12 / CXCR4 / CXCR7 chemokine axis and cancer progression". In: *Cancer Metastasis Rev.* 29.4, pp. 709–722.

Susek, Katharina Helene, Maria Karvouni, Evren Alici, and Andreas Lundqvist (2018). "The Role of CXC Chemokine Receptors 1-4 on Immune Cells in the Tumor Microenvironment." In: *Front. Immunol.* 9, p. 2159.

Suvà, Mario L, Nicolo Riggi, and Bradley E Bernstein (2013). "Epigenetic reprogramming in cancer." In: *Science (80-.).* 339.6127, pp. 1567–70.

Taghbalout, Aziz, Menghan Du, Nathaniel Jillette, Wojciech Rosikiewicz, Abhijit Rath, et al. (2019). "Casilio-ME: Enhanced CRISPR-based DNA demethylation by RNA-guided coupling methylcytosine oxidation and DNA repair pathways". In: *bioRxiv*.

Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, et al. (2009). "Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1". In: *Science (80-.).* 324.5929, pp. 930–935.

Tanenbaum, Marvin E., Luke A. Gilbert, Lei S. Qi, Jonathan S. Weissman, and Ronald D. Vale (2014). "A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging". In: *Cell* 159.3, pp. 635–646.

Thakore, Pratiksha I, Joshua B Black, Isaac B Hilton, and Charles A Gersbach (2016). "Editing the epigenome: technologies for programmable transcription and epigenetic modulation". In: *Nat. Methods* 13.2, pp. 127–137.

Tucci, Valter, Anthony R. Isles, Gavin Kelsey, Anne C. Ferguson-Smith, Valter Tucci, et al. (2019). "Genomic Imprinting and Physiological Processes in Mammals". In: *Cell* 176.5, pp. 952–965.

Turner, Bryan M. (2007). "Defining an epigenetic code". In: *Nat. Cell Biol.* 9.1, pp. 2–6.

Van Tongelen, Aurélie, Axelle Loriot, and Charles De Smet (2017). "Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes". In: *Cancer Lett.* 396, pp. 130–137.

Vettese-Dadey, M, P A Grant, T R Hebbes, C Crane- Robinson, C D Allis, and J L Workman (1996). "Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro." In: *EMBO J.* 15.10, pp. 2508–18.

Voigt, Philipp, Gary LeRoy, William J. Drury, Barry M. Zee, Jinsook Son, et al. (2012). "Asymmetrically Modified Nucleosomes". In: *Cell* 151.1, pp. 181–193.

Vojta, Aleksandar, Paula Dobrinić, Vanja Tadić, Luka Bočkor, Petra Korać, et al. (2016). "Repurposing the CRISPR-Cas9 system for targeted DNA methylation". In: *Nucleic Acids Res.* 44.12, pp. 5615–5628.

Waddington, C. H. (1957). *The Strategy of the Genes: A Discussion of some Aspects of theoretical Biology*. George Allen and Unwin, p. 262.

Wang, Tianzhen, Guangyu Wang, Dapeng Hao, Xi Liu, Dong Wang, et al. (2015). "Aberrant regulation of the LIN28A/LIN28B and let-7 loop in human malignant tumors and its effects on the hallmarks of cancer". In: *Mol. Cancer* 14.1, p. 125.

Wang, Tim, Jenny J. Wei, David M. Sabatini, and Eric S. Lander (2014). "Genetic Screens in Human Cells Using the CRISPR-Cas9 System". In: *Science (80-.).* 343.6166, pp. 80–84.

Williams, Ruth M., Upeka Senanayake, Mara Artibani, Gunes Taylor, Daniel Wells, et al. (2018). "Genome and epigenome engineering CRISPR toolkit for in vivo modulation of cis-regulatory interactions and gene expression in the chicken embryo". In: *Development* 145.4, dev160333.

Wolfe, Scot A., Lena Nekludova, and Carl O. Pabo (2000). "DNA Recognition by Cys2His2 Zinc Finger Proteins". In: *Annu. Rev. Biophys. Biomol. Struct.* 29.1, pp. 183–212.

Wu, Hao, Xiaoji Wu, Li Shen, and Yi Zhang (2014). "Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing". In: *Nat. Biotechnol.* 32.12, p. 1231.

Wu, Wei, Zhaohui Wu, Taekyung Yu, Changzhong Jiang, and Woo-Sik Kim (2015). "Recent progress on magnetic iron oxide nanoparticles: synthesis, surface functional strategies and biomedical applications". In: *Sci. Technol. Adv. Mater.* 16.2, p. 023501.

Wu, Xiaoji and Yi Zhang (2017). "TET-mediated active DNA demethylation: mechanism, function and beyond". In: *Nat. Rev. Genet.* 18.9, pp. 517–534.

Xu, Shili and Nouri Neamati (2013). "gp130: a promising drug target for cancer therapy". In: *Expert Opin. Ther. Targets* 17.11, pp. 1303–1328.

Xu, Xingxing, Yonghui Tao, Xiaobo Gao, Lei Zhang, Xufang Li, et al. (2016). "A CRISPR-based approach for targeted DNA demethylation". In: *Cell Discov.* 2.1, p. 16009.

Xue, Busheng, Jiansong Zhao, Penghui Feng, Jia Xing, Hongliang Wu, and Yan Li (2019). "Epigenetic mechanism and target therapy of UHRF1 protein complex in malignancies." In: *Onco. Targets. Ther.* 12, pp. 549–559.

Yamashita, Taro, Anuradha Budhu, Marshonna Forgues, and Xin Wei Wang (2007). "Activation of hepatic stem cell marker EpCAM by Wnt-beta-catenin signaling in hepatocellular carcinoma." In: *Cancer Res.* 67.22, pp. 10831–9.

Yeo, Nan Cher, Alejandro Chavez, Alissa Lance-Byrne, Yingleong Chan, David Menn, et al. (2018). "An enhanced CRISPR repressor for targeted mammalian gene regulation". In: *Nat. Methods* 15.8, pp. 611–616.

Yoder, J A, C P Walsh, and T H Bestor (1997). "Cytosine methylation and the ecology of intragenomic parasites." In: *Trends Genet.* 13.8, pp. 335–40.

Yoshida, K., T. Taga, M. Saito, S. Suematsu, A. Kumanogoh, et al. (1996). "Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders." In: *Proc. Natl. Acad. Sci.* 93.1, pp. 407–411.

Yu, J., M. A. Vodyanik, K. Smuga-Otto, J. Antosiewicz-Bourget, J. L. Frane, et al. (2007). "Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells". In: *Science (80-.).* 318.5858, pp. 1917–1920.

Yu, Miao, Gary C Hon, Keith E Szulwach, Chun-Xiao Song, Liang Zhang, et al. (2012a). "Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome." In: *Cell* 149.6, pp. 1368–80.

Yu, Wenyu, Emma J. Chory, Amy K. Wernimont, Wolfram Tempel, Alex Scopton, et al. (2012b). "Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors". In: *Nat. Commun.* 3.1, p. 1288.

Zalatan, Jesse G., Michael E. Lee, Ricardo Almeida, Luke A. Gilbert, Evan H. Whitehead, et al. (2015). "Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds". In: *Cell* 160.1-2, pp. 339–350.

Zhang, Feng, Le Cong, Simona Lodato, Sriram Kosuri, George M Church, and Paola Arlotta (2011). "Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription." In: *Nat. Biotechnol.* 29.2, pp. 149–53.

Zhang, Yanqing, Genevieve E Fava, Hongjun Wang, Franck Mauvais-Jarvis, Vivian A Fonseca, and Hongju Wu (2016). "PAX4 Gene Transfer Induces α-to-β Cell Phenotypic Conversion and Confers Therapeutic Benefits for Diabetes Treatment". In: *Mol. Ther.* 24.2, pp. 251–260.

Zheng, Yupeng, Paul M. Thomas, and Neil L. Kelleher (2013). "Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites". In: *Nat. Commun.* 4.1, p. 2203.

Zoghbi, Huda Y and Arthur L Beaudet (2016). "Epigenetics and Human Disease." In: *Cold Spring Harb. Perspect. Biol.* 8.2, a019497.

5 Appendix

5.1 List of publications

Tomasz P. Jurkowski#, Mirunalini Ravichandran and **Peter Stepper** (2015). Synthetic epigenetics — towards intelligent control of epigenetic states and cell identity. **Clinical Epigenetics** 7:18

Peter Stepper, Goran Kungulovski, Renata Z. Jurkowska, Tamir Chandra, Felix Krueger, Richard Reinhardt, Wolf Reik, Albert Jeltsch#, Tomasz P. Jurkowski# (2017). Efficient targeted DNA methylation with chimeric dCas9–Dnmt3a–Dnmt3L methyltransferase. **Nucleic Acids Research** 45.4

Emily A. Saunderson, **Peter Stepper**, Jennifer J. Gomm, Lily Hoa, Adrienne Morgan, Michael D. Allen, J. Louise Jones, John G. Gribben, Tomasz P. Jurkowski# & Gabriella Ficz# (2017). Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. **Nature Communications** 8(1):1450

Phil Oberacker*, **Peter Stepper***, Donna M. Bond*, Sven Höhn, Jule Focken, Vivien Meyer, Luca Schelle, Victoria J. Sugrue, Gert-Jan Jeunen, Tim Moser, Steven R. Hore, Ferdinand von Meyenn, Katharina Hipp, Timothy A. Hore#, Tomasz P. Jurkowski# (2019). Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation. **PLOS Biology** 17(1): e3000107

Phil Oberacker*, **Peter Stepper***, Donna M. Bond, Katharina Hipp, Timothy A. Hore#, Tomasz P. Jurkowski# (2019). Simple synthesis of functionalized magnetic beads for nucleic acid purification and manipulation. **Bio-Protocol** 9(20): e3394

*Co-first authors; #Corresponding authors

5.2 Primers and oligonucleotides

The following tables contain the sorted primers, oligonucleotides and gRNA sequences used in this thesis for cloning, targeting, qPCR and bisulfite sequencing.

Table 5.1: qPCR and library amplification primers.

Primer	Sequence
Q-ACTB-f	TCACCAACTGGGACGACATG
Q-ACTB-r	ACCGGAGTCCATCACGATG
Q-EPCAM-f	CTCCATGTGCTGGTGTGT
Q-EPCAM-r	TGTTTTAGTTCAATGATGATCCAGTA
Q-TFRC-f	ACCGGCACCATCAAGCT
Q-TFRC-r	TGATCACGCCAGACTTTGC
Q-TFRC-new-f	AGGGATACCTTTCGTCCCTG
Q-TFRC-new-r	CCGGATGCTTCACATTTTGC
Q-CXCR4-f	AGCATGACGGACAAGTACAGG
Q-CXCR4-r	GATGAAGTCGGGAATAGTCAGC
Q-REEP5-f	TTTGGCTACCCAGCCTACATC
Q-REEP5-r	CAACAGGAAGCCACACTTCAGC
Q-IL6ST-f	AGATAAAGCACCCTGTATCACAGAC
Q-IL6ST-r	TGGAAGGTGGAGCTTGTTTAAG
Q-LIN28A-f	CTGTAAGTGGTTCAACGTGCG
Q-LIN28A-r	CTTCAAGCTCCGGAACCCTT
Q-MAGEB2-f	TCCTGACTTCCGCTTTGGAG
Q-MAGEB2-r	GGGCACGGAGCTTACTCTTC
Q-ASCL1-fn	CGGTCTCATCCTACTCGTCG
Q-ASCL1-rn	GATCACCCTGCTTCCAAAGTC
Q-mReep5-f	GCCATCGAGAGTCCCAACAA
Q-mReep5-r	AGCATCTCAGCCCCATTAGC
Q-mIns2-f	GACCCACAAGTGGCACAAC
Q-mIns2-r	AGAGGGGTAGGCTGGGTAGT
Q-pC-Cas9-f	TTGACGCAAATGGGCGGTAG
Q-pC-Cas9-r	GCACCTTGTACTCGTCCGTA
Q-SPn-Cas9-f	TCAGGCGGCAAGAGGATTTC
Q-SPn-Cas9-r	AGTCATCCACGCGAATCTGG
Em-Ins2-fw	GGAGCGTGGCTTCTTCTACA
Em-Ins2-rev	TTCATTGCAGAGGGGTAGGC
Em-Reep5-fw	TCATCGGACTGGTGGCTTTG
Em-Reep5-rev	GTTGGGACTCTCGATGGCTT
Kapa-f	AATGATACGGCGACCACCGA
Kapa-r	CAAGCAGAAGACGGCATACGA

Table 5.2: Primers used for megaprimer mutagenesis.

Primer	Sequence
D3a-C706A-f	GAGGCAGTCCCGCCAATGACCTCTC
D3a-Cunch-f	GAGGCAGTCCCTGCAATGACCTCTC
D3a-Runch-r	CTTTATAGAGTTTGACCTGGTGGTAATGGTC
D3a-R832E-r	CTTTATAGAGTTTGACTCCGTGGTAATGGTC
mTet3-DM-f2	cccacgcGTacaaggCccaacataacctctac
mTet3-mut-r	cctctctggggaatgctgtgag

Table 5.3: First set of PCR primers and plasmids used to clone the dCas9-ED fusion plasmids.

		insert				vector			
dCas9-ED fusion	source plasmid	primer f	sequence	primer r	sequence	primer f	sequence	primer r	sequence
dCas9-Dnmt3a3L	ZNF-Dnmt3a3L (Siddique et al., 2013)	D3a3L-Cas9-f	AGCGGTTCCGGACGGGCTTCCCCAGGAATTCCCGGGTCGACTCGAAACCATGACCAGGAATTTGACCCC	D3a3L-Cas9-r	CCGGTAGGGATCGAACCCTTTCAGACTTTCCTTTTCTTCTTCGGATCTGCCCTGCTAAGAGGAAGTGAGTTTTGAGAAAAATACTTGAAG	dCas9-vector-f	TGAAAGGGTTCGATCCCTACCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-Dnmt3a	ZNF-Dnmt3a CD (Siddique et al., 2013)	D3a-Cas9-f	AGCGGTTCCGGACGGGCTTCCCCAGGAATTCCCGGGTCGACTCGAATGAACCATGACCAGGAATTTGACCC	D3a-Cas9-r	CCGGTAGGGATCGAACCCTTTCAGACTTTCCTTTTCTTCTTCGGATCTGCCCTGCTCACACAAGCAAAATATTCCTTCAGCGG	dCas9-vector-f	TGAAAGGGTTCGATCCCTACCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-KRAB	pHAGE EF1α dCas9-KRAB (Addgene Plasmid #50919)	KRABtoSPn-f	gcggttccggacgggcttccccaggaattccCgggtcgactcgaATGGACGCGAAATCACTTACG	KRABtoSPn-r	gtagggatcgaaccctttcagactttccttttcttcttcggatctgccctgctTACCAGCCAAGGTTCTTCC	dCas9-vector-f	TGAAAGGGTTCGATCCCTACCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-SID4X	pAAV_hSyn_TALEBB(HD)-NLS-SID4X_2A_phiLOV2.1... (Addgene Plasmid #47450)	SIDtoSPn-f	GCGGTTCCGGACGGGCTccccaggaattccCgggtcgactcgaGAGGCCAGCATGAACATCCA	SIDtoSPn-r	GTAGGGATCGAACCCTTTCAgactttccttttcttcttcggatctgccctgctCTCTCCACTGCCccatgg	dCas9-vector-f	TGAAAGGGTTCGATCCCTACCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
KRAB-dCas9-Dnmt3a3L	pHAGE EF1α dCas9-KRAB and dCas9-Dnmt3a3L	KRABtoNterm-f	gaaccttgccaccATGGACGCGAAATCACTTAC	KRABtoNterm-r	CCACCTTCCGCTTCTTCTTGGGGCTGCCACCTCCTGATACCAGCCAAGGTTCTTCCC	NLStoNtermSPn-f	GAAGAAGCGGAAGGTGGAGGCCAGCGGAGGTGGATCGgacaagaagtactccattggg	toNterm-SPn-r	CATggtggcaagggttcg
dCas9-Dnmt3a3L-KRAB	pHAGE EF1α dCas9-KRAB and dCas9-Dnmt3a3L	KRABtoCterm-F	GAAGAAGCGGAAGGTGGAGGCCAGCGGAGGTGGATCGATGGACGCGAAATC	KRABtoCterm-r	ctaaccggtagggatcgaaccctttcaTACCAGCCAAGGTTCTTCCC	toCterm-vect-f	ttcgatccctaccggttagtaatgag	3a3L-vect-r	CCACCTTCCGCTTCTTCTTGGGGCTGCCACCTCCTGAgactttcctttttcttcttcggatc
dCas9-AID	synthetic gBlock (amino acids 1-169 from NP_033775.1, murine AICDA)	AID-Cas9-f	agcggttccggacgggcttccccaggaattccCgggtcgactcgaatggacagccttctgatgaagcaaaag	AID-Cas9-r	ccggtagggatcgaaccctttcagactttcctttttcttcttcggatctgccctgctggtaccagaattttcatgtagcccttcccagg	dCas9-vector-f	TGAAAGGGTTCGATCCCTACCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-sl-nTet	pEXA2-nTet (synthetic construct from XP_002667965.1)	nTet-cas9-f	AGCGGTTCCGGACGGGCTTCCCCAGGAATTCCCGGGTCGACTCGAATGACGACATTTAAACAACAAACGATTAAAG	nTet-cas9-r	CCGGTAGGGATCGAACCCTTTCAGACTTTCCTTTTCTTCTTCGGATCTGCCCTGCTTTTAGTTTCCTTGTGGATTTTCAACGAG	dCas9-vector-f	TGAAAGGGTTCGATCCCTACCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-nTet	dCas9-sl-nTet	Gly4Ser	agcggttccggacgggctggaggcggcgggagtggaggcggtgggtccggaggtggcgggagcggtggcggagggtcttccccaggaattcccgg	GSvect-r	CCGGGAATTCCTGGGGA	GSvect-f	TCCCCAGGAATTCCCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-nTet A212T	pEt28a+nTet-A212T and dCas9-nTet	nTet-to Linker-f	gaattcccgggtcgactcgaATGACGACATTTAAACAACAAAC	nTet-to Linker-r	cttcttcggatctgccctgctTTAGTTTCCTTGTGGATTTTCAAC	nTet-linker-f	Agcagggcagatccgaagaag	nTet-linker-r	Tcgagtcgacccgggaattc
dCas9-nTet A212V	pEt28a+nTet-A212V and dCas9-nTet	nTet-to Linker-f	cttcttcggatctgccctgctTTAGTTTCCTTGTGGATTTTCAAC	nTet-to Linker-r	cttcttcggatctgccctgctTTAGTTTCCTTGTGGATTTTCAAC	nTet-linker-f	Agcagggcagatccgaagaag	nTet-linker-r	Tcgagtcgacccgggaattc

Table 5.4: Second set of PCR primers and plasmids used to clone the dCas9-ED fusion plasmids.

		insert				vector			
dCas9-ED fusion	source plasmid	primer f	sequence	primer r	sequence	primer f	sequence	primer r	sequence
dCas9-sl-mTet1	pET-28a-His6-mTet1CD (Ravichandran, 2017)	mTet1CD-cas9-f	agcggttccggacgggcttccccaggaattccccgggtcgactcgagaagctgcaccctgtgactgtgatg	mTet1CD-cas9-r	ccggtagggatcgaaccctttcagactttcctttttcttcttcggatctgccctgctgacccaacgattgtagggtcccg	dCas9-vector-f	TGAAAGGGTTCGATCCCTACCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-mTet1	dCas9-sl-mTet1	Gly4Ser	agcggttccggacgggctggaggcggcgggagtggaggcggtgggtccggaggtggcgggagcggtggcggagggtcttccccaggaattcccgg	GSvect-r	CCGGGAATTCCTGGGGA	GSvect-f	TCCCCAGGAATTCCCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-sl-mTet2	pET-28a-His6-mTet2CD (Ravichandran, 2017)	mTet2CD-cas9-f	agcggttccggacgggcttccccaggaattccccgggtcgactcgacaaagtcagaatggcaaatgtgaag	mTet2CD-cas9-r	ccggtagggatcgaaccctttcagactttcctttttcttcttcggatctgccctgcttacaaatgtgttgtaaggccctg	dCas9-vector-f	TGAAAGGGTTCGATCCCTACCGG	dCas9-vector-r	AGCCCGTCCGGAACCGCT
dCas9-mTet3	pET-28a-His6-mTet3CD (Ravichandran, 2017) and dCas9-nTet	mTet3-to Linker-f	Gaattcccgggtcgactcgaatgcgcccttagagttc	mTet3-to Linker-r	cttcttcggatctgccctgctgatccagcggctgtagg	nTet-linker-f	Agcagggcagatccgaagaag	nTet-linker-r	Tcgagtcgacccgggaattc
dCas9-mTet3-del1	pET-28a-His6-mTet3CD-del1 (Ravichandran, 2017) and dCas9-nTet	mTet3-to Linker-f	Gaattcccgggtcgactcgaatgcgcccttagagttc	mTet3-to Linker-r	cttcttcggatctgccctgctgatccagcggctgtagg	nTet-linker-f	Agcagggcagatccgaagaag	nTet-linker-r	Tcgagtcgacccgggaattc
dCas9-mTet3h (del1-ΔC)	pET-28aHis6-mTet3CD-del1 (Ravichandran, 2017) and dCas9-nTet	mTet3-to Linker-f	Gaattcccgggtcgactcgaatgcgcccttagagttc	mTet3-delta-dSPn-r	cttcttcggatctgccctgctgcccaggcgtgcgg	nTet-linker-f	Agcagggcagatccgaagaag	nTet-linker-r	Tcgagtcgacccgggaattc
VPR-dCas9-mTet3del1	dCas9-VPR and dCas9-mTet3-del1	VPR-toNt-f	cgggaccgatccagcctccggactATGGACGCATTGGACGATTTTG	VPR-toNt-r	tccatggtggcaagggttcgatcctGGCCTCCGCTGGTTCCTCCGGAAAACAGAGATGTGTCGAAGATG	XbaI digest			
VPR-dCas9-mTet3h	dCas9-VPR and dCas9-mTet3h	VPR-toNt-f	cgggaccgatccagcctccggactATGGACGCATTGGACGATTTTG	VPR-toNt-r	tccatggtggcaagggttcgatcctGGCCTCCGCTGGTTCCTCCGGAAAACAGAGATGTGTCGAAGATG	XbaI digest			
dCas9-VPR-mTet3-del1	dCas9-mTet3-del1 and dCas9-VPR	Gly4Ser	agcggttccggacgggctggaggcggcgggagtggaggcggtgggtccggaggtggcgggagcggtggcggagggtcttccccaggaattcccgg	mTet3CD-del1-C-r	TTTGTCTGCTCGAAGCATTAtcagatccagcggctgtaggg	VPR-linkr-f	ggatctgaTAATGCTTCGAGCAGACAAAGGG	VPR-link-r	cctccagcccgtccggaaccgctAAACAGAGATGTGTCGAAGATG
dCas9-VPR-mTet3h	dCas9-mTet3-del1 and dCas9-VPR	Gly4Ser	agcggttccggacgggctggaggcggcgggagtggaggcggtgggtccggaggtggcgggagcggtggcggagggtcttccccaggaattcccgg	mTet3CD-deltaC-link-r	TTTGTCTGCTCGAAGCATTAtcagatccctagcccaggcgtgcgg	VPR-linkr-f	ggatctgaTAATGCTTCGAGCAGACAAAGGG	VPR-link-r	cctccagcccgtccggaaccgctAAACAGAGATGTGTCGAAGATG

Table 5.5: gRNA binding sites with the PAM site underlined.

gRNA	Sequence in the genome
ASCL1 sg1	CTGTCCAAGCGGTCGAGCGACGG
ASCL1 sg2	CGGGAGAAAGGAACGGGAGGGGG
ASCL1 sg3	GCAGCCGCTCGCTGCAGCAGCGG
ASCL1 sg4	CGAAGCCAACCCGCGAAGGGAGG
ASCL1 sg5	CCTGTTTCTTTGCCACGGCCGCA
EPCAM sg1	GTTCCCGCACCCTTCCCCCCAGG
EPCAM sg2	GCAGCGCTCCTCCGGTTAAAAGG
EPCAM sg3	GGAGACGAAGCACCTGGGGCGGG
EPCAM sg4	GCGCCCCAACTGCAGCGCCGGGG
EPCAM sg5	GGCCAGAGGTGAGCAGTCCCGGG
EPCAM sg6	GCGAGCGAGCACCTTCGACGCGG
EPCAM sg7	GCCCCAGGCCTCGCGCTGCCCGG
EPCAM sg8	GCGCAGCATGGCGCCCCCGCAGG
EPCAM sg9	GTTGTGGAGCTGGGCTGGGCTGG
EPCAM sg10	GCGCTTTCCAGCGTGGAGACCGG
EPCAM sg11	GCAGCTTTGCTCGCCTTGGTAGG
EPCAM sg12	GCGCACCCACGTCCTCGGTTCGG
EPCAM sg13	GGGCCCCGGATGTGGGCCTCAGG
LIN28A sg1	CCCGCGGGGGGTTGGGTCATTGT
LIN28A sg2	TTCGAGCTCGGGACTTAGCGGGG
LIN28A sg3	AGTCTCTGGGAAACGCGTTGTGG
LIN28A sg4	CCCGAGACGGCCCTCCGATTCCG
LIN28A sg5	CTGTCCATGACCGCCCGCGCCGG
LIN28A sg6	CCTCCGGCTTACCACGTGTCTAT
MAGEB2 sg1	CCAGGGGTGAATTCTCAGGACTG
MAGEB2 sg2	GGACCCCTAGTGGAGACGAAGGG
MAGEB2 sg3	CTTGGCTTTCACGGGAATCAAGG
MAGEB2 sg4	TCGCCATTGTTAGCACCGAGAGG
MAGEB2 sg5	CCTAAGTGTTACGAGGGGGCTCT
CXCR4 sg1	TGTTTATCACTTGGCGCGTTTGG
CXCR4 sg2	CCACCGACGGTTAAACATCACAA
CXCR4 sg3	CCGTCGGCCGCACAGAGTTTAAC
CXCR4 sg4	ATAAAAACACGCTCCGAGCGCGG
IL6ST sg1	ACTCAGGGCTGCGATAATCGAGG
IL6ST sg2	CGCACGAACCCCTTGGCGCCAGG
IL6ST sg3	CCGCAGAGGTAGCGCGCCACCCC
IL6ST sg4	CCCTACCGGGACCTAGGCGCTGA
IL6ST sg5	CCGTAACGCGGCTCTTCGCGTCC

gRNA	Sequence in the genome
TFRC sg1	CTTTAATCCTCTTATCAACGGGG
TFRC sg2	GAACACCTGCGGCGTTCAGTCGG
TFRC sg3	TCAGGTTGTACCAGAGCCGAAGG
TFRC sg4	CCCGGAGCCCGGGGAACCGAAGG
TFRC sg5	CAGCCCGGAGCCGGCGCACTAGG
TFRC sg6	GGGAAGGGACGAGAGGCGAAGGG
TFRC sg7	CCCACAGGCTCGCGGCCCCGCTC
TFRC sg8	CCACCCGATATCCCGACGCTCTG
TFRC sg9	TGCTCTGACAGATCGCCCGGGGG
TFRC sg10	GAGCCGCGCGATGGGCGCACTGG
HIC1 sg1	GACTCGTCGTGGGTTCCTTTAGG
HIC1 sg2	CGAATCCATGCGCCTGAGTGAGG
HIC1 sg3	CCGTACAGTAATGTCTACGGGGA
HIC1 sg4	GTGCCTAGAAATACGTACATGGG
HIC1 sg5	ATAGCCGCAGCAGCGCCCGCCGG
RASSF1 CGI1 sg1	CCTGGTCCGGTTTGCTGAAGCAA
RASSF1 CGI1 sg2	CCGCGCAATGGAAACCTGGGTGC
RASSF1 CGI1 sg3	GCTTTGGGCGCTAGCAAGCGCGG
RASSF1 CGI1 sg4	CCCAGCGCGCCCAGCGGGTGCCA
RASSF1 CGI1 sg5	CCACTACTCACGCGCGCACTGCA
RASSF1 CG2 sg1	ATTAGAACGCTCCTTGCGCGCGG
RASSF1 CG2 sg2	CCAGCGTCCGGGCAAGCGCACAA
RASSF1 CG2 sg3	CCGGTCGCGTGCGTGCGTGTACG
RASSF1 CG2 sg4	GACGCGGCAACGGACCGGGGAGG
PTEN sg1	CCCTCGGTCTTCCGAGGCGCCCG
PTEN sg2	GCTGCGGCAGGATACGCGCTCGG
PTEN sg3	AGCGCCTGTGAGCAGCCGCGGGG
PTEN sg4	CCAGGGACCCGGGCCGGTTTTAA
p14 sg1	CCCGCCTACCGCCACTTTCCCGC
p14 sg2	CCAAGATCTCGGAACGGCTCTGA
p14 sg3	CTGCCCCCTTAACTGCAGACTGG
p14 sg4	GCACGCGCGCCGAATCCGGAGGG
p16 sg1	CGCCAGAGCCAGCGTTGGCAAGG
p16 sg2	AGGGACCGCGGTATCTTTCCAGG
p16 sg3	CCAGTCAGCCGAAGGCTCCATGC
CDKN2A CGI3 sg	TCCCGGGCAGCGTCGTGCACGGG

Table 5.6: Primers for bisulfite PCR.

Target	Primer	Sequence
EPCAM	bis-EPCAM-1f	GAAGGTTTTTTGTTTGTGTTTGTATT
	bis-EPCAM-1r	AAAAAATAAATAAACTCCCCTCCC
	bis-EPCAM-2f	GAGGGGAGTTTATTTATTTTTTTAATTTT
	bis-EPCAM-2r	ACCTCTTAATCCCCTCCCTATTATAC
	bis-EPCAM-3f	TATAATAGGGAGGGGATTAAGAGGT
	bis-EPCAM-3r	ACTAAAACTATTTTAAAACCCCAAA
	bis-EPCAM-4f	TTTGGGGTTTTAAAATAGTTTTAGT
	bis-EPCAM-4r	CATTACTCATATAATCCTTAAAAAAAA
	bis-EPCAM-new-1f	GTTTTTTGTTTGTGTTTGTATTTTTTTTAG
	bis-EPCAM-new-1r	CRAAACTACTCACCTCTAACC
	bis-EPCAM-new-2f	AGTATTAGYGGTTAGAGGTGAGTAG
	bis-EPCAM-new-2r	CCAACTCCACAACTCTACTCC
	bis-EPCAM-new-3f	GTGGAGTTGGGTTGGGTTGG
	bis-EPCAM-new-3r	ACTAAAACTATTTTAAAACCCCAAATCCACCC
	bis-EPCAM-new-4f	GGGTGGATTTGGGGTTTTAAAATAGTTTTAG
	bis-EPCAM-new-4r	aataaaTAAAAAAAAAAAAAAATACACCTAATTACT
CXCR4	bis-CXCR4-0f	GTGTGTGGGTTTTTTGTTATTTT
	bis-CXCR4-0r	CTCCCCAACACCCTATAAAAC
	bis-CXCR4-1f	GTTTTGTTTTATAGGGTGTTGGG
	bis-CXCR4-1r	ATAATTTAACCTCCCCTTTAACACC
	bis-CXCR4-2f	GGTGTTAAAGGGGAGGTTAAATTAT
	bis-CXCR4-2r	ACCTAAACTAAATTCCCAAAAAAAA
	bis-CXCR4-3f	TTTTTTTGGGAATTTAGTTTAGGTA
	bis-CXCR4-3r	CTACATATATCTCCCCCTTAAATCC
TFRC	bis-TFRC-1f	TGGTAAGATTAGGTTGTATTAGAGT
	bis-TFRC-1r	AAACCAAATATACCCCCACTATC
	bis-TFRC-2f	GATAGTGGGGGTATATTTGGTTT
	bis-TFRC-2r	ATACTATCCAACAACCATAAAAAAC
	bis-TFRC-3f	GTTTTTTATGGTTGTTGGATAGTAT
	bis-TFRC-3r	AAAACTACCAAACTACCAAAATAAA
	bis-TFRC-4f	TTTATTTTGGTAGTTTGGTAGTTTT
	bis-TFRC-4r	CAATAACAAAAATTAAATTTTATTTATTTA
Off-targets	bis-CXCR4-4-off-f	TTTGATTTAGGTGGGATTTTTGGAGTTTG
	bis-CXCR4-4-off-r	TCTAATCACAAACACAACCCAACCC
	bis-CXCR4-2-off-f	TTTATTTTTTAGGGAGAAAAGTGTTTTTG
	bis-CXCR4-2-off-r	ATCCAAAAAAACCTCTTCTCCTAAC
	bis-TFRC-2-off-f	TATTYGTTTTAGGGAAAGAAGGTGAG
	bis-TFRC-2-off-r	AATCAACRACTAAATAACCTCAACC
	bis-TFRC-3-off-f	GTTTTTGTTTTAGTATATTTGTTTTAGGGGTTG
	bis-TFRC-3-off-r	TATTCTAAACTACTACTACCTTTATAAAACCCC
Unrelated CGI	bis-KIA0179-f	TTYGTTTTTTTTTGGTTAAGTTT
	bis-KIA0179-r	RCCCACCTATCTCCCTCTA
	bis-DSCR3-f	TTTTTAAGTTTTAGAGTTTGGGGGT
	bis-DSCR3-r	ATACCTACCCTCCTCAAATATCCTTAA
	bis-SUMO3-f	GGGTTTTGGGGATTTTTGTT
	bis-SUMO3-r	AACCTCCCAATACCTCAATTTCT
	bis-WRB-f	GTAAGTGGTTGTTTGGTTTTTTAA
	bis-WRB-r	AAACACTAAAAACTCCCAATATAAAAT

Target	Primer	Sequence
HIC1	bis-HIC1-1F	GGGTTTTATGGTTATTTTTTTGGAT
	bis-HIC1-1R	ACATAAACTAATTTATCCCCAACCC
	bis-HIC1-2F	GGGTTGGGGATAAATTAGTTTATGT
	bis-HIC1-2R	ATCCCAAAAACCCTAAAACTTTAAA
	bis-HIC1-3F	TTAAAGTTTTAGGGTTTTTGGGATT
	bis-HIC1-3R	AAAACCCACCTAAAAAACATATC
RASSF1A	bis-RASSF1A-1F	TTTTTTATTTTAAAGGTTTTGGAGG
	bis-RASSF1A-1R	AAAAAAACAACCAAAAAACAAC
	bis-RASSF1A-2F	GTTGTTTTTTGGTTGTTTTTTT
	bis-RASSF1A-2R	AAAAAACCTAAACTCATTAAACTAC
	bis-RASSF1A-3F	GTAGTTTAATGAGTTTAGGTTTTTT
	bis-RASSF1A-3R	AAAACCCTCTTCCTCTAACACAATAA
	bis-RASSF1A-4F	TTATTGTGTTAGAGGAAGAGGGTTTT
	bis-RASSF1A-4R	AACCCAAAATAACAAAACCAAATAA
RASSF1B/C	bis-RASSF1B/C-1F	GGTTTTAGTTATAGTTGGATAATGTT
	bis-RASSF1B/C-1R	TCCTACCCCAATAAAACAAAAACTA
	bis-RASSF1B/C-2F	GGTATTTTTTGGTTTTTATTTGGTT
	bis-RASSF1B/C-2R	TCCTACCCCAATAAAACAAAAACTA
	bis-RASSF1B/C-3F	TAGTTTTTGTTTTATTGGGGTAGGA
	bis-RASSF1B/C-3R	AACAAACCACAATACAAACATTCTC
	bis-RASSF1B/C-4F	GTTTGTATTGTGGTTTGTTTATTTT
	bis-RASSF1B/C-4R	CAAACCTCTCTACCTATAACCTTCC
	bis-RASSF1B/C-5F	TGTTATTTAGGTTGGAGTGTAATGG
	bis-RASSF1B/C-5R	AACAAATAACAACCACAAAACAAAC
PTEN	bis-PTEN-1F	TTAGATAGGTGTTTTTTGGGTTTTT
	bis-PTEN-1R	CCCCCAAATCTATATCCTCATAATAT
	bis-PTEN-2F	ATATTATGAGGATATAGATTTGGGGG
	bis-PTEN-2R	CAAACTCTACTAAACATACCCAATATA
	bis-PTEN-3F	TATATTGGGTATGTTTAGTAGAGTTTG
	bis-PTEN-3R	AACTTCCATCATAACTACAACTTCC
	bis-PTEN-4F	GGAAGTTGTAGTTATGATGGAAGTT
	bis-PTEN-4R	ACCCCCTCCCTAAAACTACAAC
	bis-PTEN-5F	GTTGTAGTTTTAGGGAGGGGGT
	bis-PTEN-5R	CTACTTCTCCTCAACAACCAAAAAC
p14	bis-p14-1F	ATTTTTGAGGTGGGTTTAGAAGTTT
	bis-p14-1R	CTAAAAAACCAAATAAAAAAAACCCCT
	bis-p14-2F	GAGGGTTTTTTTTATTTGGTTTTTT
	bis-p14-2R	TCCCAATCTACAATTAAAAAAACAAA
	bis-p14-3F	TTTGTTTTTTTAATTGTAGATTGGGA
	bis-p14-3F	TTTGTTTTTTTAATTGTAGATTGGGA
	bis-p14-4F	ATTTTTATTATTATTTTTTTATTTTTAG
	bis-p14-4R	ACTAACATTCAACCTCCTAATTAAC
p16	bis-p16-1F	TTGATTTTAATTTTTTTGTAAATTT
	bis-p16-1R	AAAAAAAACTAACTAATCACCAAAA
	bis-p16-2F	TTTTGGTGATTAGTTAGTTTTTTTT
	bis-p16-2R	AAACCCAATCCTCCTTCCTTAC
	bis-p16-3F	GGTTTTTTTTTATTTGTTTTTTATAT
	bis-p16-3R	AAAAAAATAATTTTACTTTTTCTTATAATT