

EPIC'RISPR:

A modular and inducible platform for highly parallel synthetic epigenetics and chromatin imaging in a high-throughput format

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

Phil Oberacker

aus Karlsruhe

Hauptberichter:	Dr. Tomasz P. Jurkowski
Mitberichter:	Prof. Dr. Jörn Lausen
Vorsitzender:	Prof. Dr. Stephan Nussberger
Tag der mündlichen Prüfung:	06.07.2021

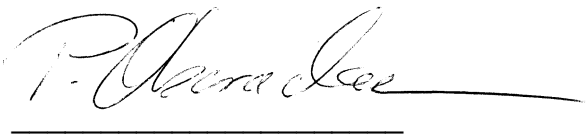
Institut für Biochemie und Technische Biochemie der Universität Stuttgart

2021

Eidesstattliche Erklärung

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

Stuttgart, Juli 2021



Phil Oberacker

Acknowledgements

First, I want to thank my supervisor, Dr Tomasz Jurkowski, for his support and guidance during my time in his group, for many constructive and reviving discussions and for giving me the chance to work on such interesting and exciting projects.

I would also like to express my gratitude to Prof. Dr Jörn Lausen and Prof. Dr Stefan Nussberger for evaluating this thesis and for being co-referees in my committee. Furthermore, I'd like to thank Prof. Dr Monilola Olayioye, Prof. Dr (apl.) Christina Wege, Prof. Dr (apl.) Günter Tovar and Dr Angelika Hausser for agreeing to review my work.

I would like to thank our collaborators Dr Timothy Hore, Dr Ferdinand von Meyenn, Dr Thomas Meany and their groups for their invaluable contributions to our research and the pleasure of working with them. I would also like to thank Dr. Timothy Hore and Dr Renata Jurkowska for the stimulating scientific discussions, their enthusiasm and suggestions.

I want to thank all the members of the IBTB in Stuttgart and the School of Biosciences in Cardiff for their help, support and the great time we had together. I would like to especially thank the other three pancakes Miru, Peter and Sven, but also Dragiza, Elisabeth, Regina, Micha, Nicole, Karl, Dan, Stephi, Angela, Inge and everyone else I had the pleasure working with. I also want to thank the students I mentored throughout my studies. Emily, Sven, Luca and Corinna, it was a pleasure working with you, and I know I learned at least as much from you as you from me. Peter and Sven, another special thanks to the two of you for reviewing this thesis and for all your constructive suggestions.

Furthermore, I would like to thank my friends outside the lab. Dennis, Nina, Michi, Juli, Thomas, Anne, Kerstin, Katharina and Robi, I want to thank all of you for your friendship throughout the years and for some of the best memories of my life.

I also want to thank my family, especially my parents, for their unconditional support throughout each day of my life and for all the sacrifices they made for me. Mama and Papa, without you this thesis would still be a distant dream, without you I would not be the person I am today, and without you, there would be a giant hole in my life. Thank you for everything!

Dear Désirée, last but by far not least I want to thank you for standing with me even though I was so far away for such a long time, for always bringing out the best in me, for loving me, for supporting me and for simply being a part of my life. I love you!

Content

Acknowledgements	V
List of publications	XI
List of abbreviations	XIII
Zusammenfassung	1
Abstract	3
1 Introduction	5
1.1 Epigenetics	5
1.1.1 Defining epigenetics	5
1.1.2 The chromatin	6
1.1.3 Epigenetic modifications and modifiers.....	8
1.2 Synthetic epigenetics	17
1.2.1 Genome targeting	18
1.2.2 Epigenetic editing	21
1.2.3 Inducible epigenetic editing	23
1.2.4 Target genes	26
1.3 Nucleic acid manipulation in a high-throughput format	29
1.4 Aims of the thesis	30
2 Results	35
2.1 Bio-On-Magnetic-Beads (BOMB)	35
2.1.1 MNP synthesis and coating	37
2.1.2 Clean-up and size exclusion.....	38
2.1.3 Plasmid isolation	41
2.1.4 RNA isolation from mammalian cells	44
2.2 Development of the EPIC'RISPR platform	45
2.2.1 The gRNA toolbox.....	45
2.2.2 The effector toolbox.....	48
2.2.3 Proof of concept and optimisation	50
2.3 EPIC'RISPR for highly parallel transcription modulation and chromatin labelling	55
2.3.1 gRNA mixtures	56
2.3.2 one effector / multiple targets	59
2.3.3 multiple effectors / multiple targets	60

2.3.4	multiple effector-copies / one target.....	62
2.3.5	multiple effectors / one target.....	62
2.4	EPIC'RISPR for inducible and tuneable transcription modulation	66
2.4.1	Switch-ON.....	66
2.4.2	Switch-OFF.....	69
2.5	EPIC'RISPR for large-scale parallel modulation of transcription.....	74
2.5.1	Controls.....	74
2.5.2	HKMTs.....	75
2.5.3	HKDMs.....	80
2.5.4	HATs and HDACs	85
2.6	CsyTag for transgene expression control.....	87
3	Discussion	95
3.1	EPIC'RISPR.....	95
3.1.1	Utilizing a modular toolbox	95
3.1.2	Parallel modulation of transcription and chromatin imaging.....	96
3.1.3	Inducibility and tuneability of transcription	103
3.1.4	Probing the epigenetic code.....	111
3.1.5	Outlook.....	116
3.1.6	Conclusion	120
3.2	CsyTag for transgene expression control.....	121
3.2.1	Scientific context	121
3.2.2	Outlook.....	123
3.2.3	Conclusion	124
3.3	Bio-On-Magnetic-Beads (BOMB).....	124
3.3.1	Benefits of BOMB.....	124
3.3.2	Outlook.....	126
3.3.3	Conclusion	128
4	Material and Methods.....	129
4.1	Plasmids	129
4.1.1	Cloning.....	129
4.1.2	gRNAs and EDs	130
4.1.3	Csy4 plasmids.....	130
4.1.4	CsyTag-modified plasmids	131
4.1.5	PuroR.....	131
4.2	Cell culture.....	132
4.2.1	HEK293 culture conditions	132

4.2.2	Transfections.....	133
4.3	Downstream processing and analysis.....	135
4.3.1	DNA	135
4.3.2	RNA	136
4.3.3	RT-qPCR	137
4.3.4	FACS analysis.....	139
4.3.5	LSM	139
4.3.6	Total protein extraction.....	139
4.3.7	SDS-PAGE and Western blot.....	140
4.3.8	<i>in vitro</i> transcription of gRNA.....	140
	Bibliography	143
	Supplementary Data	167

List of publications

Oberacker P*, Stepper P*, Bond DM*, Hoehn S, Focken J, Meyer V, et al. Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation. PLoS Biol. 2019;17. doi:10.1371/journal.pbio.3000107

Oberacker P*, Stepper P*, Bond D, Hipp K, Hore TA, Jurkowski TP. Simple Synthesis of Functionalized Paramagnetic Beads for Nucleic Acid Purification and Manipulation. Bio-Protocol. 2019;9: 1–10. doi:10.21769/bioprotoc.3394

Walker KT, Donora M, Thomas A, Philips AJ, Ramgoolam K, Pilch KS, **Oberacker P**, et al. CONTAIN: An open-source shipping container laboratory optimised for automated COVID-19 diagnostics. bioRxiv. 2020, 05.20.106625: 1-29. doi: 10.1101/2020.05.20.106625

Oberacker P, Stepper P, Jurkowski TP. CsyTag: An inducible, multipurpose RNA-tag for tuneable control of gene expression. 2021; submitted.

* shared first-authorship

List of abbreviations

abbreviation	definition
2xC-Gd2	pCAG-2xCsyTag-GFPd2
3xC-Gd2	pCAG-3xCsyTag-GFPd2
4xC-Gd2	pCAG-4xCsyTag-GFPd2
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
a	adenine
A	alanine
AB	aptamer binder
ac	acetylation
ARID5B	AT-rich interactive domain 5B
<i>ASCL1</i>	Achaete-scute family bHLH transcription factor 1
BB	binding buffer
BF	bright field
BFP	blue fluorescent protein
BOMB	Bio-On-Magnetic-Beads
boxB	stem-loop structure of prophage HK022
bp	base pairs
BrD	bromodomain
c	cytosine
CaCl ₂	calcium chloride
Cas	CRISPR-associated
CD	catalytic domain
cDNA	complementary DNA
CEM	chemical epigenetic modifiers
C-Gd2	pCAG-CsyTag-GFPd2
C-Gd2-C	pCAG-CsyTag-GFPd2-CsyTag
ChIP	chromatin immunoprecipitation
CIB	cryptochrome 2-cryptochrome-interacting bHLH
CoFe ₂ O ₄	cobalt ferrite

CpG	cytosine-guanine dinucleotide
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
Cry2	cryptochrome 2
Csy4	CRISPR-associated endonuclease Cas6f
CsyTag	Csy4 recognition site as a degradation tag
CuAAC	copper(I)-catalyzed azide-alkyne cycloaddition
CXCR4	C-X-C motif chemokine receptor 4
Cys ₂ -His ₂	C2H2
D	aspartate
D10A	protein mutant, aspartate at position 10 exchanged for alanine
d2	degradation domain
dCas9	dead Cas9
ddH ₂ O	double-distilled water
DMEM	Dulbecco's Modified Eagle's Medium
DNMT1/3A/3B/3L	DNA-methyltransferase 1/3A/3B/3L
DOT1L	DOT1 like
DPBS	Dulbecco's Phosphate Buffered Saline
dpt	days post transfection
<i>E. coli</i>	<i>Escherichia coli</i>
ED	effector domain
EED	Embryonic Ectoderm Development
EGFP	enhanced green fluorescent protein
<i>EPCAM</i>	epithelial cell adhesion molecule
ESC	embryonic stem cells
EtOH	ethanol
EZH2	enhancer of Zeste [E(z)] homolog 2
FACS	fluorescence-activated cell sorting
FBS	fetal bovine serum
FBXL10	F-Box And Leucine-Rich Repeat Protein 10
Fe ₃ O ₄	magnetic iron(II,III) oxide
FeCl ₂	iron(II) chloride
FeCl ₃	iron(III) chloride
FKBP	FK506-binding protein

FRB	FKBP-rapamycin-binding protein
G	glycine
g	guanine
Gd2	pCAG-GFPd2
Gd2-C	pCAG-GFPd2-CsyTag
gDNA	genomic DNA
GFP	green fluorescent protein
GITC	guanidine thiocyanate
GNAT	Gcn5-related N-acetyltransferases
gRNA	guide RNA
GSS	glycine-serine-serine
Gu-HCl	guanidine hydrochloride
H	histidine
H29A/G	protein mutant, histidine at position 29 exchanged for alanine or glycine
H2A	histone 2A
H2B	histone 2B
H3	histone 3
H3K14ac	acetylation at lysine 14 of histone 3
H3K27ac	acetylation at lysine 27 of histone 3
H3K27me1/2/3	mono-, di-, tri-methylation at lysine 27 of histone 3
H3K36ac	acetylation at lysine 36 of histone 3
H3K36me1/2/3	mono-, di-, tri-methylation at lysine 36 of histone 3
H3K4me1/2/3	mono-, di-, tri-methylation at lysine 4 of histone 3
H3K9ac	acetylation at lysine 9 of histone 3
H3K9me1/2/3	mono-, di-, tri-methylation at lysine 9 of histone 3
H4	histone 4
H4K20me1/2/3	mono-, di-, tri-methylation at lysine 20 of histone 4
H840A	protein mutant, histidine at position 840 exchanged for alanine
HAT	histone acetyltransferase
HDAC	histone deacetylases
HEK293	Human embryonic kidney 293
HKDM	histone-lysine demethylase
HKMT	histone-lysine methyltransferase
HOX	homeobox

HP1	Heterochromatin protein 1
hpt	hours post-transfection
JMJD	Jumonji Domain
KRAB	Krüppel-associated box
LB	Luria-Bertani Broth
<i>LIN28A</i>	Lin-28 homolog A
LM	light microscopy
LSM	laser scanning microscopy
MAA	methacrylic acid
MCP	MS2 coat protein
me	methylated
me1/2/3	mono-, di- or trimethylation
MgCl ₂	magnesium chloride
MLL	Mixed Lineage Leukemia
MNP	magnetic nanoparticle
mRNA	messenger RNA
MS2	RNA stem-loop structure of the bacteriophage MS2
MYST	MOZ, Ybf2, Sas2, and Tip60
N22p	Nun protein of prophage HK022, binds boxB
NaCl	sodium chloride
NAD ⁺	nicotinamide adenine dinucleotide
NaOH	sodium hydroxide
ncRNA	non-coding RNA
NLS	nuclear localisation sequence
NSD1/2	nuclear receptor SET domain-containing 1/2
nt	nucleotides
NURF	Nucleosome Remodeling Factor
OD600	optical density at 600 nm
ORF	open reading frame
PADI	peptidylarginine deiminases
PAM	protospacer adjacent motif
PBS	PUF-binding site
PcG	Polycomb Group
PCP	PP7 coat protein

PHD finger	plant homeodomain finger
PhyB	phytochrome B
PIF	phytochrome-interacting factor
PMAA	polymethacrylic acid
PP7	RNA stem-loop structure of <i>Pseudomonas aeruginosa</i> PP7
PRC1/2	polycomb repressive complex 1/2
PRDM2	PR/SET domain 2
PRMT	protein-arginine methyltransferase
PS	polystyrene
PTM	post-translational modification
PUF	Pumilio/FBF
<i>REEP5</i>	receptor accessory protein 5
RFP	red fluorescent protein
RIPA	Radioimmunoprecipitation assay
RT	room temperature
RT-qPCR	real-time quantitative polymerase chain reaction
SB	Super Broth
SEM	standard error of the mean or scanning electron microscopy
SET	Su(var)3-9, Enhancer-of-zeste and Trithorax
SiO ₂	silicon dioxide
SMASh	small molecule–assisted shutoff
SOB	Super Optimal Broth
SPRI	solid-phase reversible immobilisation
SUV39H1/2	Suppressor Of Variegation 3-9 Homolog 1/2
t	thymine
TAE	Tris-Acetate-EDTA
TALE	transcription activating-like effectors
TB	Terrific Broth
TE	Tris-EDTA
TEM	transmission electron microscopy
TEOS	tetraethyl orthosilicate
TET	Ten-eleven translocation methylcytosine dioxygenases
TFR1	transferrin receptor 1
<i>TFRC</i>	gene coding for transferrin receptor 1

TNA	total nucleic acid
TPE	Tris-phosphate-EDTA
tracrRNA	trans-activating CRISPR RNA
tRNA	transfer-RNA
TSS	transcriptional start site
U	units
UTR	untranslated regions
V	volumes
VP16	herpes simplex based transcriptional activator
VP64	4 consecutive copies of the activation factor VP16
VPR	synthetic construct of VP64, p65 and Rta
vSET	viral SET domain methyltransferase
w/o	without
WT	wildtype
Xi	X chromosomes inactivation
ZYM505	Zymol Medium 505
ΔNLS	removed nuclear localisation sequence

Zusammenfassung

Das Epigenom beschreibt die Gesamtheit der epigenetischen Zustände in einem Organismus. Es besteht aus den biochemischen Modifikationen der DNA und der Histon-Proteine, nicht-kodierenden RNAs und der dreidimensionalen Architektur des Erbguts. Diese Modifikationen und Strukturen steuern die Expression des Genoms in zelltyp-spezifischen Mustern und kontrollieren daher die Entwicklung des gesamten Organismus. Die diesbezügliche Forschung erbrachte hauptsächlich beschreibende Informationen über die Korrelation von epigenetischen Modifikationen und Genexpression. Leider ist bis heute wenig über die kausalen Zusammenhänge im epigenetischen Netzwerk bekannt. Mit der Entdeckung der bahnbrechenden CRISPR/Cas9 Technologie ist es heute allerdings möglich, das epigenetische Programm gezielt zu stören. Diese Methode, welche auch als Epigenetische Editierung bekannt ist, erlaubt es Effektormoleküle gezielt an spezifische Gene zu dirigieren, wo sie Modifikationen setzen oder entfernen können. Indem anschließend die Reaktion des Epigenoms beobachtet wird, können Rückschlüsse auf die Funktionsweise des epigenetischen Netzwerkes gezogen werden. Dieses System ist jedoch recht limitiert in Bezug auf die simultane Modifikation mehrerer Zielgene, was jedoch eine Notwendigkeit bei der detaillierten Untersuchung von Netzwerken darstellt.

In dieser Arbeit kombinierte ich die beiden Funktionen der Zielbestimmung sowie der Rekrutierung auf einem Molekül des CRISPR/Cas9 Systems, der gRNA. Auf diese Weise ist die EPIC'RISPR Plattform in der Lage mehrere Effektormoleküle zeitgleich an ein oder mehrere Ziele zu rekrutieren, ohne dass diese miteinander interferieren. Ich wies dies nach, indem ich zum einen zeitgleich drei Zielgene aktivierte oder reprimierte und zum anderen indem ich mehrere Fluorophore an verschiedene Orte im Erbgut dirigierte. Ich verwendete diese Technologie außerdem dazu fünf verschieden stark exprimierte Gene simultan mit einem Effektormolekül zu beeinflussen. Dafür führte ich ein Experiment durch, in welchem ich die Auswirkungen von mehr als 60 epigenetischen Effektormolekülen auf die Expression der Zielgene untersuchte. Dabei identifizierte ich mehrere vielversprechende Kandidaten, welche potenziell synergistische und damit stärkere und langanhaltendere Wirkung auf das epigenetische Programm haben könnten. Des Weiteren entwickelte ich EIN- und AUS-Schalter für das EPIC'RISPR System, welche kleine Moleküle verwenden um die eingeführten Effekte in ihrer Stärke beliebig zu steuern. Der AUS-Schalter kann

außerdem auch für die Expressionskontrolle von Transgenen verwendet werden, was die Funktionalität des Systems noch weiter ausdehnt.

Zusätzlich entwickelte unsere Arbeitsgruppe Protokolle für die Synthese und Funktionalisierung von paramagnetischen *beads*, sowie für deren Anwendung zur automatisierten Extraktion von Nukleinsäuren im Hoch-Durchsatz-Verfahren. Seitdem unsere Plattform, welche wir Bio-On-Magnetic-Beads (BOMB) nennen, veröffentlicht wurde, ist sie zu einem Dreh- und Angelpunkt für quelloffene Wissenschaft geworden, insbesondere während der COVID-19 Pandemie.

Abstract

The epigenome describes the sum of epigenetic states in an organism. It consists of biochemical modifications of the DNA and histone proteins, non-coding RNAs and the three-dimensional architecture of the genome. These modifications and structures regulate the genome expression in a cell-type-specific pattern and hence control the development of the whole organism. Research in this field yielded a lot of descriptive information about the correlation between epigenetic marks and gene expression. Unfortunately, we do not know much about the causalities within the epigenetic network. With the discovery of the groundbreaking CRISPR/Cas9 technology, it is now possible to interfere with the epigenetic program. This methodology, which is known as epigenetic editing, allows the recruitment of effector molecules to distinct targets where they introduce or remove specific modifications. By observing the response of the epigenome, we can conclude how the epigenetic network functions. However, this system is somewhat limited regarding the simultaneous modification of multiple loci, which is a necessity for investigating a network.

In this thesis, I combined the targeting and recruiting functionality of the CRISPR/Cas9 system in one molecule, the gRNA. Like this, this EPIC'RISPR platform can recruit numerous effector molecules to one or multiple targets simultaneously without interference. I demonstrated this by activating and repressing three target genes with different effector domains at once and by recruiting different fluorophores to several target loci. I further applied this technology to perturb five differently expressed target genes simultaneously with one effector molecule at a time. For this, I performed a large-scale experiment in which I probed the effects of more than 60 epigenetic effector molecules on target gene transcription. I identified several promising candidates which might exhibit synergistic behaviour and hence a stronger and longer-lasting impact on the epigenetic program. Furthermore, I developed ON- and OFF-switches for the EPIC'RISPR system which utilize small molecules to fine-tune the introduced effects arbitrarily. The OFF-switch was further applied for transgene expression control, extending the functionality of this system even further.

Additionally, our group developed protocols for the synthesis and functionalisation of paramagnetic beads and their application in the automated high-throughput extraction of nucleic acids. Since its publication, our platform, which we call Bio-On-Magnetic-

Beads (BOMB) has since become a hub for collaborations in open-source science, especially during the COVID-19 pandemic.

1 Introduction

1.1 Epigenetics

1.1.1 Defining epigenetics

The first mention of the term epigenetics appeared in 1957 [1]. In his work, Conrad Waddington described the process of cell differentiation metaphorically as a marble on the top of a hill (**Figure 1-1A**). The hillside, the epigenetic landscape, is laced with forking valleys. These valleys represent the intermediate stages of differentiation, which a cell has to pass to achieve the transition from pluripotency to a fully differentiated cell type. In this representation, the epigenetic landscape is defined and formed by the genes which act as pegs that pull the slopes of the valley and determine the respective pathways (**Figure 1-1B**).

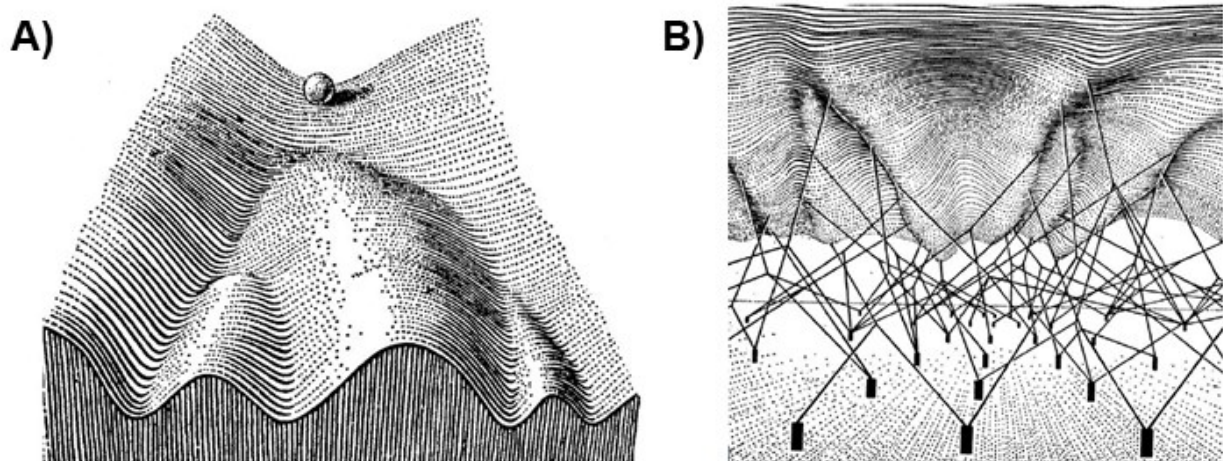


Figure 1-1: The epigenetic landscape. Figures were taken from Waddington [1]. **A)** Marble on the top of a hill represents a pluripotent cell rolling down the valleys to become a fully differentiated cell type. **B)** The genes act as pegs which form the shape of the valley and determine the fate of the rolling marble.

Since Waddington established this metaphor, our knowledge and understanding of what epigenetics is and how it functions on a molecular level have drastically improved. Today, we know that epigenetics can be interpreted as a regulatory network that works above genetic regulation. Epigenetic mechanisms offer the possibility to form genetic switches that are independent of the primary genetic code, such as e.g. the lac-operon [2]. They allow for a far greater biochemical complexity within the organism and contribute to the evolution of complex beings [3]. In 2009, the scientific community eventually settled on the following definition for the term epigenetics: “An epigenetic

trait is a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence” [4]. Whether the term “heritable” only describes the dissemination of traits from a mother to a daughter cell or also transgenerational inheritance from one organism to another is still widely debated [5–7].

Altering the phenotype of a gene in a heritable way, without introducing genetic mutations, can only be achieved by changing the respective expression level in a lasting way. A first layer of expression regulation is the condensation of DNA by proteins called histones. These proteins structure the DNA into subunits called nucleosomes that play a crucial role in the steric accessibility of the nucleic acid. By adding various chemical modifications to the amino-terminal regions of the histone proteins, their binding affinity to DNA can alter. This change in affinity can either create or close openings for the transcriptional machinery, depending on the respective modification [8,9]. Another layer of regulation represents the chemical modification of DNA with methyl-groups which can hinder DNA-binding proteins from detecting their target sequence and thus altering their effects on transcription [10,11]. Besides chromatin structure, histone modifications and DNA methylation, long non-coding RNAs (ncRNA) have been shown to influence gene expression in various ways [12].

1.1.2 The chromatin

To fully grasp how epigenetic modifications can affect the transcriptional state of a gene, it is paramount to understand the chromatin organisation (**Figure 1-2**). To compress the human DNA (~ 2 m long) into a single nucleus of a few μm , it is wrapped around an octameric histone core [13]. These octamers contain two copies of each of the histone proteins (H2A, H2B, H3, H4) and typically bind 146-147 base pairs (bp) of DNA in 1.67 helical turns. The resulting subunit is called a nucleosome, which can then be further condensed into chromatin and eventually the chromosome [14]. Histones, as described here, have only been reported in eukaryotic nuclei and some archaea, but not in bacteria [15].

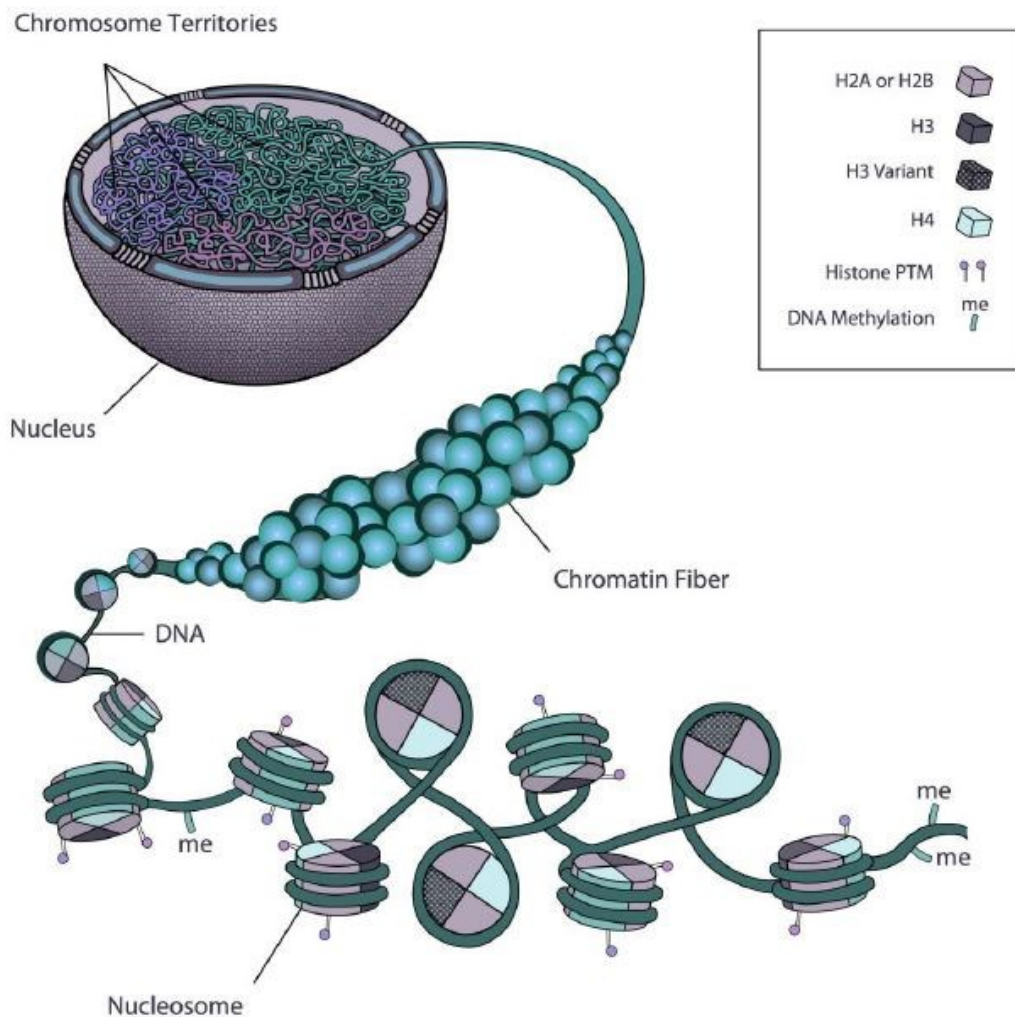


Figure 1-2: The chromatin structure. The figure was taken from Rosa and Shaw [16]. Two copies of each histone protein (H2A, H2B, H3, H4) form the core of the nucleosome, which is enwrapped in a DNA strand of 146-147 bp. DNA methylation (me) and post-translational modifications (PTM) of the histone tails contribute to the epigenetic programming. The nucleosome can be arranged in a 30 nm chromatin fibre and fitted into the nucleus.

For the transcription machinery to reach the DNA and promote gene expression, the chromatin must be accessible. The chromatin already assumes a regulatory role by arranging the DNA in a defined order. However, chemical modifications can change the binding dynamics between the DNA and the histone cores and hence alter the density and conformation of the chromatin. Depending on the respective modifications (see chapter 1.1.3 for details) the chromatin can either form densely packed areas, the heterochromatin, which is virtually inaccessible for the transcription machinery or arrange in openly approachable “beads on a string”, the euchromatin, which can be easily contacted by transcription factors and polymerases [17]. The heterochromatin typically contains silenced genes and is positioned in the laminar areas of the nucleus.

At the same time, euchromatin is associated with actively transcribed genes, and it is located in the core-centre. However, the chromatin structure, as well as the applied histone modifications, can be altered dynamically, which means that the expression of genes can be changed consistently by the cell, usually by recruiting huge remodelling complexes of several proteins [18]. This enables complex regulatory dynamics between chromatin structure and gene expression within the cell.

1.1.3 Epigenetic modifications and modifiers

Histone modifications

The epigenetic marks most relevant for this thesis are the chemical modifications of the amino-terminal histone tails. These marks are placed post translation on various amino acids of the histone, and so far, an extensive catalogue of different modifications has been reported. Prominent examples are methylation, acetylation, phosphorylation, and ubiquitination in the histones H3 and H4, with the first two being studied the most. In the majority of cases, the precise function of the histone marks is not known yet. Most studies in the recent past focussed on single modifications though, determining the exact position for a histone mark in an active, bivalent or repressed gene as well as in regulatory element [19]. This associative data represents the basis of how we typically classify an epigenetic mark today. It is assumed, however, that they underlie a histone code, where a distinct combination of marks determines a specific meaning, and consequently a precise expression level of the gene [20].

To crack such a complex code, it is paramount to understand how it is set, read, and erased. Up to today, several hundred proteins have been identified which either interact with histone modifications directly or with proteins that do, representing a whole new network of regulatory possibilities [21–27]. I summarized a selection of prominent epigenetic players as well as their targets and functions in **Figure 1-3** for a better overview. In the following paragraphs, I will describe the most important histone modifications step by step, explain their association with gene expression and what their typical interaction partners are.

H3K4

Modifications of the lysine residue at position four of the H3 tail (H3K4) are strongly associated with actively transcribed genes. H3K4 methylation is by far the most studied epigenetic mark and can occur in three states: mono-, di- and trimethyl [28,29].

Especially trimethylated H3K4 (H3K4me3) is strongly associated with active genes and histone acetylation. It usually locates at the 5' regions of genes where it flanks the transcriptional start site (TSS). However, it has also been observed in large clusters covering multiple genes in the context of developmental homeobox (HOX) genes [30,31]. The positioning of dimethylated H3K4 (H3K4me2) differs depending on species and cell type. In vertebrates, it usually correlates with H3K4me3. However, there are also reports of H3K4me2 sites that are free of trimethylation and not located at a TSS, suggesting the modification might be involved in lineage-specific chromatin states rather than general genomic regulation [29]. Monomethylated H3K4 (H3K4me1) on the other hand, is highly enriched in the gene body of actively transcribed genes as well as in enhancers, where it strongly correlates with acetylated lysine 27 of H3 (H3K27ac) [32,33].

H3K4 methylation is a modification introduced by histone-lysine methyltransferases (HKMTs), of which the majority contains a catalytic Su(var)3-9, Enhancer-of-zeste and Trithorax (SET) domain. The first HKMT identified to specifically methylate H3K4 was SET1 [37]. Today we know multiple other H3K4-specific HKMTs, including the Mixed Lineage Leukemia (MLL) family, which contains five known and one predicted HKMT: MLL1-4 (KMT2A-2D), SETD1A and SETD1B (predicted) [34]. MLL-family HKMTs are part of multiprotein complexes and share a common core complex of WDR5, RbPB5 and ASH2. This complex guides MLL1 to its target histones, by WDR5 specifically binding to H3K4me2. The SET domain then introduces the third methyl group, leaving H3K4me3 as the final product [35].

Several protein domains can specifically detect and differentiate H3K4 methylation. The Nucleosome Remodeling Factor (NURF) complex e.g. is recruited to open heterochromatin by sequence-specific transcription factors (e.g. GAGA). It contains a BPTF subunit which specifically recognizes H3K4me3 via an aromatic cage in its plant homeodomain (PHD) finger motif. Simultaneously, the subunit detects lysine acetylation with a bromodomain (BrD) [29], representing a good example for inter-modification cross-talk. In this context, H3K4me3 serves as a stabilizing factor that allows the NURF complex to increase the accessibility of the target site and recruit further factors that facilitate transcription [36].

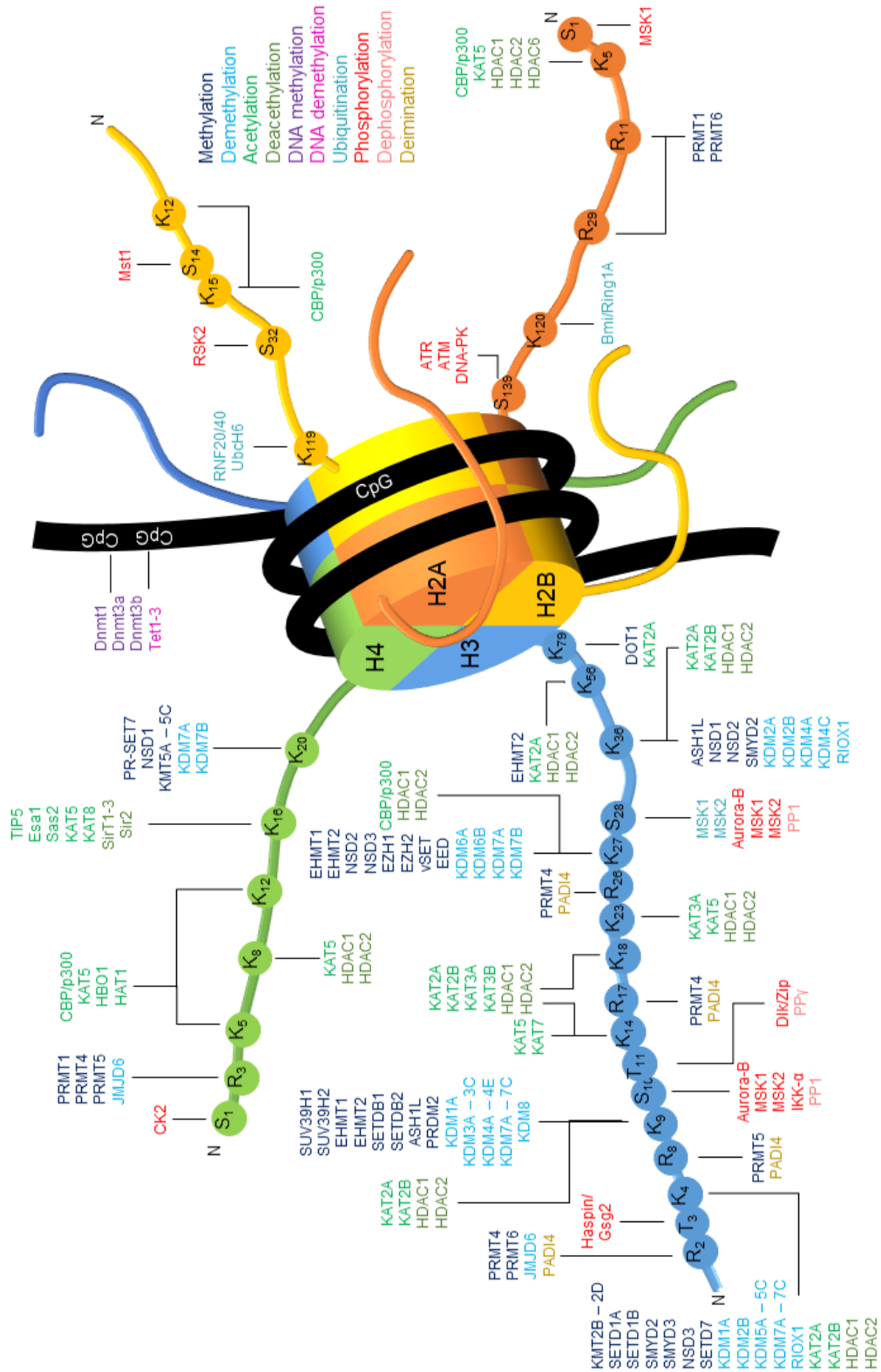


Figure 1-3: Post-translational nucleosome modifications. An overview of the most important epigenetic editors and their respective targets in the histone and DNA. Colours indicate the respective functionality. Data were obtained from several sources [37–39].

Furthermore, H3K4me3 can be bound by Ying1 which, in association with H3K36me3-binding Eaf3, can recruit the NuA4 complex, a well-known recruiter for histone acetyltransferases (HAT) such as HTATIP (KAT5) [40]. In this example, H3K4me3 actively promotes transcriptional activity in collaboration with another mark. In contrast, H3K4me3 has also been shown to recruit histone deacetylases (HDAC) in the case of DNA damage by offering an anchor point to the PHD-domain in ING2 [41]. Although this represents a direct link between H3K4me3 and transcriptional repression, it makes sense that this mark is used as a recruiting point for DNA damage repair mechanisms, as it is usually present in actively transcribed genes where the need for DNA repair is more immediate. Another example for H3K4 methylation specificity constitutes the ADD domain of the *de novo* DNA-methyltransferase 3A (DNMT3A), which is only able to bind unmethylated H3K4, but not H3K4me3 [42]. This way, H3K4me3 directly prevents active promoters from DNA methylation and stabilizes transcriptional activity.

To erase the H3K4 methylation marks, one of two classes of histone lysine demethylases (HKDM) is required that exhibit different specificities for distinct methylation states. LSD1 (KDM1A) is the first reported HKDM, and it specifically demethylates H3K4me2, but not H3K4me3 [43]. The Jumonji/AT-Rich-Interactive Domain 1 (JARID1/KDM5)-proteins, on the other hand, target H3K4me3. It has been reported that both proteins can act synergistically in global demethylation, but also that KDM5 might prevent the silencing of expressed genes by LSD1 [44].

H3K9

H3K9 can participate in both transcriptional repression and activation. The acetylated version of H3K9 (H3K9ac) is typically located in active promoters and co-localizes with the histone marks H3K4me3 and H3K14ac [45]. Mono-methylated H3K9 (H3K9me1) also correlates with active genes, as it is typically enriched around the TSS of active genes, while di- and trimethylated H3K9 (H3K9me2/3) are usually found in silenced genes [28]. Interestingly, the same study also demonstrates that some active genes contain H3K9me2 and me3 in their promoters.

The enzymes responsible for methylation of H3K9 are mostly KMTs of the Suppressor Of Variegation 3-9 (SUV39) family. Suppressor Of Variegation 3-9 Homolog 1 (SUV39H1) and Homolog 2 (SUV39H2) produce H3K9me3 by using H3K9me1 or me2 as a substrate [46]. Heterochromatin protein 1 (HP1) can bind H3K9me3 via its chromodomain and simultaneously facilitate the organization of constitutive

heterochromatin by oligomerization utilizing its chromoshadow-domain [47]. HP1 also interacts with SUV39H1 and H2, thus spreading and maintaining heterochromatin at the same time. The SUV39H1/HP1 complex has further been shown to interact with the Methyl-CpG Binding Domain 1 directly linking DNA methylation to H3K9me3 [48,49].

While H3K9me3 is a characteristic mark for heterochromatin, H3K9me2 plays an important role in the silencing of euchromatic regions and X chromosomes inactivation (Xi) [50,51], where it interacts with Cdy1, a protein that further interacts with H3K27me3 and recruits another member of the SUV39 family, G9a (EHMT2). G9a forms heterodimers with GLP (EHMT1) and facilitates the epigenetic inheritance of euchromatic H3K9me2 [52]. Triggered by binding H3K9me2 with its ANK domain, it can generate H3K9me2 on neighbouring histones and thus propagate the epigenetic modification to newly assembled nucleosomes. Furthermore, recent studies suggest that the G9a/GLP complex is directly involved in genetic imprinting by recruiting de novo DNA methyltransferases [53].

There are multiple HKDMs, mostly members of the JmjC domain-containing family, that show specificity towards H3K9 methylation. Most of them even differentiate between different methylation levels, suggesting a tight and distinct control of H3K9 methylation is crucial for cellular homeostasis. KDM3A, KDM3B and KDM7B e.g., specifically demethylate H3K9me1 and H3K9me2, but not H3K9me3 [54–56]. LSD1, which has also been known to demethylate H3K4me2, shows an even stronger distinction and only exhibits demethylation activity towards H3K9me2 [57]. H3K9me3, on the other hand, can be demethylated by KDM4A and KDM4C [58,59].

H3K27 and the Polycomb Group (PcG)

H3K27 is mostly known for its involvement in transcriptional repression. H3K27 is typically trimethylated (H3K27me3) by the enhancer of Zeste [E(z)] homolog 2 (EZH2) protein, a subunit of the polycomb repressive complex 2 (PRC2). However, other HKMTs have been reported to exhibit (weaker) H3K27 methylation activity as well [60,61]. In *Drosophila*, PRC2 is recruited at specific DNA sites called polycomb responsive elements [62], while in mammalian organisms it seems to interact with the transcription factor REST which binds to cytosine-guanine dinucleotide (CpG)-rich sequences [63]. Once recruited, PRC2 trimethylates adjacent H3K27 residues.

H3K27me3 then serves as a docking site for the polycomb repressive complex 1 (PRC1) which directly interferes with transcription [64].

H3K27me3 plays a crucial role in promoters of developmentally important genes of embryonic stem cells (ESC), where it co-exists with the antagonistic H3K4me3 mark in the same nucleosome [65]. These bivalent or “poised” promoters can either be permanently silenced (H3K9me3, DNA methylation) or activated during cell differentiation [66]. H3K27me1 and H3K27me2 are less studied than the trimethylated state of the residue but show a similar distribution. Although both are controlled by the PRC2 complex as well, H3K27me1 is associated with transcriptionally active promoters [67], suggesting multiple counter-intuitive responsibilities of the complex.

H3K27me2 and me3 demethylation are mainly carried out by UTX (KDM6A), a member of the JmjC family of HKDMs and a subunit of the MLL2/3 complex. UTX stimulates the expression of HOX genes by removing the repressive H3K27 methylation mark and simultaneously facilitates H3K4me3 by its association with MLL2/3 [68].

Besides its role in PcG-mediated repression, H3K27 can also participate in transcriptional activation and therefore, serve as an antagonist of the repression complex. If acetylated (e.g. by p300/CBP (KAT3B)), H3K27ac blocks methylation as one lysine cannot be modified both ways simultaneously and additionally facilitates an open chromatin structure by the addition of a negative charge [69].

H3K36

The function of H3K36 methylation is very diverse and differs depending on its location in the gene and the degree of methylation, highlighting the extreme complexity and importance of the mark. There is a gradual distribution of mono-, di- and trimethylated H3K36 (H3K36me1/2/3) from the promoter region of active genes to the 3'-end [70]. The methylation typically serves a repressive purpose as it prevents an aberrant start of transcription within the gene body.

H3K36me1 and me2 are typically introduced by nuclear receptor SET domain-containing 1 (NSD1/KMT3B) and NSD2 (KMT3G) directly after start of the transcription, quickly followed by trimethylation with enzymes like SETD2, NSD3 (KMT3F), ASH1L (KMT2H) or SMYD2 (KMT3C) [71]. H3K36me3 then recruits Rpd3S Histone deacetylase complex which leads to histone deacetylation and chromatin condensation, consequently prohibiting transcription within the gene body [72,73].

There are also reports that H3K36me3 is involved in epigenetically encoded, cell-type-specific splicing of exons, although the exact mechanism is not yet completely understood [74]. Furthermore, H3K36me3 is strongly associated with DNA methylation [75].

Tri- and dimethylated H3K36 can be demethylated by KDM2B and KDM4A. As the latter is known to associate with HP1, H3K36 demethylation is directly linked to heterochromatin formation [76,77]. H3K36me1 can further be removed by FXBL11 (KDM2A), which cannot demethylate H3K36me3 [78].

Not many studies focused on H3K36 acetylation (H3K36ac) yet. However, the mark shows a similar distribution as other H3 acetylations and is involved in transcriptional activation [79].

H4K20

While most lysines on H4 are typically acetylated, H4K20 can be mono-, di- and trimethylated (H4K20me_{1/2/3}). Introduced by KMT5A (SET8/PR-SET7), H4K20me₁ is mostly associated with actively transcribed genes and plays an important role in cell cycle control, where it is involved in chromatin modification [80]. The function of H4K20me₂ is more diverse, as it seems to be involved in DNA repair, as well as in DNA replication, where it serves as a starting-marker [81,82]. H4K20me₃, on the other hand, participates in the silencing of repetitive DNA, such as transposons and can, therefore, be directly associated with the repression of transcriptionally active genes [80,83]. Both H4K20me₂ and me₃ are introduced by KMT5B and 5C (SUV420H1/H2) [84].

H2A and H2B

While histones 3 and 4 are the subject of intensive studies, histones 2A and 2B have been mostly overlooked in the past. Although several modifications, ranging from methylation, acetylation, ubiquitination, and phosphorylation have been reported, little is known about their function and distribution [85]. It has been shown, though, that monoubiquitinated H2A plays a role in the repression of HOX genes, where it is introduced by the PRC1 complex [86].

Arginine methylation and deimination

Besides the modification of lysine, arginine is another common target for PTM in the amino-terminal regions in histones. Methylation of arginine residues is mostly associated with transcriptional activation and is typically catalyzed by proteins of the protein-arginine methyltransferases (PRMTs). However, the methylation of arginine can be antagonised by deimination. The protein family of peptidylarginine deiminases (PADs) can convert unmethylated arginine to citrulline. PADs, hence, do not function as active arginine demethylases, but rather prevent methylation. So far, no enzyme has been reported that can reverse the deimination of arginine [87,88].

Histone acetylation and deacetylation

So far, I mostly discussed histone methylation and the critical enzymes for placing and erasing these marks, although histone acetylation is a significant modification of the histone proteins as well. It is a vital marker for the dynamic and precise control of transcription and is positively associated with H3K4me3 [89]. By shrouding the positive charge of the lysine residues, acetylation hinders the interaction between histones and DNA, consequently opening up the chromatin structure, which allows BrD-containing regulators and transcription factors to access the DNA and transcriptionally activate the associated gene [90,91]. In contrast to most HKMTs and HKDMs, HATs and HDACs are usually not specific for a single lysine modification but rather exhibit a broader catalytic potential.

There are three prominent families of HATs relevant for this thesis, namely Gcn5-related N-acetyltransferases (GNAT)/PCAF [92], the MOZ, Ybf2, Sas2, and Tip60 (MYST)-family [93], and the p300/CBP (CREB-binding protein) HATs [94]. They interact with other proteins to form large acetylation complexes and hence contain multiple reading domains like BrDs, PHD-fingers or tudor domains [95]. Although following different catalytic strategies, all of them utilize Acetyl-CoA as a donor for acetyl groups.

HDAC can be distinguished in four classes (I, II, III and IV) [96]. Class I contains HDAC1-3 and HDAC8, which all associate in complexes. HDAC 1 and 2 have even been reported to be part of several complexes, namely NuRD, CoResr and mSin3A, which can lead to both stable and transient repression of transcription [97]. Class II (HDAC4-7, 9a-c and 11) share the same catalytic mechanisms with class I and directly

hydrolyse acetyl-lysine [96]. HDACs of class III, also known as sirtuins, utilize a completely different catalytic mechanism which depends on nicotinamide adenine dinucleotide (NAD⁺) as a co-factor [96]. The only known class IV member is HDAC11. The protein shares a similar sequence with HDACs from class I and II and is involved in the regulation of protein stability [98].

DNA methylation and demethylation

Besides the modification of histone-tails and remodelling of the chromatin, methylation of DNA represents a crucial element in epigenetic programming and cell development. In mammals, the methylation of CpG dinucleotides is the most commonly found modification of DNA [99]. The methyl group is hereby added to the 5th carbon atom in the pyrimidine ring forming 5-methylcytosine (5mC). The modified base can be the target of spontaneous deamination, which turns it into thymine (T), creating a T:C mismatch in the DNA sequence. This mismatch can then be detected and repaired by the DNA repair machinery [100]. In somatic cells, about 70-80% of the CpG sites are methylated, except for CpG-islands. These are CpG-rich regions that occur in about 70% of all human genes, and approximately 50% of all CpG sites are usually found in the promoter regions [101,102]. CpG-methylation of the promoter region of a gene results in transcriptional repression [103], while methylation in the gene body is mainly found in actively transcribed genes, where it correlates with H3K36me3 to prevent aberrant transcription [104].

In humans, three enzymes of the DNA methyltransferase (DNMT) family are responsible for introducing DNA methylation and its maintenance: DNMT1, DNMT3A and DNMT3B. While the former is responsible for detecting hemimethylated DNA (e.g. after replication) and reintroduce the palindromic 5mC, DNMT3A and 3B interact with the catalytically inactive DNMT3L to introduce methyl groups *de novo* in unmethylated CpG sites [105]. Before the discovery of the Ten-eleven translocation methylcytosine dioxygenases (TET), it was suggested that demethylation only occurs by dilution as a consequence of DNA replication. It has since been shown that 5mC can be oxidized in a step-wise manner to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and eventually to 5-carboxylcytosine (5caC), which can then be exchanged to unmethylated C by the base excision repair mechanism or thymine DNA glycosylases [106].

VPR and KRAB: Transcriptional modulators

Although the primary purpose of the epigenetic machinery lies in the regulation of the cell's genome, it also functions as a repressor for the spread of transposable elements and repetitive sequences as well as viral infections. It comes to no surprise that most viruses have developed counteracting elements themselves, that drive their transcriptional activity. The mature virions of herpes simplex virus type 1 e.g. contain a factor termed VP16 that exhibits *cis*- and *trans*-activating functionality [107]. Multiple copies of this element can be synthetically combined (VP64) to achieve even stronger effects. VP64 has since been a major effector for transcriptional synthetic activation experiments [108,109]. In another example, the Epstein-Barr virus encodes for two immediate-early proteins: Rta and ZEBRA. In the initiation of the lytic cycle, these two proteins enhance their own expression in a positive feedback loop, but also stimulate each other's expression as well as the expression of other viral components downstream [110]. Together with the NF- κ B subunit p65 [111], VP64 and Rta have been combined to the tripartite VP64-p65-Rta (VPR), one of the strongest known synthetic transactivators [109].

The Krüppel-associated box (KRAB) can be found in about 30% of the all known zinc finger proteins in humans and has long been identified as a robust transcriptional repression domain [108,112,113]. Located at the N-terminus of Cys₂-His₂ (C2H2) zinc finger proteins, KRAB interacts with chromatin remodelling factors (e.g. NuRD complex), eventually recruiting HKMTs such as SETDB1 and effectively shuts down gene transcription via H3K9me3 introduction and deacetylation [114]. Therefore, KRAB has been a primary choice for synthetic transcriptional repression in multiple studies [108,115,116].

1.2 Synthetic epigenetics

So far, studies of the epigenetic network have mostly provided associative and descriptive data, which has been compiled in massive libraries. This assembled knowledge was vital for identifying alterations in the epigenome during the development of diseases and for discovering epigenetic drugs to treat them [117–123]. Although these approaches have proven their worth, they are unable to unravel major questions about the epigenetic system: What is the exact functionality of an epigenetic mark? Is it the cause for a change in the transcription status, or is it introduced as a

consequence? How do epigenetic marks interact with the epigenetic protein machinery? Synthetic epigenetics provides multiple tools that could help to answer these questions. The term synthetic epigenetics was defined by Jurkowski et al. as “the design and construction of novel specific artificial epigenetic pathways or the redesign of existing natural biological systems, to intentionally change epigenetic information of the cell at desired loci” [124]. Following this definition, there are two primary strategies for applying synthetic epigenetics: indirect top-down and direct bottom-up. Indirect top-down approaches include somatic cell nuclear transfer experiments, transdifferentiation and induced pluripotency in ESC [125–128]. In these experiments, the differentiation is initiated by oocyte-specific factors or expression of transcription factors, meaning the occurring epigenetic changes are a byproduct of the conversion rather than the actual cause. The bottom-up approach, on the other hand, utilizes targeting domains that specifically bind an arbitrary DNA sequence and recruit an epigenetic effector domain (ED) that introduces a specific epigenetic mark at the desired target and hence perturbs the epigenetic code at this location. This set-up, which usually creates a single, defined change in the epigenetic network, is called epigenetic editing [129] and it is the main focus of this thesis.

1.2.1 Genome targeting

The idea of epigenetic editing requires the possibility to recruit an ED to an arbitrary but distinct genetic locus. Although there are hundreds of proteins known that contain DNA interacting domains, most of them relate to a specific sequence, and there is no simple way of changing their target. The first candidates that allow the modulation of their binding specificity were C2H2 zinc fingers, 30 amino acid long units which specifically bind three bp of DNA [130]. This specificity can be altered by changing the amino acids that interact with the nucleobases, which consequently results in new zinc finger unit variants with distinct specificities. Multiple of these units (typically 3 to 6, **Figure 1-4A**) can be arranged in arrays of tandem repeats to cover a specific sequence long enough to be unique in the genome [131,132]. However, neighbouring zinc fingers, as well as base modifications, influence the binding specificity, creating the necessity of validating each array separately, which is a rather time-consuming process [133].

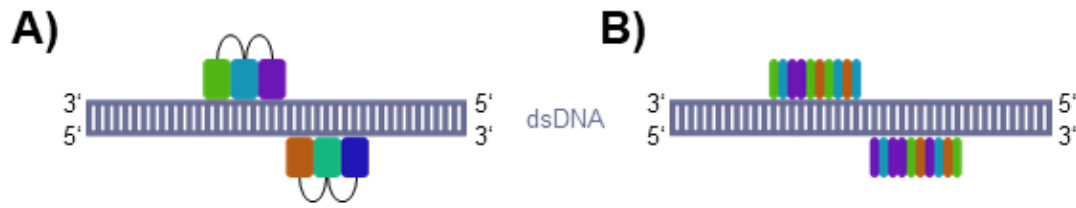


Figure 1-4: Schematic representation of genome targeting with **A)** zinc finger- and **B)** TALE-arrays. While zinc finger units detect three consecutive bp, TALE units are specific for a single nucleotide.

The first alternatives to zinc fingers are the transcription activating-like effectors (TALE) [134]. TALE occur in tandem repeats of 34 amino acids, with each repeat recognizing a single, distinct bp utilizing the amino acids in position 12 and 13 of each repeat [135]. This single readout allows an elegant design of custom TALE arrays that can be easily assembled in a modular fashion (**Figure 1-4B**) [136,137]. However, TALE suffer from the same time-consuming engineering process as zinc fingers, and each array has to be assembled separately.

Although zinc finger and TALE arrays have contributed to many important discoveries, they were outmatched by the versatility, flexibility and modularity of the CRISPR/Cas9 system. The clustered regularly interspaced short palindromic repeats (CRISPR) have been known for many years before their role as an adaptive prokaryotic immune system was revealed [138]. In this context, exogenous DNA elements (e.g. from phages or plasmids) are processed into short fragments (spacers) and inserted into the bacterial genome between the repetitive CRISPR sequences. The CRISPR array is then transcribed into a precursor RNA which is further processed into the CRISPR RNAs (crRNA) which each contains an individual spacer sequence. The crRNA guides CRISPR-associated (Cas) proteins to complementary invading genetic elements, which it recognizes by Watson-Crick base pairing. After the target sequence is identified, the Cas protein exhibits an endonucleolytic activity, effectively cleaving the intruder sequence. To be recognized by the complex though, the target sequence requires a so-called protospacer adjacent motif (PAM) at its 3'-end. This PAM is usually three nucleotides long and specific for each Cas protein [139]. Three distinct types of CRISPR systems have been described so far, from which type II is the most utilized in genome engineering and synthetic epigenetics, as it only requires one protein, Cas9.

The type II CRISPR/Cas system also entails a trans-activating CRISPR RNA (tracrRNA) for the maturing of the crRNA from the precursor. The tracrRNA binds to

the crRNA precursor, and both molecules are co-processed by a double-strand specific RNA endoribonuclease RNase III [140,141]. The mature crRNA:tracrRNA duplex then associates with Cas9 and guides the endoribonuclease to a complementary double-stranded DNA where it introduces a site-specific double-strand break, using its catalytic domains (HNH and RuvC) [142]. The crRNA:tracrRNA duplex can also be fused with a small stem-loop into a single guide RNA (gRNA) which reduces the CRISPR/Cas9 system to only two components, making it even more powerful and easy to apply [139,142,143]. A schematic representation of CRISPR/Cas9 is shown in **Figure 1-5**.

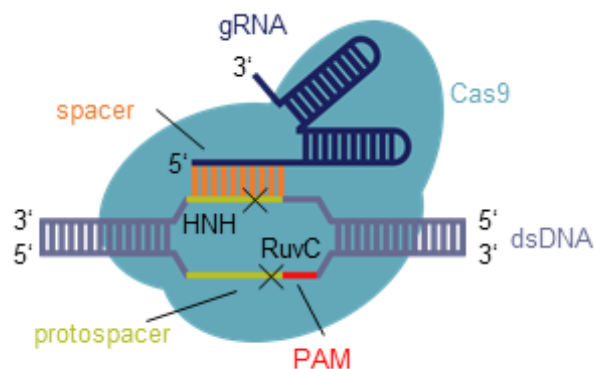


Figure 1-5: Schematic representation of the Cas9:gRNA targeting complex. The recognition of the target, as well as the insertion of a double-strand break (HNH, RuvC), requires complementarity of spacer and protospacer and an appropriate PAM site. In the case of Cas9, the PAM sequence is NGG, with N being any base.

The distinct advantage of the CRISPR/Cas system over zinc fingers and TALE is the way of how it recognizes its target sequence. Watson-Crick base pairing allows the simple exchange of the crRNA or gRNA with the desired complementary spacer sequence to change the target. Additionally, orthologous Cas9 species utilize different PAMs which enable the recognition of multiple targets simultaneously [144]. Zinc fingers and TALE arrays, on the other hand, recognize their binding sequence in the DNA by interacting with the major groove [145,146]. As described in chapter 1.1.3, DNA is methylated in the majority of the cases at CpG and 5mC, 5fC and 5caC are located in the major groove [147]. Hence, they may interfere with the target recognition of zinc fingers and TALE [148,149].

CRISPR/Cas9 became especially interesting for synthetic epigenetics in 2013 when a catalytically dead Cas9 (dCas9) was developed [150]. A single point mutation in each

of the catalytic centres (HNH_H840A, RuvC_D10A) is sufficient to remove all catalytic activity while maintaining the ability to recognize and bind a specific sequence. This allows the recruitment of any protein to any target locus in the genome, simply by fusing it to dCas9. This discovery cleared the way for epigenetic editing as we know it today.

1.2.2 Epigenetic editing

Epigenetic editing represents one of the most powerful tools to study the epigenetic network. By combining an ED with a sequence-specific targeting domain, an epigenetic alteration can be introduced at almost any location in the genome (**Figure 1-6**) [129,151]. The consequences of this alteration can be analysed by quantifying the expression level of the target gene [152], measuring its methylation status [153,154] and determining the presence of chromatin marks [155].

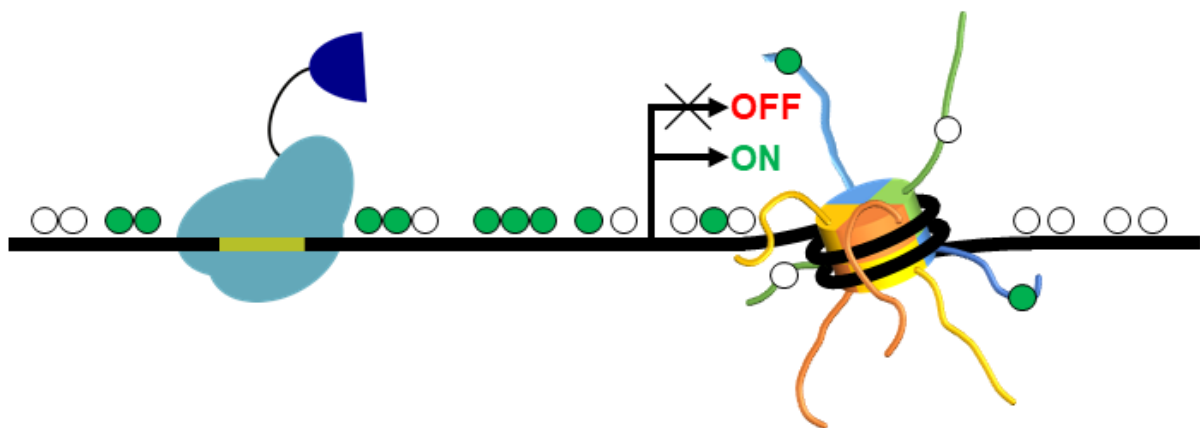


Figure 1-6: Concept of epigenetic editing. Figure inspired by Jurkowski et al. [124]. The fusion construct of DNA binding domain and epigenetic modifier is recruited to the desired, reprogrammable DNA sequence, where the modifier can alter the epigenetic state and consequently introduce a change in expression level (activation or repression). Green dots represent epigenetic modifications on the DNA or histone tails.

As the choice for practical targeting domains was very limited in the past, the development of new epigenetic editing tools initially progressed very slowly. Despite their limitations, however, zinc fingers and TALE arrays participated in pioneering research. During this time, the scientific community put a strong focus on DNA methylation and consequently, the primary candidates for epigenetic editing were DNMTs [156–159] and TET enzymes [160,161]. With the discovery of the dCas9 mutant [150], it became very soon apparent that the CRISPR/Cas system would be the new major player in epigenetic editing [162]. Shortly after, the first study utilizing

CRISPR/Cas for direct epigenetic editing with a dCas9-p300 (KAT3B) fusion construct was published [163], soon followed by a great variety of research papers demonstrating epigenetic editing by recruiting DNMTs [103,164]) and TET enzymes [165], HKMTs [166] and HKDMs [167], fusion constructs of multiple effectors [103,168,169], synthetic transcriptional modulators [109] and even multiple copies of one effector [170]. However, there are contradictory results about the stability of the introduced marks. While the group around Kungulovski et al. demonstrated that synthetically introduced methylation of DNA and H3K9 deteriorates within a few days after transfection [171], two other groups show stable propagation of methylation over a period of several weeks [168,172].

The real strength of CRISPR/Cas as an epigenetic editing tool lies in its modularity, especially when it is combined with aptamers. An aptamer is a small RNA molecule that forms a secondary structure and can be specifically bound by a ligand [173]. The most notable example for such a ligand or aptamer binder (AB) is the bacteriophage MS2 coat protein (MCP) which binds to a small stem-loop structure, referred to as MS2 [174]. MS2 can either be used to replace the loop-structures in a gRNA [175], or it can be added at its 3' end [176]. Both constructs allow the system to recruit MCP-tagged EDs to a gene of interest (**Figure 1-7**). Besides MS2, there are several other aptamer/AB combinations from which PP7/PCP (RNA stem-loop structure of *Pseudomonas aeruginosa* PP7 / PP7 coat protein) [177] and boxB/N22p (stem-loop structure of prophage HK022 / Nun protein of prophage HK022) [178] are the most important representatives for this thesis. Alternatively, conserved Pumilio/FBF (PUF) RNA-binding domains can specifically interact with certain 8-mer RNA sequences (PUF-binding site, PBS), which allows ED-recruitment in the same manner as described for MS2/MCP [179]. As the use of RNA structures for recruitment combines this functionality with the targeting specificity on one molecule, the system gains a vast boost in modularity. Different EDs can potentially be recruited to various but distinct targets, simply by exchanging the used gRNAs. It has even been shown for fluorophores that different RNA stem-loops in one gRNA allow the co-recruitment of different molecules to one target locus [180]. This functionality has a strong potential for studies focussing on the interaction of different epigenetic marks and/or effector molecules. It will provide knowledge about their interaction and eventually allow controlling the epigenetic program to an unprecedented extent.

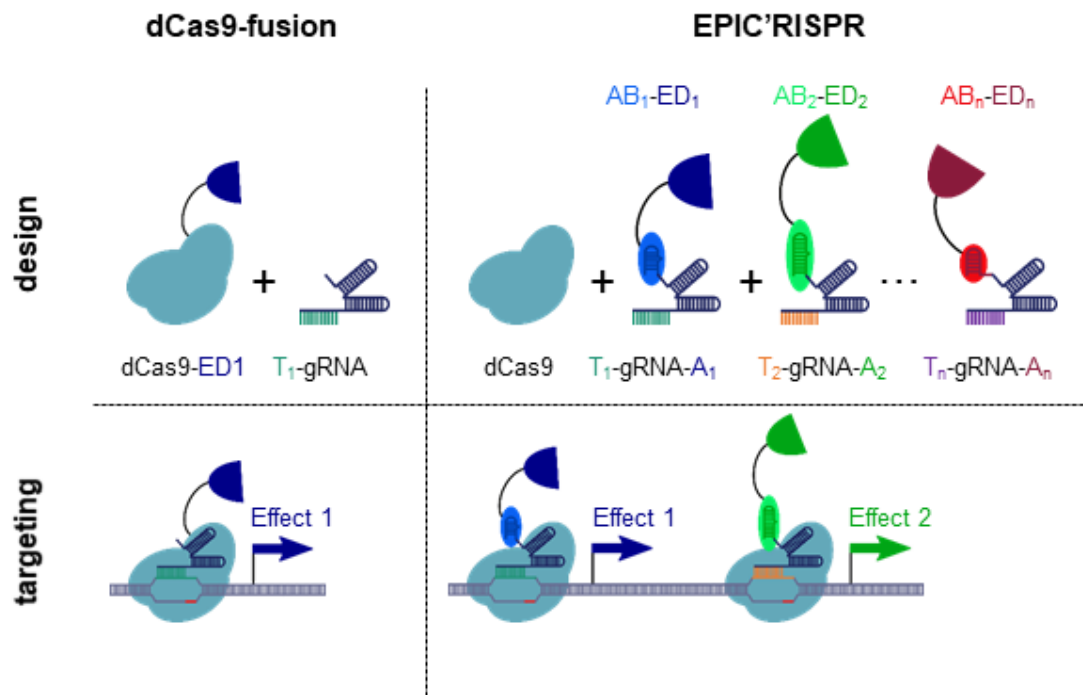


Figure 1-7: Epigenetic editing strategies with dCas. The classical system uses a dCas9-ED fusion construct. The exhibited function can be recruited to one or many target (T) sites. The EPIC'RISPR system utilized aptamer (A) RNA scaffolds in the gRNA to link targeting and recruitment specificity. The system, therefore, can modify individual targets with different regulatory effects simultaneously.

1.2.3 Inducible epigenetic editing

Maintaining homeostasis but simultaneously adapting to changes in the environment requires cellular programming to be flexible and dynamic. Studying the causality of gene functions in biochemical processes hence requires the spatio-temporal modulation of gene expression, a functionality which the classical tools do not contain. A variety of elegant solutions for adding this functionality can be provided by the light-sensitive proteins found in many plants and microbes [181–188]. It did not take long before these optogenetic modules were integrated into the existing binding domain- [189–191], zinc finger- [192], TALE- [193,194] and CRISPR/Cas-based [195] editing systems to create two-hybrid systems that assemble when exposed to a specific wavelength. In these set-ups, one of the binding partners is commonly fused to the recruiting module, while the other is linked to the desired ED. Upon induction, the ED is recruited to the location of interest. Alternatively, systems in which two modules hybridize in the presence of specific, non-toxic chemicals can be used. The respective set-ups allow the initiation of epigenetic modulation by merely adding the individual compound to the cell culture media [196–200].

FKBP and FRB

One of the most thoroughly characterized systems for drug-induced dimerization employs the dimerization domains of the FK506-binding protein (FKBP) and the FKBP-rapamycin-binding protein (FRB), termed DmrA and DmrC. The two domains form heterodimers by interacting with the chemical rapamycin, or one of its analogues (**Figure 1-8A**) [201].

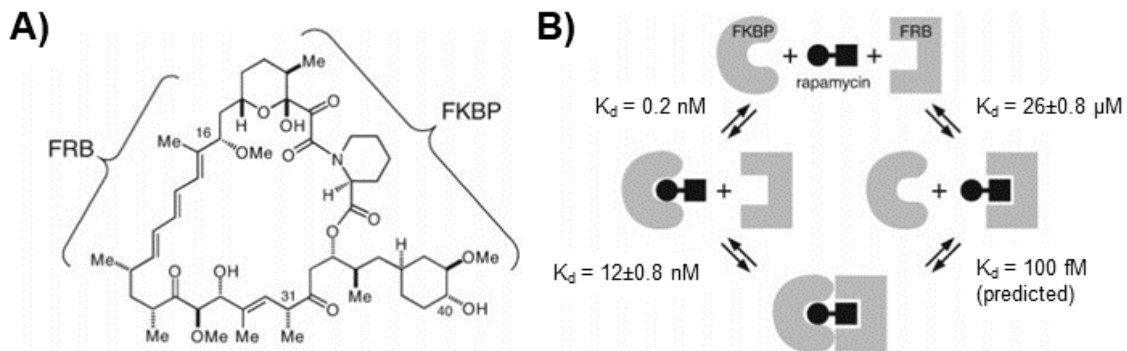


Figure 1-8: FRB and FKBP dimerize in the presence of rapamycin. These figures were taken and slightly modified from Banaszynski et al. [201] **A)** Chemical structure of rapamycin and the interaction areas with the two binding domains. **B)** Schematic of the possible binding states of FRB, FKBP and rapamycin.

The binding strength of FRB and FKBP to rapamycin varies depending on the respective binding state (**Figure 1-8B**). FKBP shows a strong individual binding affinity towards rapamycin ($K_d = 0.2 \text{ nM}$), while the interaction between the chemical and FRB is only modest in the absence of FKBP ($K_d = 26 \pm 0.8 \text{ }\mu\text{M}$). However, this interaction is about 2000-fold stronger when FKBP is already bound to rapamycin ($K_d = 12 \pm 0.8 \text{ nM}$) [201]. The reason for this is a conformational change in FKBP when binding to rapamycin. This change creates a binding site for FRB and allows direct interaction between the two proteins, significantly increasing the binding probability of FRB to the complex [202,203]. Due to the great detail we know about FRB and FKBP, these heterodimers have been prime candidates for adding the functionality of inducibility to epigenetic editing systems [197–200].

Csy4 as a potential OFF-switch

All of the described dimerization complexes assemble within seconds when triggered but typically need hours to reverse the binding [181,183,188,200]. Light-inducible systems are quicker and usually separate within 6 to 10 minutes [184,188]. Some of

them are even faster and disassemble in a few seconds, but require the exposure to a particular wavelength to do so [188,190,191]. However, certain experiments (e.g. investigating repair mechanisms) require a system with ON as the status *quo*, which can be turned off within minutes without the need for elaborate set-ups. A potential candidate for filling this role as an OFF-switch for epigenetic editing is another member of the CRISPR/Cas family, the CRISPR-associated endonuclease Cas6f (Csy4) [204]. Its native function is the processing of the initial crRNA transcript into mature crRNA molecules. For this, it binds and digests a small stem-loop in the CRISPR repeats, the Csy4 recognition site. This tag has recently been implemented as an RNA degradation tag for the control of transgenes where it is implemented in the untranslated regions (UTR) of the reporter gene and allows the Csy4-induced cleavage of the respective UTR [205,206]. This not only prohibits the recruitment of the translation machinery [207] but also enables the digest of the messenger RNA (mRNA) by exonucleases [208]. The catalytic activity of Csy4 depends strongly on the histidine (H) in its active centre (**Figure 1-9A**). Similar as described for Cas9 above, a mutant in which the H is exchanged for alanine (A) or glycine (G), loses all catalytic activity while maintaining the specific binding to its target sequence [209]. It has been shown *in vitro*, however, that imidazole can rescue this activity by assuming the role of the imidazole ring of the lost H (**Figure 1-9B**) and quickly degrade the respective RNA [210]. If this rescue effect can be achieved *in vivo*, Csy4 and its binding site would be a valuable tool for induced degradation of gRNAs and therefore a potential fast OFF-switch for CRISPR/Cas-based epigenetic editing.

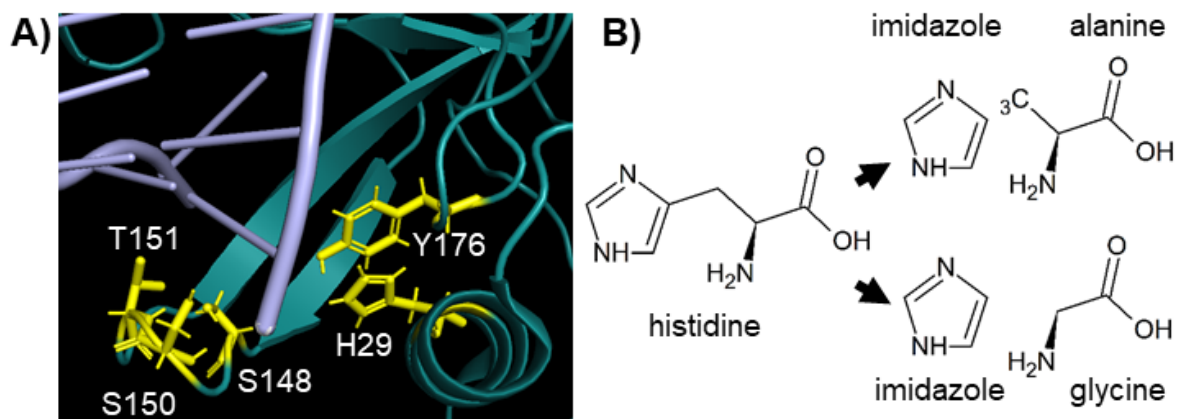


Figure 1-9: Csy4 and imidazole. The figure was taken from Oberacker et al. ([211] submitted) **A**) Crystal structure of the active centre of Csy4 binding to CsyTag RNA stem-loop (grey) with the essential amino acids highlighted in yellow. The figure is based on RCSB PDB 4AL5 [212]. **B**) Schematic representation of the substitution of H by A or G in combination with imidazole.

1.2.4 Target genes

There is a great variety of possible target genes for developing epigenetic editing tools. For the editing experiments in this thesis, I chose a selection of five genes with different expression levels and different epigenetic environments. All genes have been shown to be valid targets for synthetic activation and repression by recruiting epigenetic effectors to their promoters. The spacer sequences for all genes are listed in **Table 0-1** of the Supplementary Data section and have been designed, verified and kindly provided by my colleague Peter Stepper.

CXCR4 (Figure 1-10A, annotated in orange)

C-X-C motif chemokine receptor 4 (*CXCR4*) is a gene coding for a transmembrane receptor, and its overexpression is associated with several cancer types [213–215]. Furthermore, it is one of the prime requirements for HIV entry into the cell [216]. The gene has been successfully activated [176] and repressed [116] using CRISPR-based editing tools. However, sample preparation must be executed quite carefully, as *CXCR4* is a monoexonic gene and remaining genomic DNA (gDNA) cannot be distinguished from complementary DNA (cDNA).

EPCAM (Figure 1-10B, annotated in yellow)

The gene for the oncogene epithelial cell adhesion molecule (*EPCAM*) encodes a transmembrane protein that plays a role in several biochemical pathways, such as cell cycle control, several signalling processes as well as in proliferation [217,218]. *EPCAM* is upregulated in multiple cancers, and it has been the target of epigenetic editing studies [219,220].

TFRC (Figure 1-10C, annotated in red)

TFRC or *CD71* is the gene that encodes for the transferrin receptor 1 (TFR1) protein. TFR1 is typically located in the cell membrane where it is required for the uptake of iron from transferrin and its delivery into the cell [221]. As *CXCR4* and *EPCAM*, *TFRC* is frequently upregulated in cancer [222,223], including acute leukaemia [224], and it was one of the first targets for CRISPR-based repression with dCas9-KRAB fusion proteins [116].

ASCL1 (Figure 1-11A, annotated in green)

Achaete-scute family bHLH transcription factor 1 (*ASCL1*) is the gene of a helix-loop-helix transcription factor and regulates cell differentiation in neurons [225]. It is almost completely repressed in Human embryonic kidney 293 (HEK293) cells, making it a target perfectly suited for transcriptional activation [175,226]. However, this also means that repression experiments on *ASCL1* need to be evaluated very carefully, as even small mistakes in handling or concentration calculation might lead to huge errors.

LIN28A (Figure 1-11B, annotated in purple)

Lin-28 homolog A (*LIN28A*) encodes a miRNA- and mRNA-binding protein and is mostly expressed stem cells [227]. Like its homolog *LIN28B*, it is overexpressed in a variety of severe cancer types [228] and has been a target for synthetic gene activation [175].

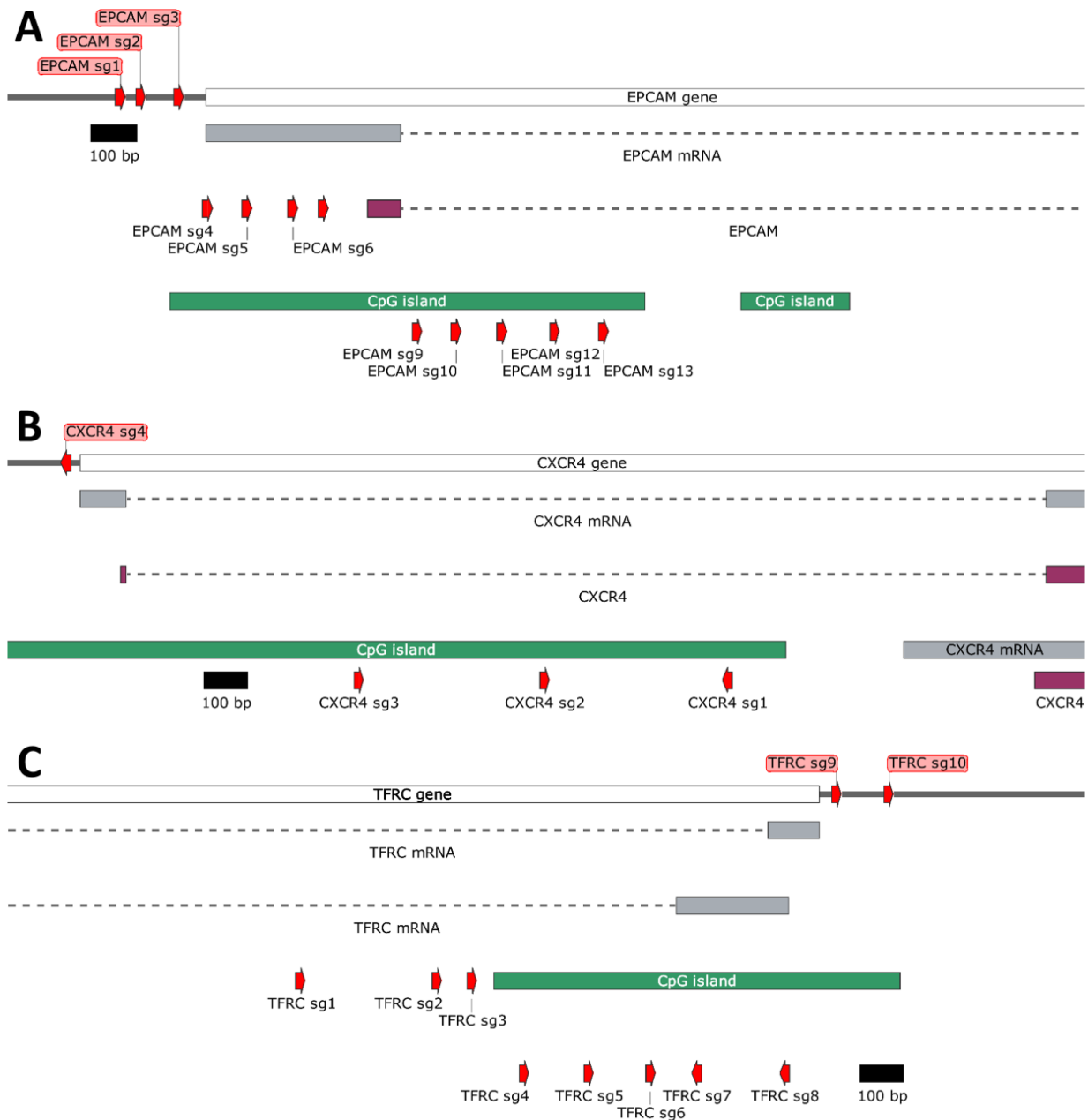


Figure 1-10: Target gene promoter regions 1. These figures were made and kindly provided by Peter Stepper. The genes are indicated as white boxes, with exons depicted in grey and introns as dashed lines. Purple bars represent protein-coding regions, and green bars indicate CpG islands. The gRNA target sequences are shown as red arrows that indicate the orientation of the PAM. **A)** *EPCAM* on chromosome 2. **B)** *CXCR4* on chromosome 2. **C)** *TFRC* on chromosome 3.

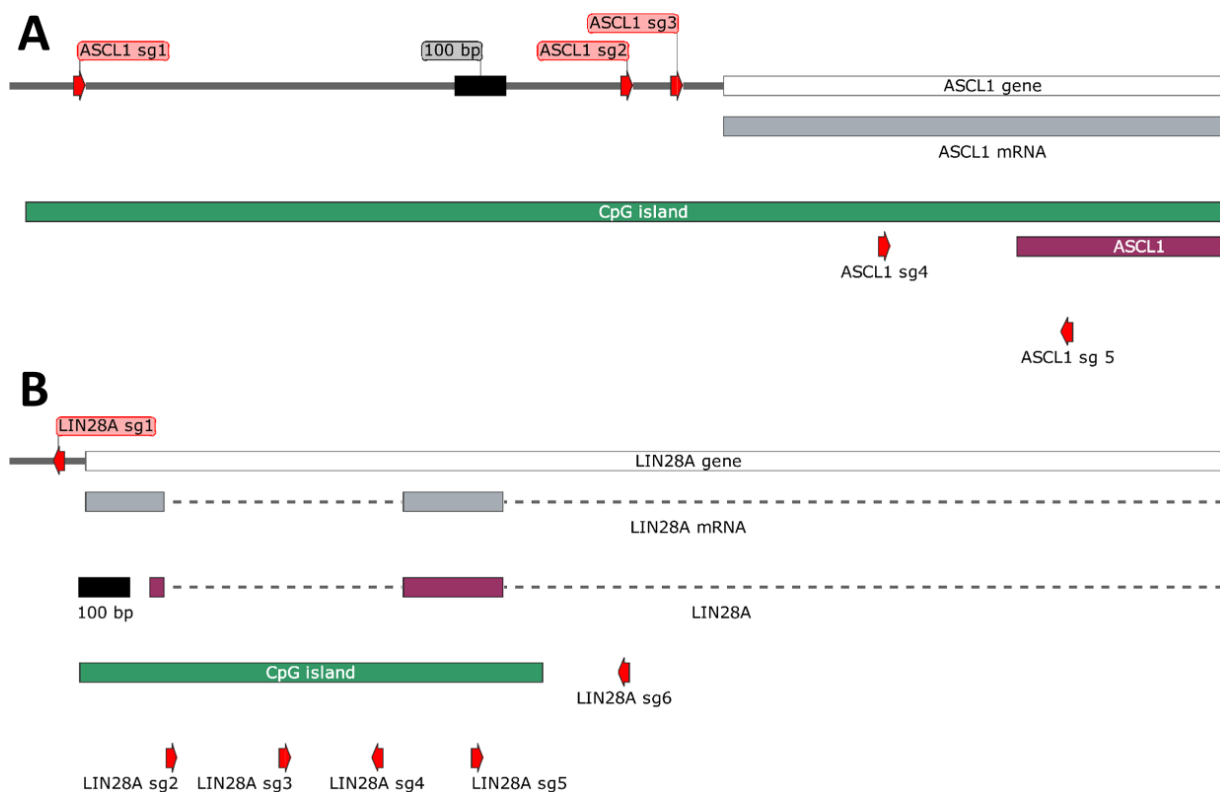


Figure 1-11: Target gene promoter regions 2. These figures were made and kindly provided by Peter Stepper. The genes are indicated as white boxes, with exons depicted in grey and introns as dashed lines. Purple bars represent protein-coding regions, and green bars indicate CpG islands. The gRNA target sequences are shown as red arrows that indicate the orientation of the PAM. **A)** *ASCL1* on chromosome 12. **B)** *LIN28A* on chromosome 1.

1.3 Nucleic acid manipulation in a high-throughput format

The vast majority of medical or research institutions that focus on molecular biology, genetics or epigenetics rely on the purification or manipulation of nucleic acids. DNA and RNA can be isolated from cultured bacterial, plant or animal cells, but also from tissue samples, clinical swaps, plasma, enzymatic reactions or even environmental sources. Most commercial purification or extraction kits utilize silica-based column systems. However, column-based protocols usually require a centrifuge, limiting their applications in high-throughput set-ups. Single-tube systems can be scaled according to the used table-top centrifuge, which typically only holds between 24 and 32 tubes. As the system further prohibits the use of multichannel pipettes as well as most liquid handling systems, there are limited ways to reduce the required time per sample. Although solutions for 96-well formats are available, the system still needs a centrifuge and therefore specialised robot systems for automated high-throughput workflows.

These are usually very expensive and somewhat limited in their applications, making them not financially feasible for most laboratories.

An elegant alternative solution for effective and scalable high-throughput nucleic acid extraction can be provided by magnetic nanoparticles (MNPs), commonly known as magnetic beads [229,230]. These particles utilize solid-phase reversible immobilisation (SPRI) [230]. Under dehydrating conditions, MNPs can reversibly bind DNA or RNA, while being immobilized by a magnet during several wash- or manipulation-steps to either remove contaminants or perform on-bead enzymatic reactions like digests. As such bead-based protocols neither require a centrifuge nor expensive materials, they are easily and arbitrarily scalable, making them perfectly suited for most common liquid handling machines.

MNPs are usually composed of a magnetic core functionalised with a chemical coat. There are several possible methods for MNP synthesis [231–234]. The core typically consists of either polystyrene (PS) coated ferrites [231] or co-precipitated highly magnetic iron(II,III) oxide particles (Fe_3O_4) [232]. These core-particles are then functionalised with a chemical surface containing e.g. silica- or carboxyl-groups. Silica-layers are usually synthesised by hydrolyzing tetraethyl orthosilicate (TEOS) [232,235], while carboxyl-surfaces are produced by polymerizing methacrylic acid (MAA) [236].

Despite the striking advantages over column-based systems, very little community effort has been conducted to develop and promote MNP-based extraction protocols.

1.4 Aims of the thesis

Designing a toolbox for modular, parallel and inducible epigenetic editing (Figure 1-12A)

The epigenetic network regulates the expression of genes and is therefore essential for cell differentiation, the crosstalk between cells, and consequently for the development of any multicellular organism. Dysregulation of the epigenetic program can have severe consequences for the affected cell and even for the whole organism. Several diseases, including cancer, have been associated with malfunctions in the epigenome. To decipher the underlying epigenetic code, researchers initially focused on descriptive approaches. However, these approaches seemed unable to answer pressing questions about the functionality and causality of epigenetic modifications.

The CRISPR/Cas system offers a flexible and easy way of distinctly perturbing the epigenetic code; a functionality which is exceptionally well suited to answer these questions. Although this system has already proved to be an invaluable tool, it is relatively inflexible when it comes to multiplexing approaches and modularity.

Hence, the first aim of my thesis was the development and validation of a method that does not use dCas9-fusion constructs, but instead utilizes aptamer structures and the respective binding proteins to unite the targeting- and recruiting-function in one molecule, the gRNA. I designed these gRNAs in a way that the incorporated aptamer sequences can be arbitrarily exchanged in a simple 4-step procedure. Additionally, I wanted to create a system that allows the easy combination of an aptamer binder (AB) with the desired ED. I then validated this system by recruiting the transcriptional modulators VPR and KRAB to five differently expressed target genes and monitoring the induced changes in transcription. I further planned to design ligand-dependent mechanisms to turn this system ON or OFF. These switches are based on the FKBP/FRB dimerizers and the Csy4 protein. Both methods work independently from each other and can be finetuned by altering the ligand concentration. This toolbox was later named EPIC'RISPR.

Highly parallel large-scale application of EPIC'RISPR (Figure 1-12B)

I planned to apply the EPIC'RISPR system for large-scale experiments to investigate the effects of several dozens EDs on five differently expressed target genes simultaneously. I aimed to identify the most efficient epigenetic modulators by observing the changes in the transcriptional states of the targets and possibly determine general principles of epigenetic regulation. For this, I focused on histone lysine methylation and demethylation as well as histone acetylation and deacetylation.

Development of imidazole-induced enzymes and induced gene control (Figure 1-12C and Figure 1-12D)

In many enzymes which utilize H in their catalytic centre, the enzymatic activity can be disturbed by exchanging the H for an A residue. It has been shown for several enzyme classes, that the respective activity can be rescued by imidazole, which might assume the role of the imidazole ring of H. In this thesis, I aimed to design and validate this technique for Csy4, and eventually apply it for controlling the expression of endogenous and transgenic genes.

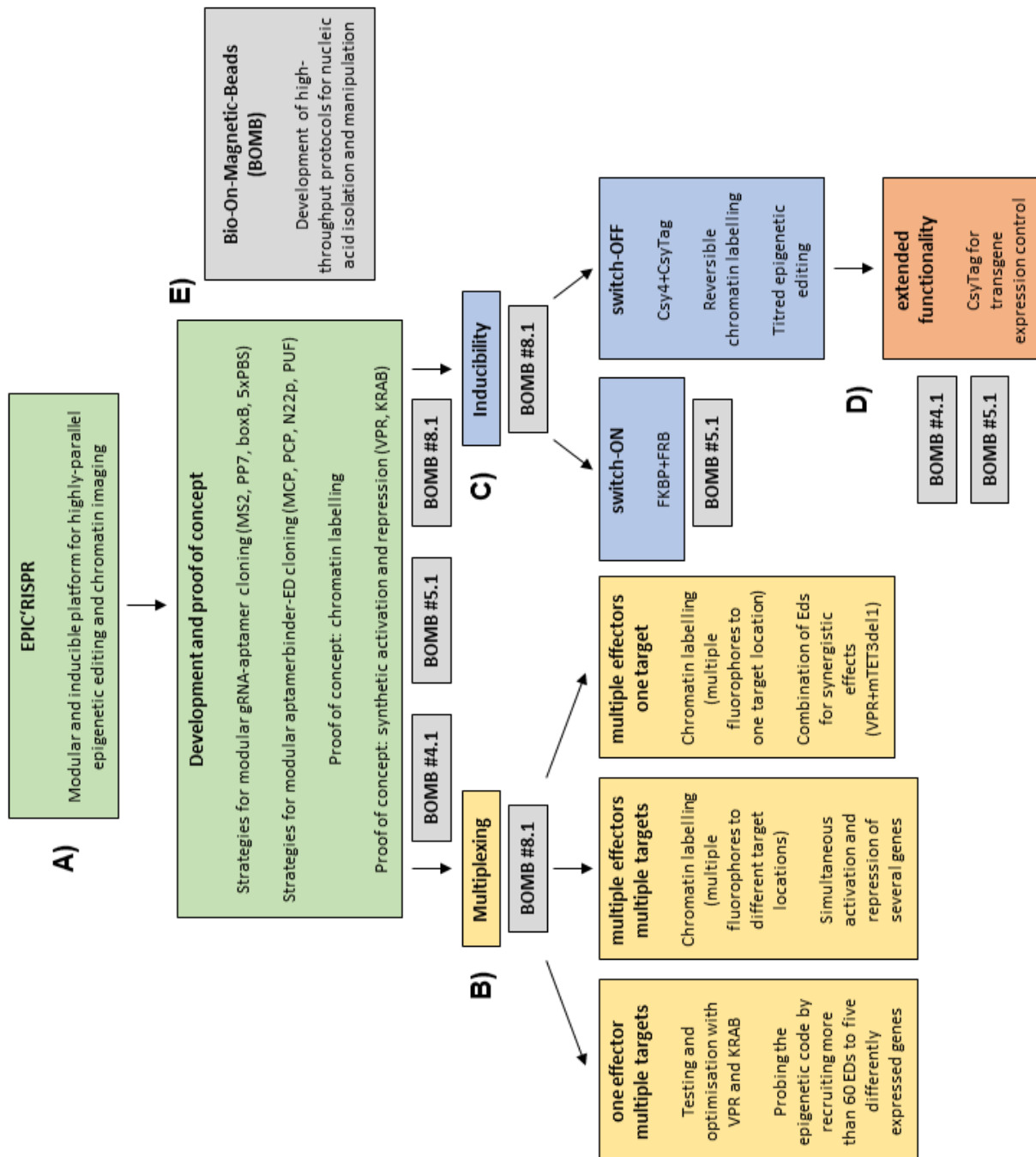


Figure 1-12: Roadmap of the thesis. **A)** To design the EPIC'RISPR platform, several cloning strategies needed to be developed and the proof of concept had to be established. **B)** A first application of the EPIC'RISPR system was the parallel modification of one or multiple genes with one or several EDs. This system was further used for multiplexed chromatin labelling. **C)** To titre the transcriptional modulation effect, I wanted to develop an ON- as well as an OFF-switch. For this, I wanted to utilize the FKBP/FRB system as well as a mutant of Csy4, and its recognition site as a degradation tag. **D)** Using the extreme versatility of the CsyTag system, I wanted to apply it for transgene expression control as well. **E)** To process the samples necessary for these projects, we created the BOMB platform for high-throughput extraction and purification of nucleic acids.

Developing high-throughput protocols for the cost-effective handling of nucleic acids (Figure 1-12E)

Confronted with the enormous amount of samples required for the experiments in this thesis, we aimed to develop high-throughput procedures for cost-effective extraction of nucleic acids. In this project, I focused on protocols for the isolation of DNA from bacterial cultures and various solutions, such as PCR reactions, as well as on the isolation of RNA from mammalian cells. Furthermore, I participated in the optimisation of protocols for the synthesis and functionalisation of magnetic beads.

2 Results

2.1 Bio-On-Magnetic-Beads (BOMB)

In epigenetics we not only investigate the factors that potentially read, introduce or erase a specific mark, but also the meaning and stability of said marks at various positions in the genome and at different time points. This leads to experiments with a notoriously high sample count and therefore extremely high costs, both regarding money and time. To achieve the goals of this thesis and perform the planned experiments, we first needed to employ techniques that would allow us to process large quantities of samples quickly, efficiently, reliably and financially feasible. Although the development of such techniques was not the main focus of my PhD studies, it still was a crucial factor for the success of all the results in this thesis.

As most of the experiments in our research revolve around the handling of nucleic acids, we decided to focus on the improvement of methods for the isolation and manipulation of DNA and RNA. The established commercial solutions use either column-based systems or employ MNPs with a modified surface to capture nucleic acids. As MNP-based protocols are easily scalable, both in the number of samples and used volumes, we decided to develop open-source protocols for their synthesis, functionalisation and application in nucleic acid extraction. We first started synthesizing and functionalising MNPs ourselves (**Figure 2-1 (1)-(3)**) by reverse- and forward-engineering existing methods. Together with Dr Timothy Hore and his group in New Zealand we established a broad range of protocols for their use in extracting RNA, DNA, and total nucleic acid (TNA) from a variety of sources (**Figure 2-1 (4)-(9)**). These include but are not limited to cultured mammalian cells, *Escherichia coli* (*E. coli*), *Saccharomyces cerevisiae*, plant and mammalian tissues. Considering the costs of commercial kits and MNPs, this represented a real game-changer for our research.

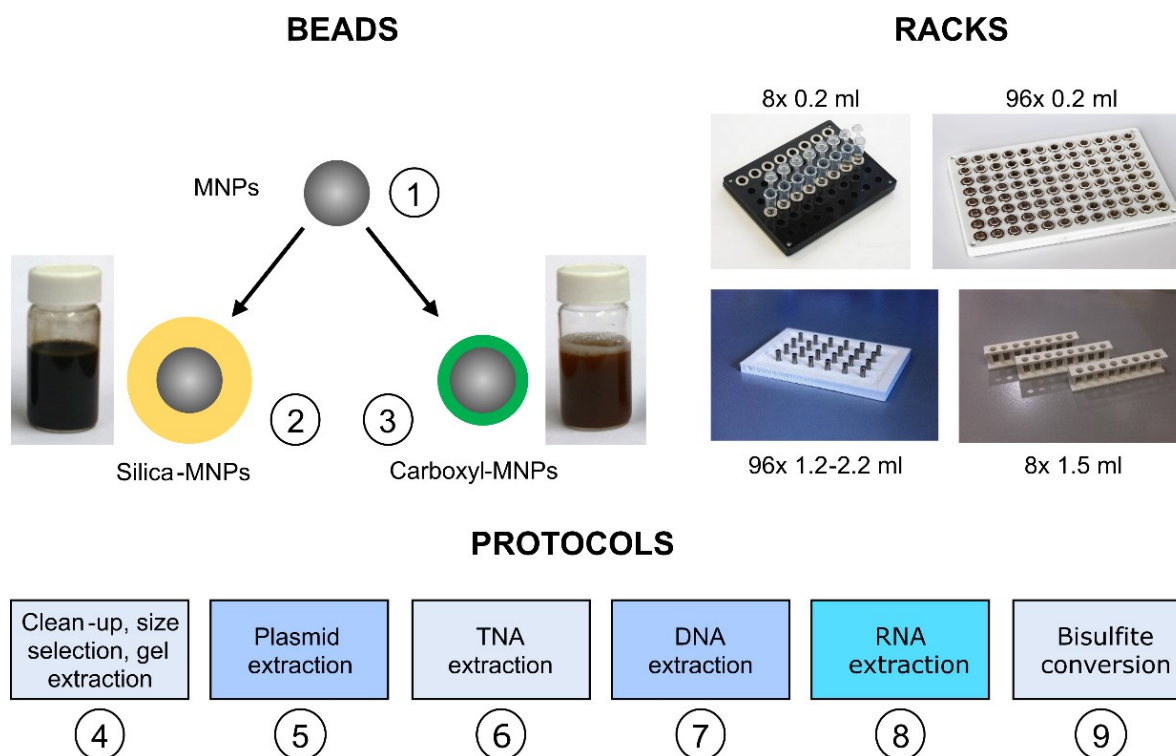


Figure 2-1: The BOMB platform. Figures and figure text were taken from Oberacker et al. [237]. The BOMB platform is composed of (1) ferrite MNPs that can be coated with either a (2) silica or (3) carboxyl surface, BOMB magnetic racks produced by laser cutting or 3D printing, and (4–9) the basic BOMB protocols for purification of nucleic acids from various sample origins. The circled numbers indicate the protocols for the respective procedure.

When designing these protocols, we focused on reliability and robustness as well as on modularity and automation friendliness. Almost all protocols follow a pattern of:

lyse → bind → wash → (manipulate → wash →) elute

Most of these protocols share similar or identical buffer systems, which further enables us to combine various protocols in a modular fashion for specialised applications. We further designed and published various magnetic rack systems that can easily be crafted with a 3D printer or a laser cutting device (**Figure 2-1**). These can be used for common microcentrifuge tubes, 96-well microtiter- or deep-well plates. The racks utilize small neodymium magnets that are obtainable online for a small price. Together with the synthesised beads and the developed protocols, these racks form the open-source platform Bio-On-Magnetic-Beads (BOMB), which we made publicly available for the broader research community [237]. A more detailed protocol for the synthesis of MNPs and their functionalization was later published with bio-protocol [238]. To promote community engagement in developing new and modifying existing BOMB

protocols, we installed a web-forum for open communication and exchange of knowledge¹. Considering the major financial and technical advantages of MNP-based protocols for nucleic acid extraction, we hope the BOMB platform to become a pioneer in open-source science.

My focus in this project was on the following four parts: Firstly, synthesis and functionalization of MNPs, secondly, designing clean-up and size-exclusion protocols for *in vitro* reactions like PCRs or enzyme digests, thirdly, establishing procedures for the isolation of plasmid DNA from *E. coli* and fourthly, developing and optimising RNA extraction protocols from mammalian HEK293 cells.

2.1.1 MNP synthesis and coating

The initial protocol for the synthesis and functionalization of the BOMB MNPs was designed by Dr Tomasz Jurkowski. I was involved in the later optimisation process of both the core particle synthesis as well as the coating protocols.

We based our synthesis on the co-precipitation method published by Choi et al. [232]. The reaction is very robust and uses a set-up of common laboratory equipment (**Figure 2-2A**), an optimal feature for a broader community. In short, a solution containing iron(II) chloride (FeCl_2) and iron(III) chloride (FeCl_3) in 1:2 molar ratio is slowly dripped into preheated and/or degassed sodium hydroxide (NaOH). The forming black precipitate consists of highly magnetic Fe_3O_4 (**Figure 2-2B**) of about 5 to 20 nm in size (**Figure 2-2C**). We then successfully encased these particles in a layer of silica-groups using the Stöber method [235]. In this reaction, TEOS hydrolyzes under basic conditions, depositing a silicon dioxide (SiO_2) layer around the ferrite core. Conveniently, the radius of this layer depends on the ratio of core particles to TEOS [239] and can hence be altered arbitrarily. Our standard ratio yielded silica-MNPs with an approximate diameter of 400 nm (**Figure 2-2D**). The silica-coat not only prevents oxidation of the core MNPs but also provides an inert surface for the precipitation of nucleic acids. Alternatively, the core can be coated with a layer of carboxyl-groups (**Figure 2-2E**). For this, we used MAA monomers, which can enter a free-radical retrograde precipitation polymerization reaction to form polymethacrylic acid (PMAA)

¹ <https://bomb.bio>

as a coating around the ferrite MNP core [236]. Although it does not necessarily provide the same stability as a layer of silica, the carboxyl-layer grants the MNPs a weak negative charge, thereby altering the bead functionality.

These protocols describe a fast, easy and inexpensive procedure for the synthesis and functionalization of MNPs that are well suited for the isolation of nucleic acids from a variety of sources.

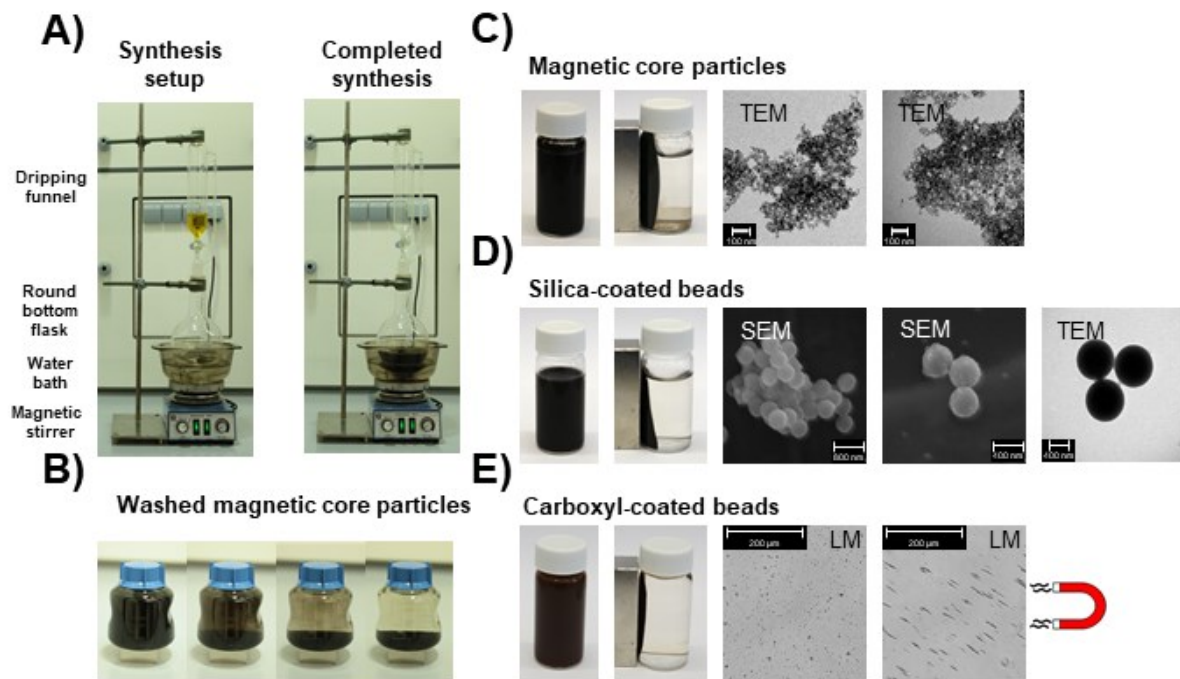


Figure 2-2: Synthesis and functionalization of MNPs. Figures and figure text were taken from Oberacker et al. [238] and have been slightly modified. **A)** The basic set-up of equipment for core particle synthesis. **B)** Time-lapse of the behaviour of synthesized and washed core particles on a strong magnet. **C)** Core particles in transmission electron microscopy (TEM). **D)** Silica-coated MNPs in scanning electron microscopy (SEM) and TEM. **E)** Carboxyl-coated MNPs in light microscopy (LM), with and without an applied magnet.

2.1.2 Clean-up and size exclusion

To purify nucleic acids from enzymatic reactions, we exploited the fact that their affinity towards siloxane- and/or silanol-groups alter in the presence of chaotropic salts [240,241]. I composed a guanidine hydrochloride (Gu-HCl)-based binding buffer that allows the capture of DNA using the above-described silica-MNPs (see chapter 2.1.1). After several wash steps with 80% ethanol (EtOH), the purified nucleic acid can be eluted from the beads by the addition of double-distilled water (ddH_2O) or Tris-EDTA (TE) buffer. Using this method, I was able to successfully extract plasmid DNA from a

solution with known concentration with a yield of $95\pm 1.6\%$ (**Figure 2-3A**). I further compared and combined the binding media (silica-MNPs) and the binding buffer with a commercial, column-based solution. A clean-up of the same sample with the commercial kit resulted in a recovery of $80\pm 3.3\%$, demonstrating the high quality of the BOMB protocols compared to commercial solutions. The combination of a commercial column and house-made binding buffer yielded $99\pm 7.5\%$ of the input, and the commercial binding buffer together with our silica-beads $33\pm 0.7\%$.

I then tested the capacity of the beads by purifying different amounts [μg] of plasmid DNA and analysed the recovered yields (**Figure 2-3B**). The protocol achieved an excellent linear correlation ($R^2=0.99$) between input and output up to $95\ \mu\text{g}$ with an average yield of $78\pm 1.2\%$ for a one-year-old bead solution. At higher amounts of input, the beads reached their capacity limit and started to form flakes during the washing steps (**Figure 2-3C**). This is also reflected in a sudden drop in recovery between 100 and $250\ \mu\text{g}$ ($42\pm 3.0\%$).

To further investigate the binding capabilities of the silica-MNPs, I tested the effect of different salt concentrations during the binding step as well as the influence of DNA length. For this, I purified a molecular weight marker with various amounts of binding buffer (**Figure 2-3D**). This enabled me to isolate single-stranded DNA fragments from 100 to 3000 bp selectively. Inspired by these results, I probed the behaviour of even smaller fragments in high salt concentrations (~ 2.7 M, **Figure 2-3E**). I wanted to see whether a high enough salt concentration would allow the removal of single nucleotides while small DNA fragments such as primers could still be attached to the beads. I was able to successfully purify single-stranded oligonucleotides as small as 25 nucleotides (nt) with a yield of approximately 30% , while larger oligos (71 nt) yielded about 80% of the input. As the lengths of these oligonucleotides are comparable to the sizes of small ncRNA, I performed the high-salt protocol on samples of *in vitro* transcribed transfer-RNA (tRNA, 75 nt) and gRNA (130 nt). While the efficiencies of recovery are lower compared to DNA molecules of the same length, tRNA still yielded $22\pm 1.4\%$ of the input and gRNA even $36\pm 3.1\%$.

In conclusion, this protocol states an effective solution for the clean-up and selective size exclusion of nucleic acids (plasmid DNA, single-stranded DNA, small ncRNA) from a range of 25 to several thousand nts. The procedure is not only comparable to commercial solutions but even outperforms them in certain applications.

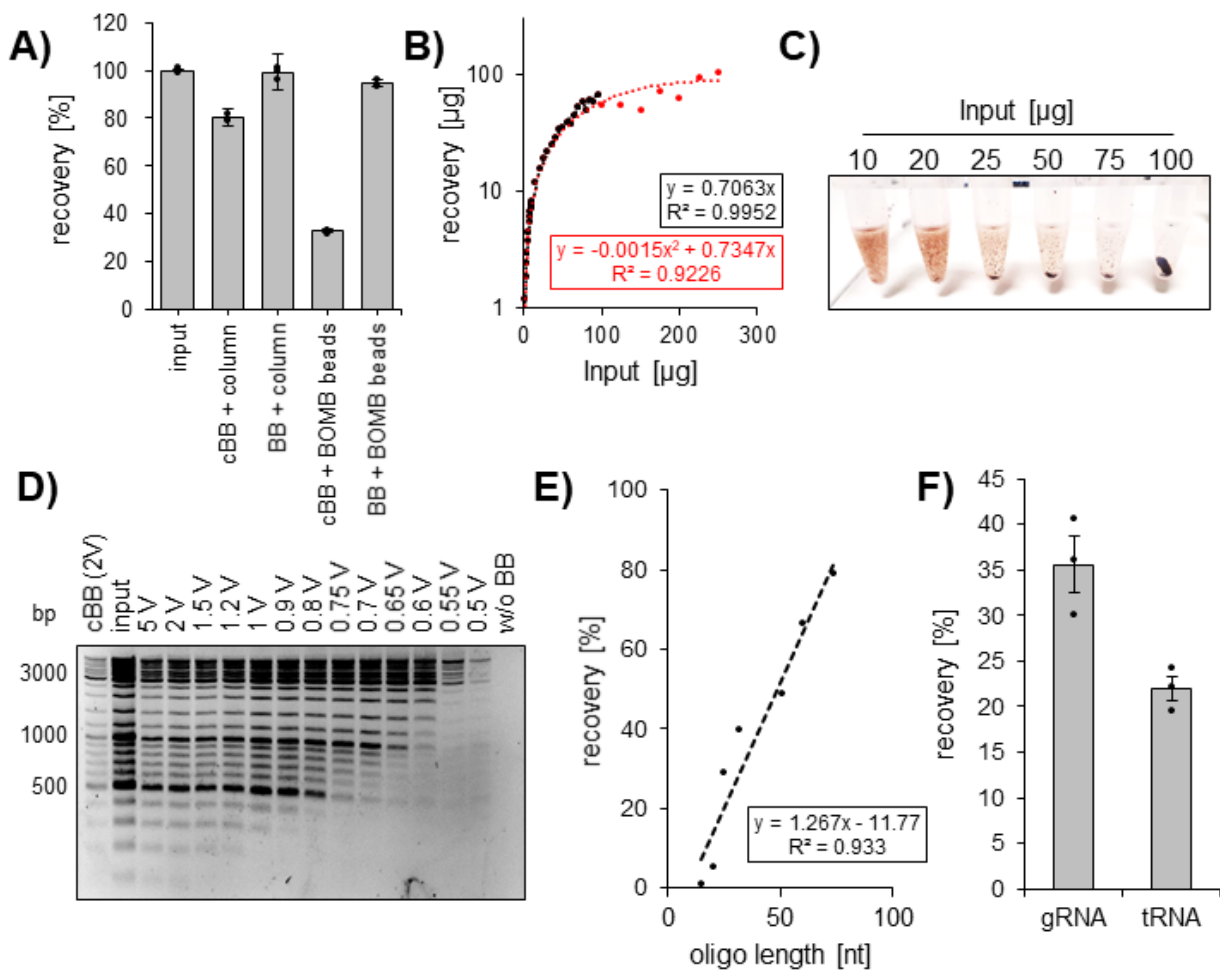


Figure 2-3: BOMB clean-up and size-exclusion with B silica-coated MNPs. Figures A and D were taken from Oberacker et al. [237] and have been slightly modified. **A)** Total recovery of ~6 µg plasmid DNA (input) using either a commercial kit that includes silica-packed columns or the here developed clean-up protocol with silica-coated beads (BOMB). For the binding, 2 volumes (V) of either commercial binding buffer (cBB) or the binding buffer (BB) described in the BOMB protocol above was used. Error bars represent the standard error of the mean (SEM), n=3. **B)** Recovered plasmid DNA [µg] over the respective input [µg]. The graph shows a good linear correlation ($R^2=0.99$) between input and output up to 100 µg (black) with a yield of about 71% using 5 µl of a 12-month-old bead stock solution (1:1 of wet mass to d_4H_2O). Higher amounts of input do not yield more DNA due to bead capacity (red). **C)** Clumping behaviour of BOMB silica-beads loaded with different amounts of plasmid DNA during washing with EtOH. **D)** Size exclusion of GeneRuler DNA Ladder Mix (Thermo) using BOMB silica-coated magnetic beads. Different V of binding buffer compared to sample V were used to achieve size exclusion. 2 V of cBB was used as a control relative to the input. **E)** Total recovery of single-stranded oligonucleotides of different lengths using 10 V BB. **F)** Total recovery of two types of *in vitro* transcribed small ncRNAs using 10 V BB. Error bars represent SEM, n = 3.

2.1.3 Plasmid isolation

To perform the screening experiments that will be described in chapter 2.5, it was necessary to develop strategies that allow the cloning of hundreds of constructs in a reasonable amount of time. A crucial time-limiting step is hereby the extraction of plasmid DNA from bacterial cells. Most commercial or home-made solutions apply alkaline lysis protocols [242–244]. With the help of Sven Höhn and Luca Schelle, two students I mentored during their respective MSc and BSc theses, I was able to reverse-engineer the published protocols utilizing our BOMB silica-MNPs for the isolation of plasmid DNA in a 96-well format. In short, cells are lysed using NaOH. After lysis, the solution is neutralized using a potassium acetate buffer with a high concentration of Gu-HCl. Under these conditions, the gDNA will precipitate and can be separated from the plasmid DNA by centrifugation. Furthermore, the high salt concentration allows the plasmid DNA to bind to the silica-MNP, which are added subsequently. After a few washes with EtOH, the DNA can be eluted. Using this protocol, we were able to obtain an average yield of 6.2 ± 0.2 µg per 2 ml of bacterial culture with an average $A_{260 \text{ nm}}/A_{280 \text{ nm}}$ ratio of 1.81 ± 0.003 ($n = 96$). The total distribution is plotted in **Figure 2-4A**, together with exemplary samples that were purified using commercial kits (red). We further tested the performance of the purified plasmid DNA in restriction digests (**Figure 2-4B**) as well as in Sanger sequencing (**Figure 2-4C**), with comparable results to commercially purified DNA.

Although the conversion into a 96-well system was already an outstanding improvement for our daily workflow, we were thinking further and wanted to make the protocol feasible for automation. The centrifugation step after the neutralization was a critical limitation, as it is challenging to implement the protocol into a fully automated format. I bypassed this step by designing a buffer system that splits the neutralization buffer in two: a potassium acetate buffer for neutralization and a Gu-HCl buffer for capturing the DNA. After neutralization, I added silica-MNPs to the solution and applied a magnet. While the cellular debris and the gDNA were physically pulled to the magnet, the plasmid DNA stayed in solution under these conditions. After combining the supernatant with the high-salt buffer and a fresh batch of beads, the plasmid DNA could be captured and purified as described before. The obtained DNA is comparable in amount and quality to the centrifugation-based protocol (**Figure 2-4D**).

As the bacteria are cultured in 2.2 ml deep-well plates to accommodate the 96-well format, the oxygen transfer during the incubation step is limited compared to a test tube set-up. To fully optimise the protocol, I adjusted the growth conditions of the bacterial cultures. I first compared different growth media and volumes to optimise the ratio of bacterial growth and oxygen supply (**Figure 2-4E**). Terrific Broth (TB) as a basis in a 2 ml culture performed best and resulted in an average yield of $7.5 \pm 0.3 \mu\text{g}$ ($n = 3$) plasmid DNA after 22 h of growth. I then tested a variety of culturing times from 15 to 24 h (**Figure 2-4F**). With an average yield of $9.7 \pm 0.8 \mu\text{g}$ ($n = 5$) after 24 h of incubation, this time point was significantly higher ($P < 0.05$) than any sample below 22 h of incubation.

Here I demonstrate the conversion of a column-based, low-throughput system for plasmid DNA isolation into a 96-well format fitted for automation processes that perform as well as its commercial counterpart regarding quality and downstream applications.

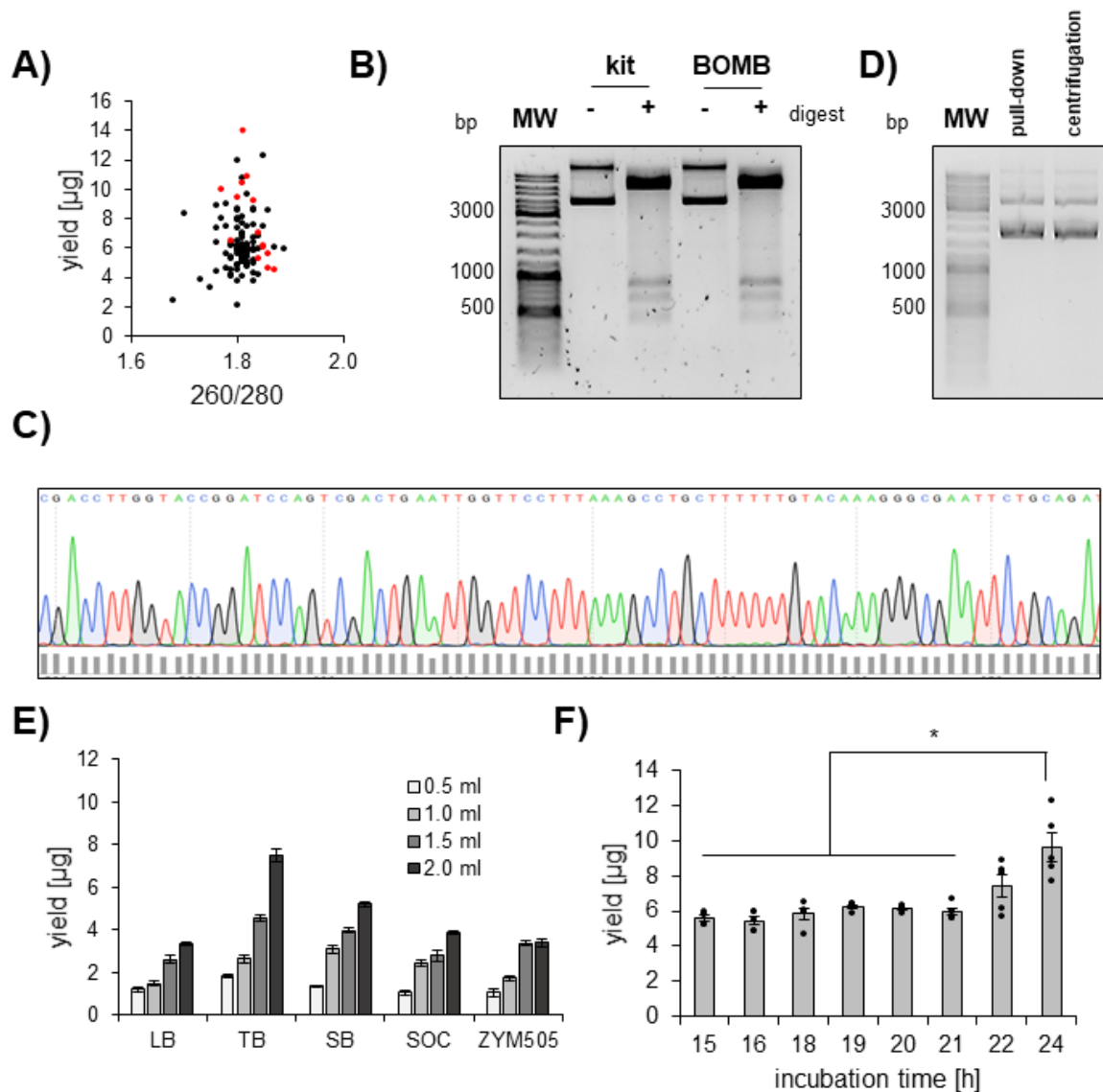


Figure 2-4: Quality control and optimisation of BOMB plasmid extraction. Figures and figure text were taken from Oberacker et al. [237] and have been slightly modified. **A)** Total plasmid DNA yield [µg] extracted from *E. coli*, plotted against the $A_{260\text{ nm}}/A_{280\text{ nm}}$ ratio for each sample. Black dots represent samples extracted using the plasmid extraction protocol; red dots represent samples processed using a commercial kit. **B)** Comparison of commercially purified pUC19 plasmid DNA (kit) and BOMB extracted DNA (BOMB) with (+) and without (-) restriction enzyme digestion. MW: GeneRuler DNA Ladder Mix (Thermo Scientific). **C)** Comparison of EFS plasmid DNA extracted with BOMB either using a centrifugation-based protocol or a protocol utilizing MNPs for the physical pull-down of cellular debris. MW: Gene Ruler DNA Ladder (Thermo Scientific). **D)** Exemplary sequencing trace of BOMB extracted plasmids with Sanger sequencing. A sequencing read length of at least 800-1000 nts is typically observed. **E)** Optimisation of reaction volume and media. LB: Luria-Bertani Broth, TB: Terrific Broth, SB: Super Broth, SOC: Super Optimal Broth, ZYM505: Zymol Medium 505. Error bars represent the SEM, n = 3. **F)** Optimisation of incubation time. Error bars represent the SEM, n = 5, *P < 0.05, two-sided t-test.

2.1.4 RNA isolation from mammalian cells

In this thesis, I am mainly focusing on the analysis of expression levels. Therefore, the isolation of total RNA, including mRNA is an essential step of almost all my experiments. Initially, we performed these isolations using commercial column-based systems, which were not suitable for the experimental workflow that we had planned with several hundreds of isolations simultaneously. Together with my colleague Peter Stepper, and the students we mentored (Sven Höhn, Jule Focken, and Vivien Meyer), I developed and optimised a protocol for the high-throughput extraction of total RNA from mammalian HEK293 cells.

We utilized a highly effective TRI buffer system [245,246] for the lysis of the cells and the retainment of the RNA. If combined with a Gu-HCl-containing buffer, this reagent allows nucleic acids to be captured by added silica-MNPs. While nucleic acids stick to the beads, contaminants including the phenol in the TRI reagent can be washed away with 90% EtOH. An on-bead digest with DNase I degrades the co-purified gDNA. After several more wash steps, the RNA can be eluted and used for further reactions. RNA purified with this protocol yields an average of 7.2 ± 0.3 μg per $0.5 \cdot 10^6$ cells with an $A_{260 \text{ nm}}/A_{280 \text{ nm}}$ ratio of 1.89 ± 0.001 ($n = 96$) (**Figure 2-5A**). The RNA is fully intact with mRNA, tRNA and ribosomal RNA fully preserved (**Figure 2-5B**) and it is further suited for delicate downstream-processes such as real-time quantitative polymerase chain reaction (RT-qPCR, **Figure 2-5C**). As an alternative to the TRI based isolation system, our collaborator Dr Timothy Hore and his group designed a variety of guanidine thiocyanate (GITC)-based protocols. These protocols avoid the toxic phenol used in this method, which provides a safer working environment.

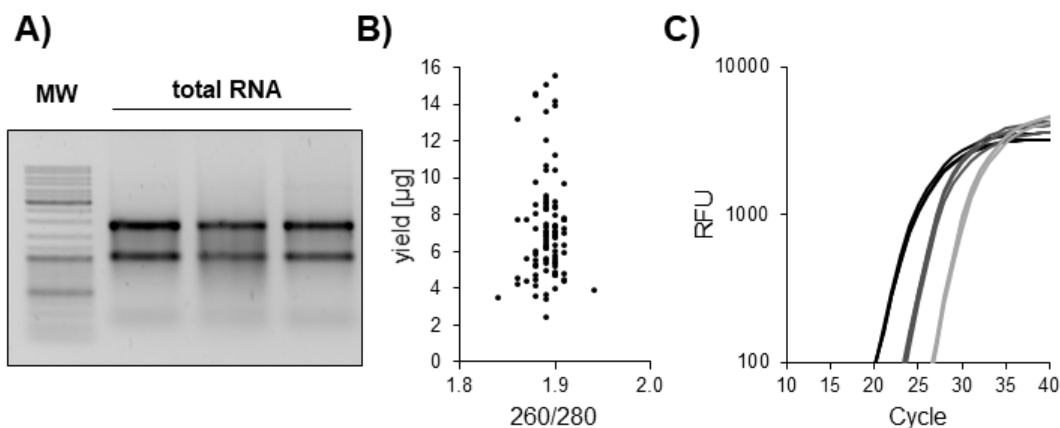


Figure 2-5: Quality control of total RNA isolated using BOMB protocols Figures and figure text were taken from Oberacker et al. [237] and have been slightly modified. **A)** Total RNA isolated from mammalian HEK293. MW: Gene Ruler DNA Ladder (Thermo Scientific). **B)** Total RNA yield [μg] extracted from *E. coli*, plotted against the $A_{260\text{ nm}}/A_{280\text{ nm}}$ ratio for each sample. **C)** The amplification curve of cDNA reverse transcribed from BOMB extracted total RNA. dilution series: black = undiluted, dark grey = 10-fold dilution, light grey = 100-fold dilution.

2.2 Development of the EPIC'RISPR platform

The most significant advantages of the EPIC'RISPR platform over any dCas9-fusion based system are its modularity and the ability to be easily used for multiplexing set-ups. To achieve these two characteristics, I needed to develop cloning strategies that resulted in toolbox systems for attaching arbitrary aptamers to the 3'-end of the gRNAs as well as for fusing ABs and EDs.

2.2.1 The gRNA toolbox

To construct functional gRNAs, a three-component toolbox was designed. These components are the spacer sequence, the scaffold and the aptamers (**Figure 2-6**). A previously published gRNA vector [103] which contains a stuffer region for a one-step insertion of spacer sequences (Supplementary Data **Table 0-1**) by golden gate assembly [247,248] served as a basis. This vector was modified with a new, optimised gRNA scaffold for effective targeting [180]. To obtain this construct, firstly, a BsaI restriction site was removed from the original plasmid [103] using ccdB_out_f and ccdB_out_r (Supplementary Data **Table 0-2**) via PCR and ligation. The construct was further modified in the same manner in two successive steps using U6_crRNA_f and _r and U6_gRNA_in_f and _r (all Supplementary Data **Table 0-2**). The resulting construct contains not only the optimised gRNA scaffold but also a second stuffer

region at the 3'-end of the gRNA (Supplementary Data **Table 0-3**). This second stuffer contains two BsaI restriction sites with internal binding sequences, which allows the insertion of aptamer sequences utilizing a hierarchical cloning strategy that has already been established for TALEN [137].

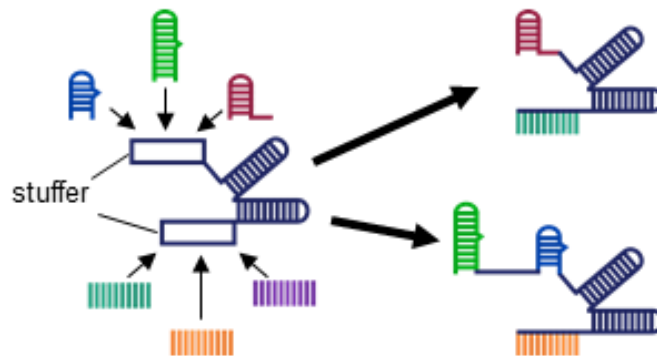


Figure 2-6: Schematic representation of the EPIC'RISPR gRNA toolbox. Two stuffer systems can be exchanged for an arbitrary combination of spacer sequence and aptamer in a hierarchical Golden Gate assembly procedure.

In this 4-step hierarchical cloning procedure, the aptamer sequences were successively modified, digested and ligated to assemble in a specific order to an oligomer that can be introduced into the vector (**Figure 2-7**). The desired aptamers were provided as oligonucleotides with specific flanking sequences. In the first step (**Figure 2-7 I**), these flanking sequences were used to amplify the aptamers in a PCR reaction simultaneously modifying the flanks with BsaI and/or BsmBI restriction binding sites. After digestion with BsmBI, these amplicons contained specific sticky ends and were ligated to a circular oligomer in a predetermined order (**Figure 2-7 II**). This oligomer was then amplified in a PCR reaction and digested with BsaI, resulting in a linear oligomer with specific sticky ends (**Figure 2-7 III**). To fit this construct into the target vector, the stuffer region was removed by a BsaI digest, leaving it with matching ends. In a final ligation step, the oligomer was introduced into the vector (**Figure 2-7 IV**). The final product was purified and transformed into bacteria. All sequences are provided in **Table 0-3** and **Table 0-4** of the Supplementary Data section. This toolbox represents an efficient and quick 4-step step process for the equipment of a gRNA with an arbitrary number and combination of aptamer sequences.

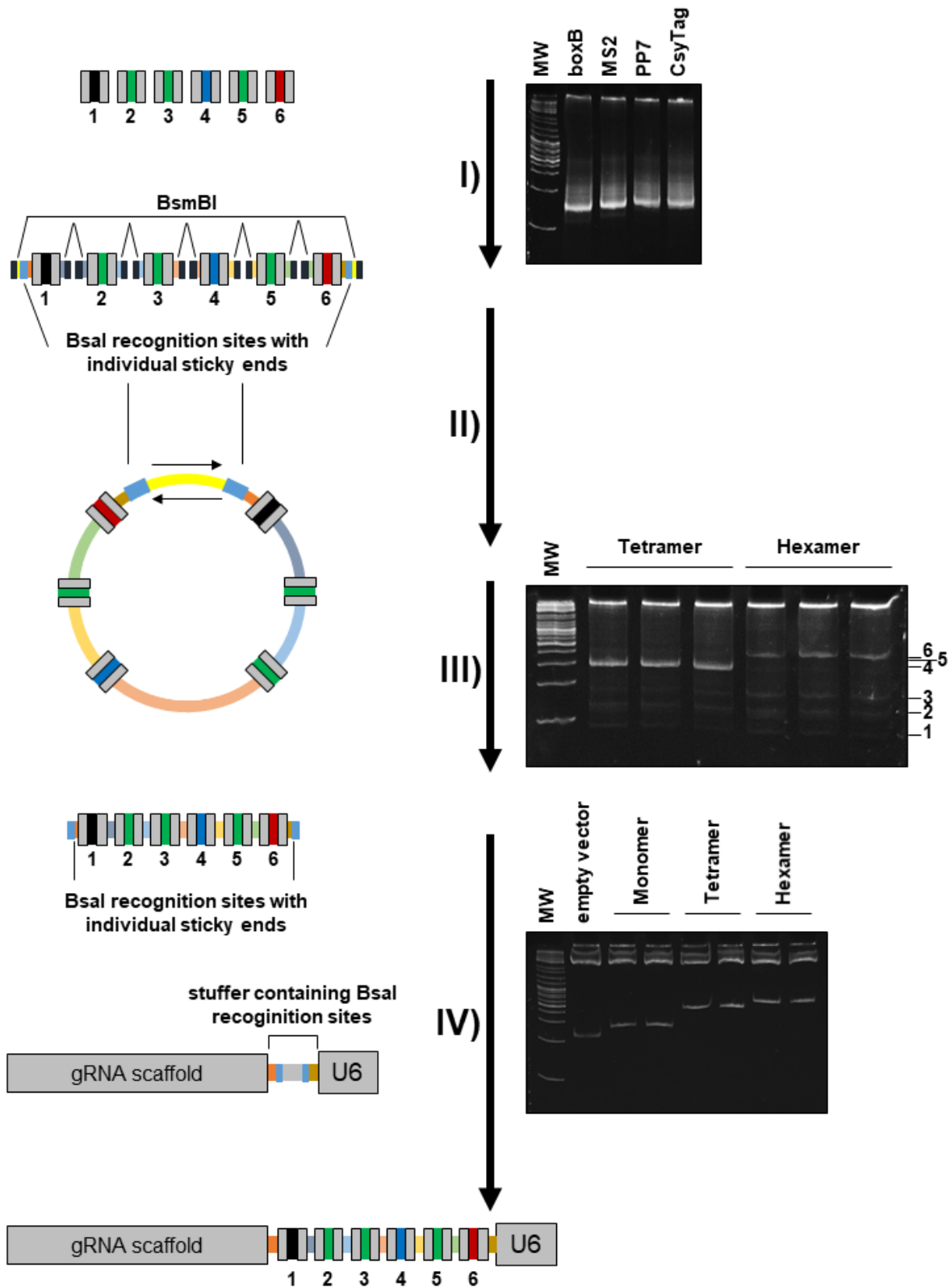


Figure 2-7: Hierarchical cloning strategy for the modular functionalization of gRNAs with arbitrary combinations of aptamers. I) Addition of ligation adapters containing BsmBI and Bsal recognition sites with individual restriction products (sticky ends) by PCR. II) Restriction digest allows the ligation of an arbitrary combination of cassettes to circular molecules (e.g. hexamers). An additional exonuclease treatment ensures the removal of remaining linear molecules amplification. III) The circular

molecules can be amplified in a PCR reaction, leading to linear oligomers (1-6) with individual BsaI restriction sites on either side. **IV)** A BsaI restriction digest prepares the oligomer for the insertion into the target vector. The vector itself is also treated with BsaI to release the stuffer sequence, creating fitting sticky ends for the combination with the oligomer in a golden gate assembly reaction. MW: GeneRuler DNA Ladder Mix (Thermo Scientific).

2.2.2 The effector toolbox

For the modular assembly of any combination of AB and ED, a system was created that allows the arbitrary exchange of both proteins while connecting them by a linker that grants good sterical flexibility (**Figure 2-8A**). For this, a series of vectors were modified, which had previously been established for targeted chromosome labelling [180]. These vectors contain either MCP, PCP or N22p at the C-terminus of the open reading frame (ORF) followed by a glycine-serine-serine (GSS) linker and three successive fluorophores. To modify them, a synthetic 39 amino acids long sequence was inserted. By replacing the fluorophores, this sequence extends the linker while also providing a unique PstI restriction site at the C-terminus, upstream of FLAG tag and a nuclear localisation sequence (NLS). To obtain this construct, the fluorophores were removed by restriction digest with XhoI and BamHI, and the vector backbone was isolated using BOMB protocol #4.3 [237]. The synthetic linker (Supplementary Data **Table 0-5**) was modified in a PCR reaction using gBLOCK-rain-f and -r (Supplementary Data **Table 0-6**) and inserted into the vector via ligation. The sequence for this construct is provided in **Table 0-7** of the Supplementary Data section. The desired EDs were amplified from cDNA and modified with flanking sequences identical to the flanks of the PstI site in the vector (Supplementary Data **Table 0-8**). To ensure that the respective EDs were functional, we either amplified the full-length protein or determined the smallest functional subunits for which crystal structures have been published [39,212]. The synthetic transactivator domains VP64 and VPR were amplified from SP-dCas9-VPR (Addgene #63798, gift from George Church) [109], KRAB from pHAGE EF1 α dCas9-KRAB (Addgene #50919, gift from Rene Maehr & Scot Wolfe) [249], and mTET3del1 from pET28a (+)_mTET3-CD-del1 (gift from Mirunalini Ravichandran) [250]. FKBP and FRB were amplified using the primers pHAGE_DmrA_PstI_f and _r or pHAGE_DmrC_PstI_f and _r (Supplementary Data **Table 0-9**), respectively. As templates, DmrA_FLAG_CDS and DmrC_HA_CDS (Supplementary Data **Table 0-5**) were utilized. The amplified domains were subsequently inserted into the linearized vector by Gibson assembly. The success of

the cloning was controlled by an insert release with PstI (**Figure 2-8B**) and validated by Sanger sequencing.

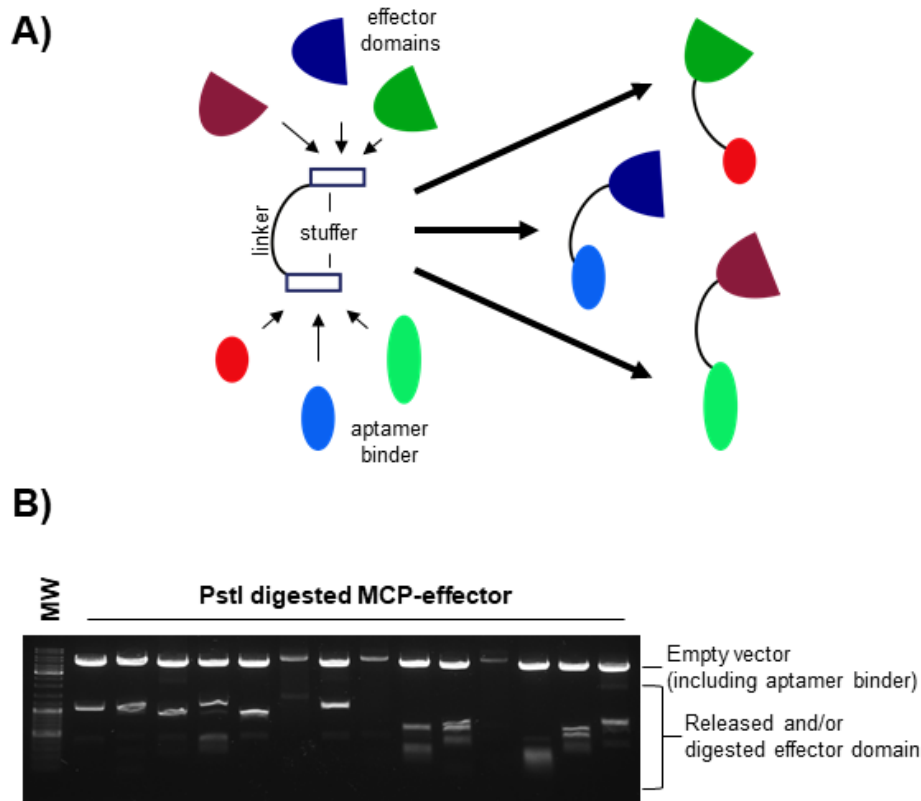


Figure 2-8: The EPIC'RISPR effector toolbox allows the arbitrary combination of aptamer binder and effector domain. A) Schematic representation of the EPIC'RISPR effector toolbox. Two stuffer systems, separated by a GSS linker, can be exchanged for an arbitrary combination of aptamer binder (N-terminus) and ED (C-terminus) using Gibson Assembly. The constructs also contain a FLAG-tag as well as an NLS at the C-terminus of the respective effector molecules (not shown for better clarity of the schematic). **B)** Exemplary insert release of MCP-effector constructs with PstI. The enzyme releases the inserted ED and, in some cases, also cuts inside the ED itself, leaving a specific pattern that helps with identification. All constructs have been confirmed with Sanger sequencing. MW: GeneRuler DNA Ladder Mix (Thermo Scientific).

The AB domains were exchanged similarly. The respective domain of interest was amplified using PCR and equipped with specific flanking sequences (Supplementary Data **Table 0-9**), while the vector was linearized either by PCR amplification (open_gBLOCK_f, _r, Supplementary Data **Table 0-6**) or an enzyme digest. The PCR product was then introduced by Gibson assembly. To insert FKBP or FRB, the respective sequence was amplified using pHAGE_EFS_DmrA_NcoI_f and pHAGE_EFS_DmrA_SpeI_r or pHAGE_EFS_DmrC_NcoI_f and pHAGE_EFS_DmrC_SpeI_r (Supplementary Data **Table 0-9**). PUF domains were

obtained from pAC1355-pmax-NLSPUFa_VP64 (Addgene #71881), pAC1356-pmax-NLSPUFb_VP64 (Addgene #71882), pAC1357-pmax-NLSPUFw_VP64 (Addgene #71883) or pAC1358-pmax-NLSPUFc_VP64 (Addgene #71884) [179] using PUF_pcDNA3.1_Gib_f and _r (Supplementary Data **Table 0-9**).

2.2.3 Proof of concept and optimisation

CRISPR/Cas9 is an established two-component system (gRNA and dCas9-effector) for reliable recruitment of molecules to genomic sequences. EPIC'RISPR, however, consists not only of three components (gRNA, dCas9, effector) but also utilizes a new gRNA scaffold with one or even multiple aptamers. These aptamers, by nature, form secondary structures, and the correct folding of the gRNA is essential for efficient assembly of the complex. To verify the functionality of EPIC'RISPR, I equipped an MS2-tagged gRNA with the spacer specific for the repetitive sequences in the telomere region in human cells (TELS-gRNA-MS2, Supplementary Data **Table 0-1** and **Table 0-3**). I then transfected 750 ng of this gRNA together with 150 ng dCas9 and 20 ng MCP-coupled blue fluorescent protein (BFP) into cultured HEK293 cells, as this ratio was suggested by similar studies [180,251]. One day post transfection (dpt) I monitored the fluorescence using laser scanning microscopy (LSM). As indicated by the blue spots in **Figure 2-9** the gRNA was successful in recruiting the MCP-3xBFP to the telomeres of the cells. In comparison, untargeted N22p-3xRFP (red fluorescent protein) showed an equal distribution in the nucleus. This not only shows that EPIC'RISPR can efficiently target specific sequences and recruit AB domain-coupled molecules to the respective locus of interest, but it also demonstrates that effectors utilizing other aptamer binders are not affected, therefore opening the possibility for multiplexing approaches.

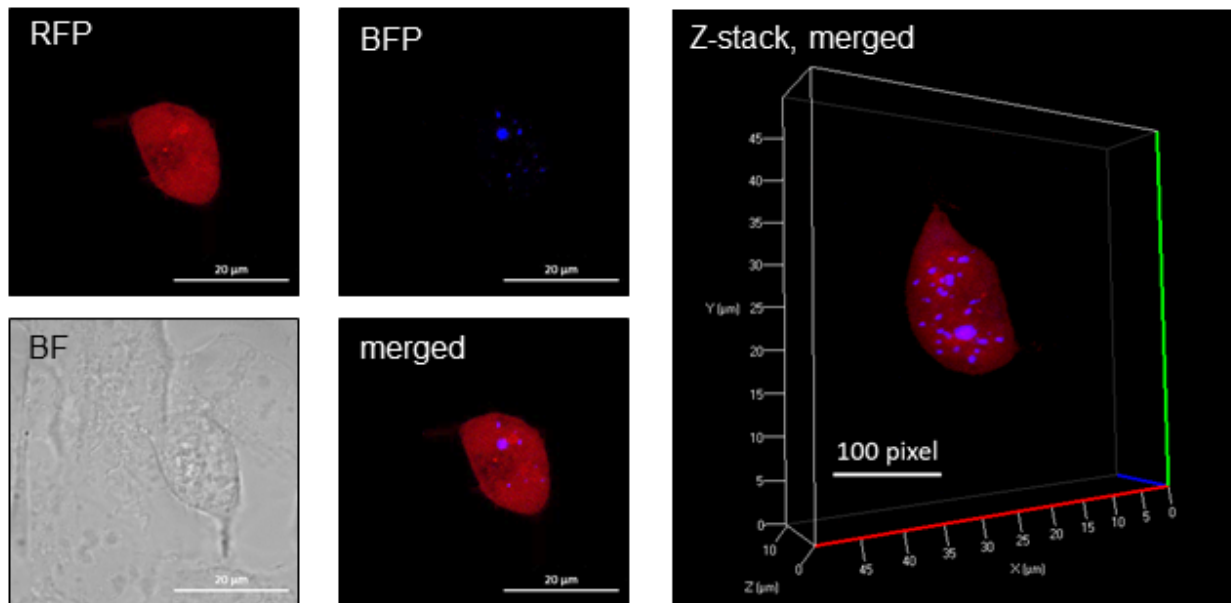


Figure 2-9: The EPIC'RISPR platform efficiently labels specific genetic loci by fluorophore recruitment. LSM generated pictures of EPIC'RISPR labelled telomers by MCP-3xBFP recruitment. As a control, untargeted N22p-3xRFP was transfected. Specifics: LSM 710, AxioObserver Plan-Apochromat 63x/1.40 Oil DIC M27, N22p-3xRFP: 576-748 nm, mRFP1.2, Ex561, Em662, MCP-3xBFP: 410-474 nm, EBFP, Ex405, Em442.

Having established the basic functionality of the system for chromatin labelling, I wanted to apply EPIC'RISPR for epigenetic editing. I coupled the synthetic transcriptional activator VPR to MCP and recruited it to the *ASCL1* promoter region. I chose *ASCL1* due to its low base expression level, making it an optimal candidate for activation experiments, as already a weak activation effect should increase the gene expression to detectable levels.

For this experiment, I used an equimolar mixture of five *ASCL1*-specific gRNA-MS2 molecules (Supplementary Data **Table 0-1**) and MCP-VPR. The gRNA spacer sequences were designed and validated by my colleague Peter Stepper. I transfected the constructs into HEK293 cells, together with dCas9 in two different vector backbones (pHAGE, dSPn) and two different ratios (r1, r2). 3 dpt, the cells were harvested, the total RNA was extracted and reverse-transcribed into cDNA. Using RT-qPCR, the transcriptional state of *ASCL1* in each sample was determined (**Figure 2-10A**). The strongest activation effect (191 ± 18 -fold) was achieved with a pHAGE vector backbone at a ratio of 1:1:5 (r2, gRNA:dCas9:effector), which is comparable to the established two-component system using dCas9-VPR and an untagged gRNA (104 ± 4 -fold). Similar results can be achieved with dSPn-based dCas9 (117 ± 5 -fold).

Using a ratio of 5:1:1 (r1, gRNA:dCas9:effector) resulted in a significantly lower transcriptional activation effect. Furthermore, MS2-tagged gRNA used for recruiting a dCas9-VPR fusion protein accomplished comparable effects to the untagged gRNA (81±4-fold).

To determine if these conditions are also optimal for transcriptional repression, I recruited MCP-coupled KRAB to the *CXCR4* promoter in the same manner (**Figure 2-10B**) using two *CXCR4*-specific spacer sequences which were also designed and validated by Peter Stepper (Supplementary Table **Table 0-1**). *CXCR4* is an actively transcribed gene in HEK293 and therefore a good target for synthetic repression. Significant reduction of transcriptional activity was achieved with EPIC'RISPR constructs transfected in r2 (pHAGE: 0.41±0.03-fold, dSPn: 0.57±0.02-fold) and with the direct fusion construct (0.30±0.01-fold). However, there was a difference between pHAGE and dSPn vector backbone, with the former yielding a significantly stronger repressive effect. Hence, in both experimental set-ups, for activation and repression, the same conditions generated the strongest effect.

Besides MS2/MCP, multiple other aptamer/aptamer binding domain combinations have been established as efficient tools in CRISPR-based systems, such as PP7/PCP or boxB/N22p. Furthermore, the PBS/PUF system has been utilized for epigenetic editing as well [179]. Therefore, I tested these combinations by transcriptionally activate *ASCL1* with coupled VPR (**Figure 2-10C**). Besides boxB/N22p all combinations achieved a significant relative activation, ranging from 42±3-fold to 422±13-fold, with MS2/MCP being the most effective combination in this set-up. This experiment was performed as described for **Figure 2-10A** with the additional transfection of 5% puromycin resistance plasmid. During the culturing, I added 2.5 µg/ml puromycin (InvivoGen) to the culture medium for a selection of transfected cells. The puromycin selection increased the achieved effect of MCP-VPR by 2.2-fold to 422±13-fold relative activation.

So far, I activated a gene with a low basic expression level (*ASCL1*) and repressed a gene that was actively transcribed (*CXCR4*). However, the question remained how the base expression level of a gene correlates with the respective epigenetic perturbation, being it activation or repression. I, therefore, targeted MCP-VPR and MCP-KRAB to a variety of five genes with different expression states and determined their expression level compared to the untargeted housekeeping gene receptor accessory protein 5

(*REEP5*, **Figure 2-10D**). For targeting, I utilized equimolar mixtures of target-specific spacer sequences which, as for *ASCL1* and *CXCR4* were designed and validated by Peter Stepper in his doctoral thesis (see Supplementary Data **Table 0-1**). Synthetic transcriptional activation worked best when the target has a low base expression such as *LIN28A* and *ASCL1*. In these genes, 100-fold overexpression was easily achieved with MCP-VPR, while active targets like *TFRC* only allowed a 2-fold increase in transcription. Transcriptional repression followed the inverse principle, with active genes like *TFRC* and *CXCR4* representing optimal targets for transcriptional repression. *ASCL1*, on the other hand, was not repressed at all by MCP-KRAB, showing no significant change in expression compared to mock-transfected cells.

Based on these results, I decided to perform all other editing experiments in this thesis with a ratio of 1:1:5 (gRNA:dCas9:effector) and pHAGE-based dCas9 vectors. Furthermore, I used puromycin selection, if not stated otherwise. Due to its better performance compared to other recruiters, I implemented the MS2/MCP system as the main aptamer/AB combination.

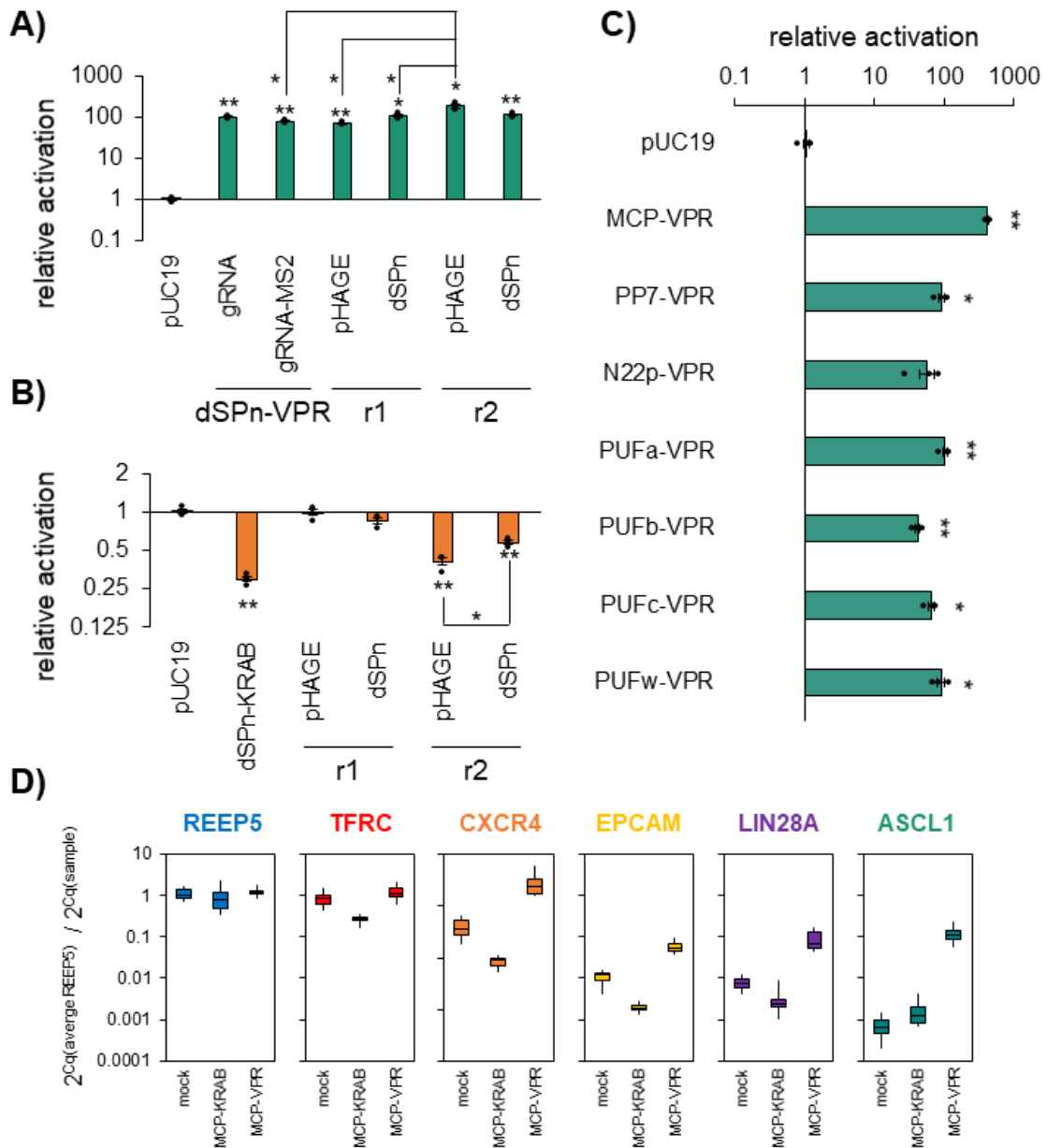


Figure 2-10: Proof of concept and optimisation of the EPIC-RISPR platform as an epigenetic editing tool. A) Optimisation of plasmid backbones and ratios (r1 and r2, gRNA:Effector:dCas9) for activation using MCP-VPR targeting the *ASCL1* promoter. **B)** Optimisation of plasmid backbones and ratios (r1 and r2, gRNA:Effector:dCas9) for repression using MCP-KRAB targeting the *CXCR4* promoter. **C)** Comparison of recruitment efficacy to the *ASCL1* promoter of different aptamers and their respective binding protein using binding protein-VPR as a reporter system for transcriptional activation. **A)** to **C)** Error bars represent SEM of at least three biological replicates with two technical replicates each. *P < 0.05, **P < 0.01, two-sided t-test. **D)** Basic expression levels of untargeted (*REEP5*) and targeted (*TFRC*, *CXCR4*, *EPCAM*, *LIN28A*, and *ASCL1*) genes after perturbation with either MCP-VPR or MCP-KRAB. The boundary of the box plots represents the 25th and 75th percentile, with the median indicated as a black line. Whiskers below and above show the 10th and 90th percentiles, respectively. n = 14 independent experiments.

2.3 EPIC'RISPR for highly parallel transcription modulation and chromatin labelling

The melding of targeting- and recruitment functionality on one molecule, the gRNA, distinguishes the EPIC'RISPR system from any dCas9-effector fusion-based set-up for epigenetic editing. Any desired combination of spacer sequences can be combined with any one or multiple effector molecules in a specific manner without interference. It gives the system incredible flexibility and capability as the simultaneous treatment of multiple target sequences with various combinations of effector molecules is easily doable. There are three basic possibilities of multiplexing approaches using the EPIC'RISPR platform (**Figure 2-11**): Firstly, the targeting of multiple target sequences with one effector molecule. A mixture of gRNAs specific for multiple loci, but all tagged with the same aptamer allow the recruitment of one type of effector to several target sites at the same time. Secondly, gRNAs tagged with different aptamers and spacer sequences can recruit different effectors to separate targets without interference. Thirdly, multiple effector molecules can be recruited to one target site. This can be achieved either by adding multiple aptamers to the gRNA and the consequential recruitment of multiple AB-fused effectors or by recruiting one multi-effector protein by one aptamer. In chapter 2.2.3, I already demonstrated that the recruitment of such a multi-effector protein is doable by targeting a fusion-construct of MCP and three consecutive fluorescent proteins (3xBFP) to the telomere region of HEK293 cells. Furthermore, VPR is a synthetic multi-domain protein that contains three transcriptional activators. In the following chapters, I will present evidence for the feasibility of all multiplexing approaches described here.

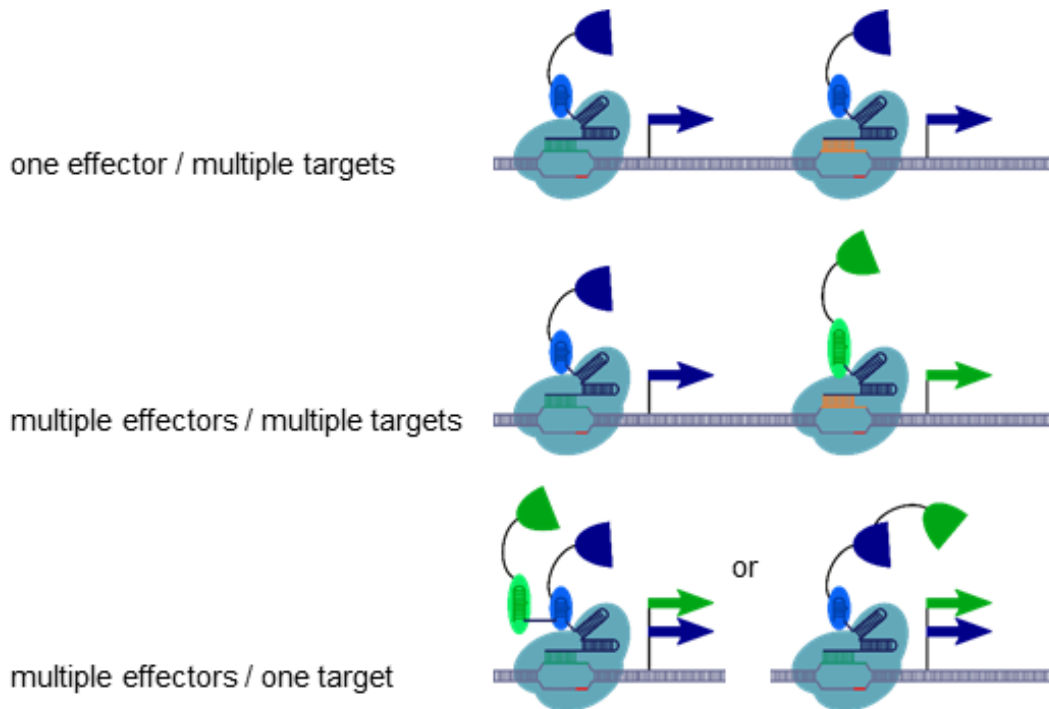


Figure 2-11: Schematic representation of possible EPIC' Multiplexing set-ups. Upper panel: simultaneous targeting of multiple target gene promoters with one effector molecule. Middle panel: simultaneous recruitment of different EDs to separate target gene promoters. Lower panel: simultaneous recruitment of multiple effector molecules to one target, either by using multiple aptamer sequences or multi-effector fusion constructs.

2.3.1 gRNA mixtures

To contain the determined optimal ratio between dCas9, AB-ED and gRNA-aptamer for multiplex experiments, the total amount of transfected gRNA should not be changed. This means, however, that a smaller amount of each individual gRNA has to be used, which might affect the efficacy of the transcriptional alteration. So far, I used equimolar mixtures of two to eight gRNAs which were designed and validated by my colleague Peter Stepper (Supplementary Data **Table 0-1**). I quantified the effects of a selection of single gRNAs in the EPIC'RISPR set-up by recruiting MCP-VPR (**Figure 2-12**). For each target, at least two gRNAs were sufficient to upregulate the transcription ($P < 0.05$) significantly. I also tested the equimolar mixtures of all gRNAs specific for one target as described in chapter 2.2.3. In the case of *ASCL1*, this mixture achieved a stronger activation of the target gene than any single gRNA, while for *LIN28A* and *CXCR4*, the effects were similar. For *TFRC* and *EPCAM*, on the other hand, at least one single gRNA yielded an effect more potent than the mixture. For

multiplexing, I selected the two to three best-performing gRNAs for each target and combined them in an equimolar mixture, which I then transfected together with dCas9 and MCP-VPR into HEK293 cells (**Figure 2-12**, m:). In all cases but *LIN28A*, the achieved efficacies were lower than when using a mixture of all available gRNAs for one target. For *LIN28A* the multiplexing set-up achieved only 43% of the effect of the single target activation experiment, while *ASCL1* still yielded 94%. Interestingly, the activation efficacies between multiplexing experiments and single gRNA set-ups varied depending on the target gene. For *ASCL1* the mixture of the three best performing gRNA outperformed the strongest single gRNA by ~2.8-fold, while *LIN28A* and *TFRC* achieved only ~0.5-fold of the most efficient single gRNA. For *EPCAM* and *CXCR4*, however, the multiplexing mixtures activated the target genes similar to the most efficient single gRNA.

In conclusion, the equimolar mixture I assembled, containing two to three gRNAs per target was sufficient to upregulate all five target genes significantly. In the following chapters (2.3.3 to 2.3.5), I hence used equimolar mixtures of the two to three most effective gRNAs specific for the desired targets. For the experiments described in chapters 2.3.2 and 2.5, I utilized the same gRNA mixture for all five genes as for the experiment described above (**Figure 2-12**).

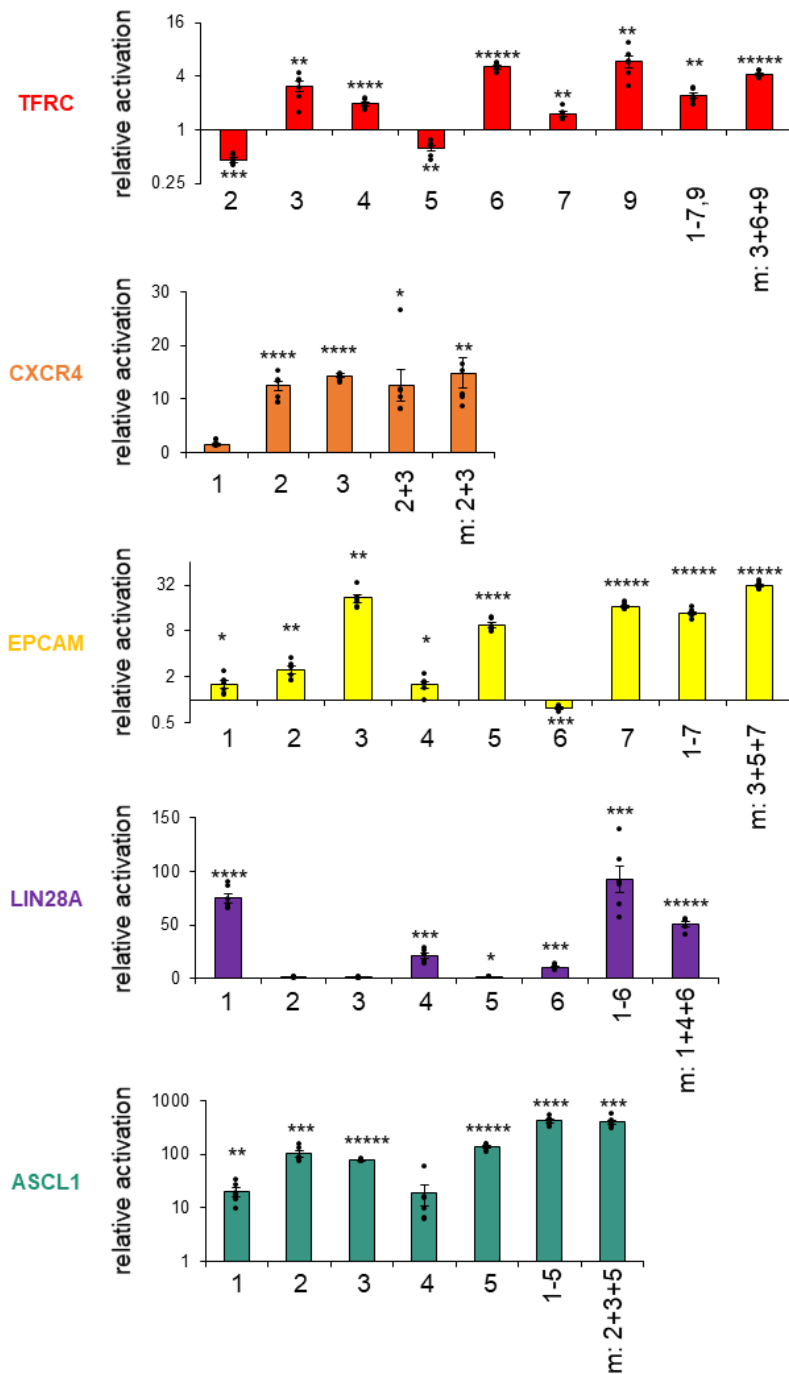


Figure 2-12: Validation of target-specific MS2-modified gRNAs using MCP-VPR. The spacer sequences were initially designed and validated by Peter Stepper during his doctoral thesis (Supplementary Data **Table 0-1**). I chose several gRNAs (numbers) for each target that performed well for Mr Stepper and tested them individually as well as in an equimolar mixture of all. I then performed a multiplex experiment (m:) with an equimolar mixture containing the three best gRNAs of each target (two for *CXCR4*) and compared the transcriptional activation efficacies. Error bars represent SEM of three biological replicates with two technical replicates each. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001 *****P < 0.00001, two-sided t-test.

2.3.2 one effector / multiple targets

For recruiting one ED to multiple genetic loci, I determined the three most effective gRNAs for each target (2 for *CXCR4*, see chapter 2.3.1) and prepared an equimolar mixture containing the MS2-coupled gRNAs for each target. In then transfected these gRNAs together withdCas9 and MCP-VPR or -KRAB, respectively, and analysed the relative transcriptional activation 3 dpt (**Figure 2-13**).

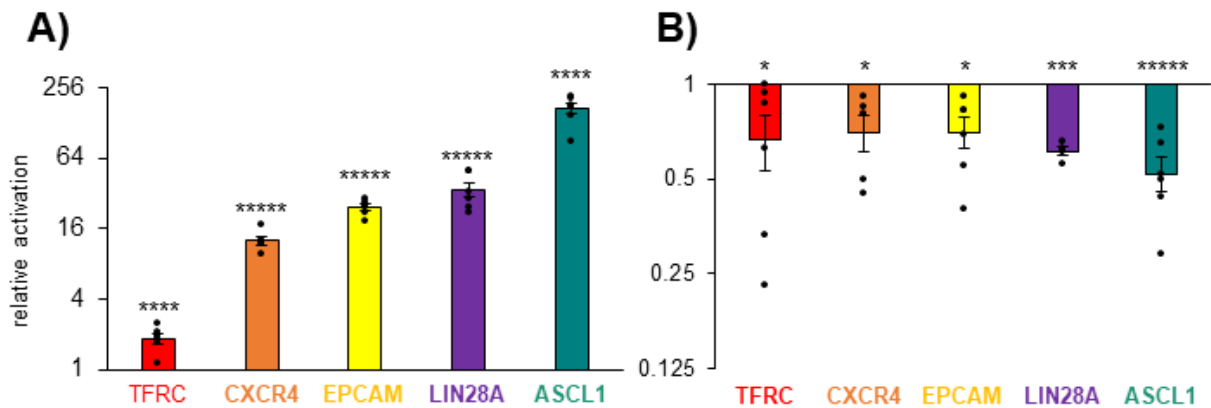


Figure 2-13: EPIC multiplexing allows the recruitment of one effector molecule to five target gene promoters. A) recruitment of MCP-VPR via MS2-tagged gRNAs. **B)** recruitment of MCP-KRAB via MS2-tagged gRNAs. Error bars represent SEM of four to six biological replicates with two technical replicates each. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, ***** $P < 0.00001$, two-sided t-test.

Both transcriptional effector molecules influenced the transcription of all five target genes significantly, demonstrating the effective multiplexing capabilities of EPIC'RISPR. Similar to the effects seen before (**Figure 2-10D**), the achieved relative activational effect of MCP-VPR strongly was strongly dependent on the target gene's basic expression level. Active genes showed a lower response to the effector than repressed genes. However, the achieved overall effects are slightly lower compared to non-multiplexing experiments (e.g. *ASCL1* with 169 ± 19 -fold); an effect that can be explained by the lower portion of gRNA specific for each respective gene.

The repression experiment with MCP-KRAB showed a higher deviation and lower changes in the transcriptional state as well (e.g. *CXCR4* with 0.7 ± 0.1 -fold) compared to the non-multiplexing experiment (0.41 ± 0.03 -fold, **Figure 2-10B**). An exception depicts *ASCL1*, which in this set-up could be effectively repressed by MCP-KRAB (0.52 ± 0.06 -fold). Due to its low base expression level, repression experiments on *ASCL1* should be interpreted very carefully though.

These results emphasise the extreme versatility of the EPIC'RISPR platform and the possibilities of utilizing it in high-throughput experiments. Together with Sven Höhn and Luca Schelle, I used this multiplexing set-up to recruit more than 60 human EDs to these five genes and monitored the induced changes in the transcriptional state of the targets. The results of this large-scale experiment are presented in detail in chapter 2.5.

2.3.3 multiple effectors / multiple targets

After demonstrating that the targeting of multiple genomic loci with EPIC'RSIRP is highly effective and yields significant changes in expression levels, I planned to extend the application of multiplexing by using differently tagged gRNAs for different loci. This allowed me to specifically recruit different AB-ED fusion proteins in parallel to a variety of genetic loci.

To obtain a visual proof of concept, I targeted the repetitive sequences in the telomere regions using a PP7-tagged gRNA (TELS-gRNA-PP7, Supplementary Data **Table 0-1**) as well as a gRNA tagged with MS2 that contained a spacer specific for the repetitive sequences in chromosome 9 (Chr9-gRNA-MS2, Supplementary Data **Table 0-1**). I utilized these gRNAs to recruit PCP-3xGFP (green fluorescent protein) and MCP-3xBFP to the respective loci, without any co-localisation of the signal as indicated by the distinct green and blue dots in **Figure 2-14A**. Untargeted N22p-3xRFP served as a negative control and showed an equal distribution throughout the cells.

I then implemented this set-up for the epigenetic editing of three different genes simultaneously. For this I directed the EPIC'RISPR modulator constructs MCP-KRAB, PCP-VPR and PUFa-VPR to three different target genes (*CXCR4*, *LIN28A*, *ASCL1*) by using target-specific gRNAs modified with fitting binding sites (MS2, PP7, PBSa). The gRNAs were combined in an equimolar mixture containing the two to three best spacer sequences for each target as described in chapter 2.3.1. I observed simultaneous and significant repression of *CXCR4* and activation of *LIN28A* and *ASCL1* when I co-transfected the named transcriptional modulators and gRNAs with pHAGE-dCas9 (**Figure 2-14B**). To verify that no off-target effects occur when co-transfecting e.g. PCP-VPR and PUFa-VPR, I also transfected samples that lacked one of the gRNA sets specific for either target. As expected, these samples did not show any significant change in transcriptional activity compared to either mock-transfected

cells or samples transfected without any ED. More importantly, EDs with a non-fitting aptamer binding domain did not influence the transcriptional state of the respective off-target. This shows that the EPIC'RISPR system can easily change the transcriptional states of multiple genetic loci in different ways without interference, emphasizing its superior modularity and flexibility over set-ups based on dCas9-effector fusion proteins.

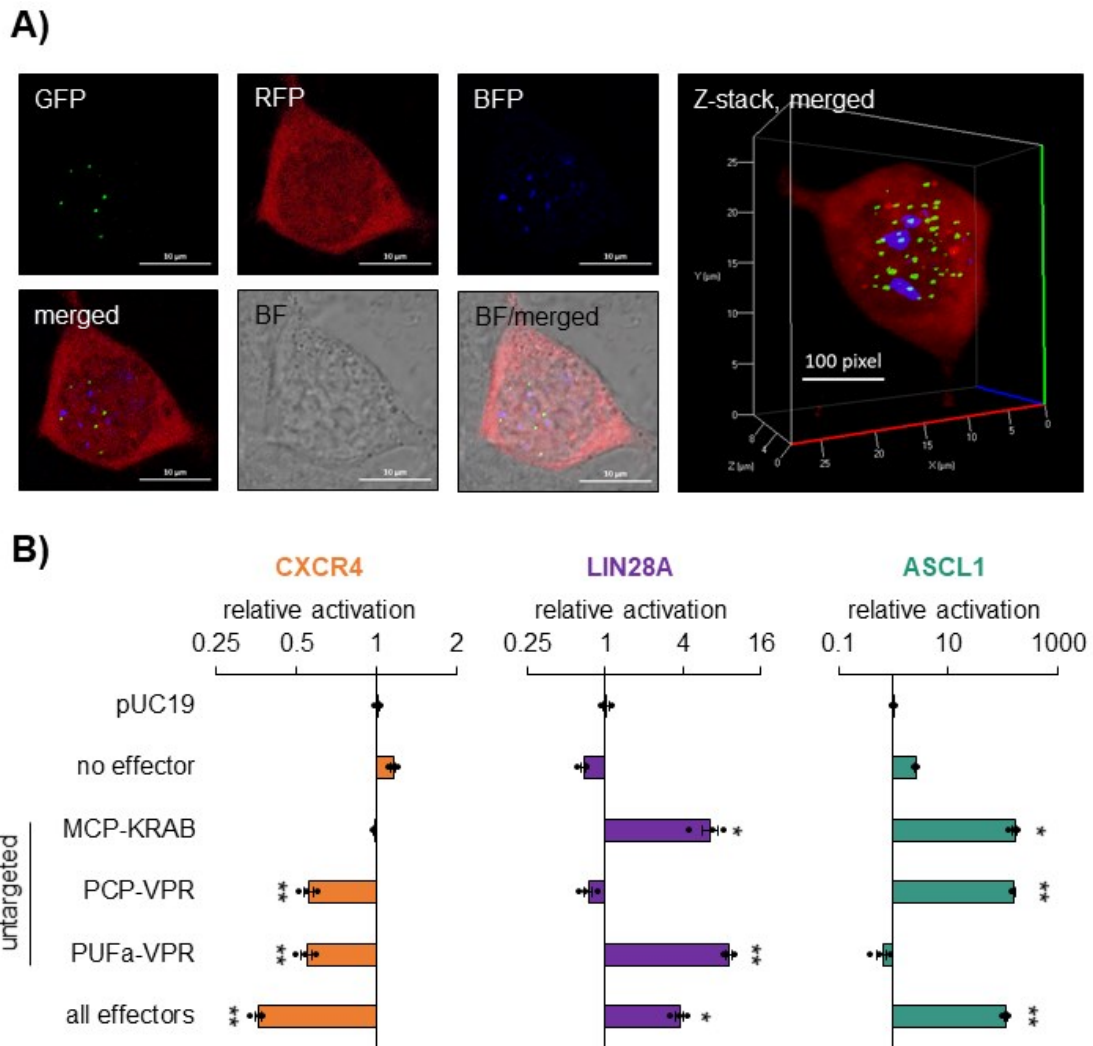


Figure 2-14: EPIC'multiplexing allows the simultaneous recruitment of multiple effectors to multiple, different target loci without interference. A) LSM generated pictures of simultaneous labelling of telomeres (PCP-3xGFP) and chromosome 9 specific repeats (MCP-3xBFP). As a control, untargeted N22p-3xRFP was transfected. Specifics: LSM 710, AxioObserver Plan-Apochromat 63x/1.40 Oil DIC M27, N22p-3xRFP: 576-748 nm, mRFP1.2, Ex561, Em662, MCP-3xBFP: 410-474 nm, EBFP, Ex405, Em442, PCP-3xGFP: 496-553 nm, EGFP, Ex488, Em525. **B)** Simultaneous activation and repression of three target genes (*CXCR4*, *LIN28A*, *ASCL1*) by effector recruitment using three different aptamer/binding protein combinations. To determine the occurrence of off-target effects samples with untargeted effectors (no gRNA-aptamer) were analysed. Error bars represent SEM of at least three biological replicates with two technical replicates each. *P < 0.05, **P < 0.01, two-sided t-test.

2.3.4 multiple effector-copies / one target

It has been shown for an epitope (SunTag)-modified dCas9, that a signal amplification can be achieved in epigenetic editing by recruiting multiple copies of scFv-VP64 to one target gene. However, aptamer-based systems so far failed to deliver similar results. There is a study showing that a 2x-aptamer system can be stable and achieve increased effects, but the respective construct has to be stabilized by a black folding sequence, limiting the maximal effects of the application. To determine the potential of signal amplification using multiple aptamers for the EPIC'RISPR system, I designed a gRNA modified with five consecutive MS2 aptamers at the 3' end (**Figure 2-15A**).

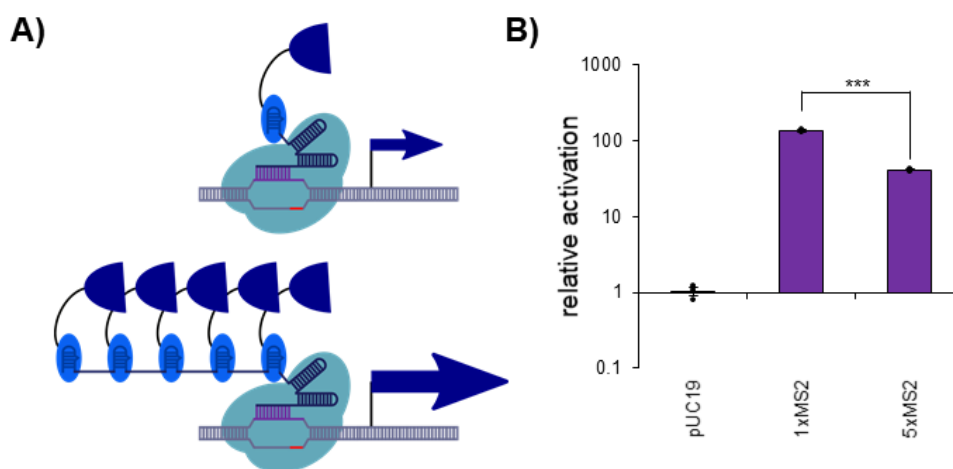


Figure 2-15: Signal amplification with EPIC'RISPR. A) Schematic representation of signal amplification by recruiting five copies of MCP-VPR to the *LIN28A* target gene. B) Relative activation of *LIN28A* by MCP-VPR. The ED was recruited by a gRNA modified with either one or five copies of the MS2 stem-loop. Error bars represent SEM of three biological replicates with three technical replicates each. *** $P < 0.001$, two-sided t-test.

I utilized this gRNA to recruit multiple copies of MCP-VPR to the *LIN28A* target gene and achieved a significant upregulation of the gene by 41 ± 1 -fold ($P < 0.001$, **Figure 2-15B**). However, the standard EPIC'RISPR gRNA containing only one stem-loop achieved 135 ± 4 -fold activation of *LIN28A*, which is significantly higher ($P < 0.001$).

2.3.5 multiple effectors / one target

VPR represents a good example that the simultaneous recruitment of multiple EDs to one promoter can have combinatorial or additive effects on the target gene's transcriptional state and epigenetic environment. In the case of VPR, these synergistic effects increase the immediate overall efficacy of transcriptional modulation.

Unpublished findings from other members of our group suggest that co-recruitment of other EDs can also have effects on the stability of the altered transcriptional status. These concepts have also been investigated by others in the past [166,168,172]. This concept represents the foundation for any immediate and long-term change in the epigenetic code, and therefore for the lasting reprogramming of biochemical processes or even differentiation states. EPIC'RISPR is an optimal tool for implementing such strategies in a large-scale format. Two possible strategies of approaching this come to mind: the recruitment of different AB-EDs via a multi-aptamer tagged gRNA or the recruitment of a multi-effector fusion construct with one aptamer/binding domain combination. While the latter has already been demonstrated by recruiting VPR as well as three consecutive fluorophores, co-recruitment with different aptamers in one gRNA was not yet validated. The general idea of a multi-aptamer set-up harbours the big advantage of increased modularity, as any aptamer binding domain-effector fusion construct of the basic EPIC'RISPR toolbox can be utilized and conveniently be interchanged to test combinations in large-scale experiments. To verify the functionality of such a multi-aptamer-tagged gRNA, I designed and cloned three different variants that contained one, two or three consecutive copies of the following aptamer cassette: PP7-MS2-boxB (**Figure 2-16**). I equipped these gRNAs with a spacer sequence for the chromosome 9-specific repeats and co-transfected them together with PCP-3xGFP, MCP-3xBFP, N22p-3xRFP and pHAGE-dCas9. In each case, a co-localization of all three fluorophores could be observed as displayed by distinct white dots in the merged pictures (**Figure 2-17**).

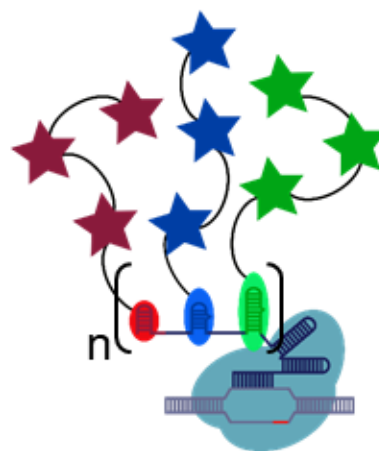


Figure 2-16: Schematic representation of a multi-aptamer gRNA recruiting three different binding domain-coupled fluorophores to a target locus. The cassette of PP7, MS2 and boxB can be introduced once, twice or three times ($n = 1, 2$ or 3) to recruit different amounts of fluorophores

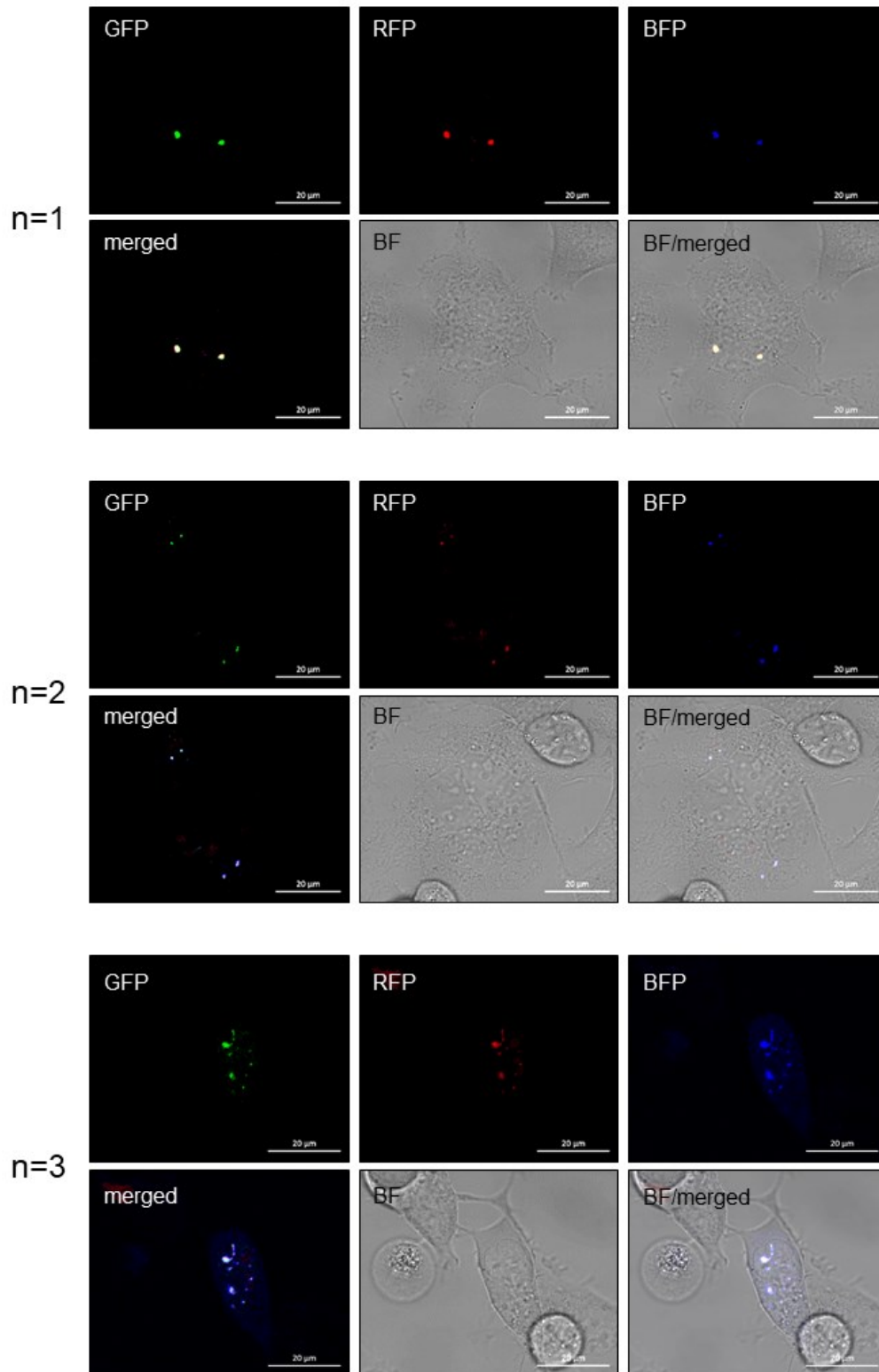


Figure 2-17: EPIC' multiplexing allows the recruitment of three different fluorophores to one target sequence. LSM generated pictures of EPIC'RISPR labelled of chromosome 9-specific repetitive sequences using gRNA-(PP7-MS2-boxB)_n, PCP-3xGFP, MCP-3xBFP and N22p-3xRFP, with n being 1, 2 or 3. Specifics: LSM 710, AxioObserver Plan-Apochromat 63x/1.40 Oil DIC M27, N22p-3xRFP: 576-748 nm, mRFP1.2, Ex561, Em662, MCP-3xBFP: 410-474 nm, EBFP, Ex405, Em442, PCP-3xGFP: 496-553 nm, EGFP, Ex488, Em525

I then tried to implement this system for epigenetic editing and co-targeted VPR and mTET3del1 to one target gene promoter. For this experiment, I decided to choose *EPCAM* as the promoter is heavily methylated in HEK293 cells and offers a good target for mTET3del1 mediated demethylation. In contrast to other experiments performed in this thesis, I decided to analyse samples 5 dpt, instead of 3 dpt, as recent findings of other experiments in our group suggest stronger effects in transcriptional upregulation by target promoter demethylation at later time points. I co-targeted *EPCAM* with a gRNA harbouring the MS2 stem-loop as well as PBSa (gRNA-MS2-PBSa), MCP-VPR and PUFa-mTET3del1 and achieved significant upregulation by 3.9 ± 0.6 -fold 5 dpt ($P < 0.05$, **Figure 2-18**). However, this effect is not significantly different when recruiting only VPR with either gRNA-MS2 (3.3 ± 0.4 -fold) or gRNA-MS2-PBSa (2.8 ± 0.1 -fold). Recruiting mTET3del1 alone with either gRNA also resulted in upregulation of *EPCAM*, although to a lower extent (2.0 ± 0.3 -fold and 1.6 ± 0.1 -fold).

The possibility of targeting multiple EDs to a single target suggests a powerful functionality for the EPIC'RSIPR platform, as this ability would revolutionize the way we approach epigenetics. I was able to successfully recruit multiple fluorophores to a single genetic locus, either by using numerous aptamers or by multi-fluorophore fusion constructs. However, more research has to be done to effectively utilize this system for co-recruitment of multiple EDs to one target.

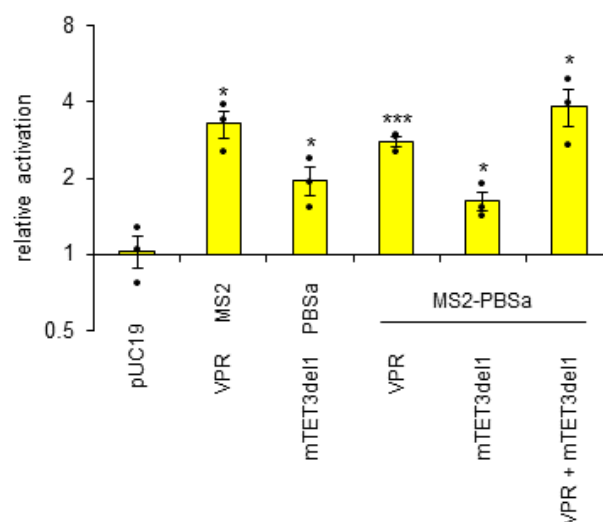


Figure 2-18: Simultaneous recruitment of two different EDs to the *EPCAM* target gene promoter. MCP-VPR and PUFa-mTET3del1 were first recruited individually using single aptamer gRNAs (gRNA-MS2, gRNA-PBSa) or gRNA containing both binding sites (gRNA-MS2-PBSa). Finally, both EDs were recruited simultaneously using gRNA-MS2-PBSa. Error bars represent SEM of three biological replicates with two technical replicates each. * $P < 0.05$, *** $P < 0.001$, two-sided t-test.

2.4 EPIC'RISPR for inducible and tuneable transcription modulation

Xi, imprinting or cell differentiation are all examples of naturally occurring regulation of gene expression in the mammalian organism [252–254]. To study these natural processes of epigenetic modulation, it is crucial to investigate the influence of different levels of mRNA on the respective biochemical pathway. One way to do so is the synthetic alteration of the transcriptional state of the respective gene. CRISPR/Cas9-based epigenetic editing methods such as the EPIC'RISPR platform show great potential as tools for studies like this. Yet, most epigenetic EDs typically used in these systems are selected for their strong effect on the transcriptional state of the target gene. By recruiting e.g. VPR, the EPIC'RISPR system can increase the transcriptional activity of a gene by several hundred-folds. However, this does not represent a natural change in the expression of mRNA in a human cell. To overcome this issue, I modified the EPIC'RISPR platform with an ON- as well as an OFF-switch, which established the functionalities of inducibility and tunability. These switches allowed me to introduce defined shifts in gene expression in naturally occurring dimensions.

2.4.1 Switch-ON

The results described in this subchapter were obtained together with Corinna Kersten, a student I mentored during her Bachelor thesis.

As a candidate for the switch-ON functionality, we selected the dimerization pair FKBP and FRB [255]. The two proteins form heterodimers in the presence of the chemical rapamycin or one of its analogues. We hypothesised that such an induced dimerization process could be utilized for tuneable epigenetic editing with the EPIC'RISPR platform. By fusing one of the two molecules to an epigenetic effector and the respective partner to e.g. MCP, this complex could assemble in the presence of rapamycin, recruiting the effector to a gene of interest by an MS2-tagged gRNA (**Figure 2-19A**) [200]. To find the best sterical positioning of the molecules, we designed several constructs with permuted orientations of FKBP and FRB (**Figure 2-19B**).

When recruited to the promoter region of *ASCL1* as described above (see chapter 2.2.3), an equimolar combination of MCP-2xFKBP and FRB-VPR achieved a relative activation of 34 ± 1.7 -fold activation compared to pUC19 transfected cells in the presence of 100 nM a rapamycin analogue (A/C Heterodimerizer, TaKaRa) (**Figure 2-19C**). This is significantly higher than any other FKBP/FRB-combination tested in

this experiment and also higher than the same construct accomplished at 5 nM rapamycin (3.6 ± 0.2 -fold).

Based on these results, we wanted to determine the optimal concentration of A/C Heterodimerizer for this combination of the two dimerization partners. We hence exposed the transfected cells to different concentrations of rapamycin, ranging from 25 to 2000 nM (**Figure 2-19D**). While at all concentrations, a significant increase of transcriptional activity was observed, 250 nM rapamycin yielded the highest activation effect of 66 ± 9 -fold. This corresponds to $3.2\pm 0.5\%$ of the effect achieved by MCP-VPR. Our goal, however, was to improve the system to reach a maximum activation effect closer to MCP-VPR to cover the whole bandwidth of the effector's capabilities. As VPR is a large molecule, we hypothesized that its size might sterically hinder the assembly of the complex. To test this theory, we performed the same experiment using the smaller subunit VP64 as an effector (**Figure 2-19E**). We achieved significant activation effects from 50 to 250 nM in a dose-dependent manner, while observing the strongest effects at 150 nM with a relative activation of 5.1 ± 0.2 -fold. This corresponds to $49\pm 1\%$ of the effects achieved by the basic EPIC'RISPR construct MCP-VP64 (10.3 ± 1.5 -fold).

Here we developed a rapamycin-induced FKBP/FRB-based ON-switch for the EPIC'RISPR platform that allows the dose-dependent modulation of transcriptional activity by the simple addition of a chemical to the cell culture. We show that multiple orientations of the dimerization partners can be utilized with different efficiencies. The achieved effects are significant compared to pUC19-transfected cells as well as to cells transfected with only the dCas9/gRNA complex but without an effector. We further demonstrate that the achieved results compared to the direct MCP-effector differ depending on the used effector molecule and, in our most successful set-up, covers about 50% of the effect-range of the uninduced system.

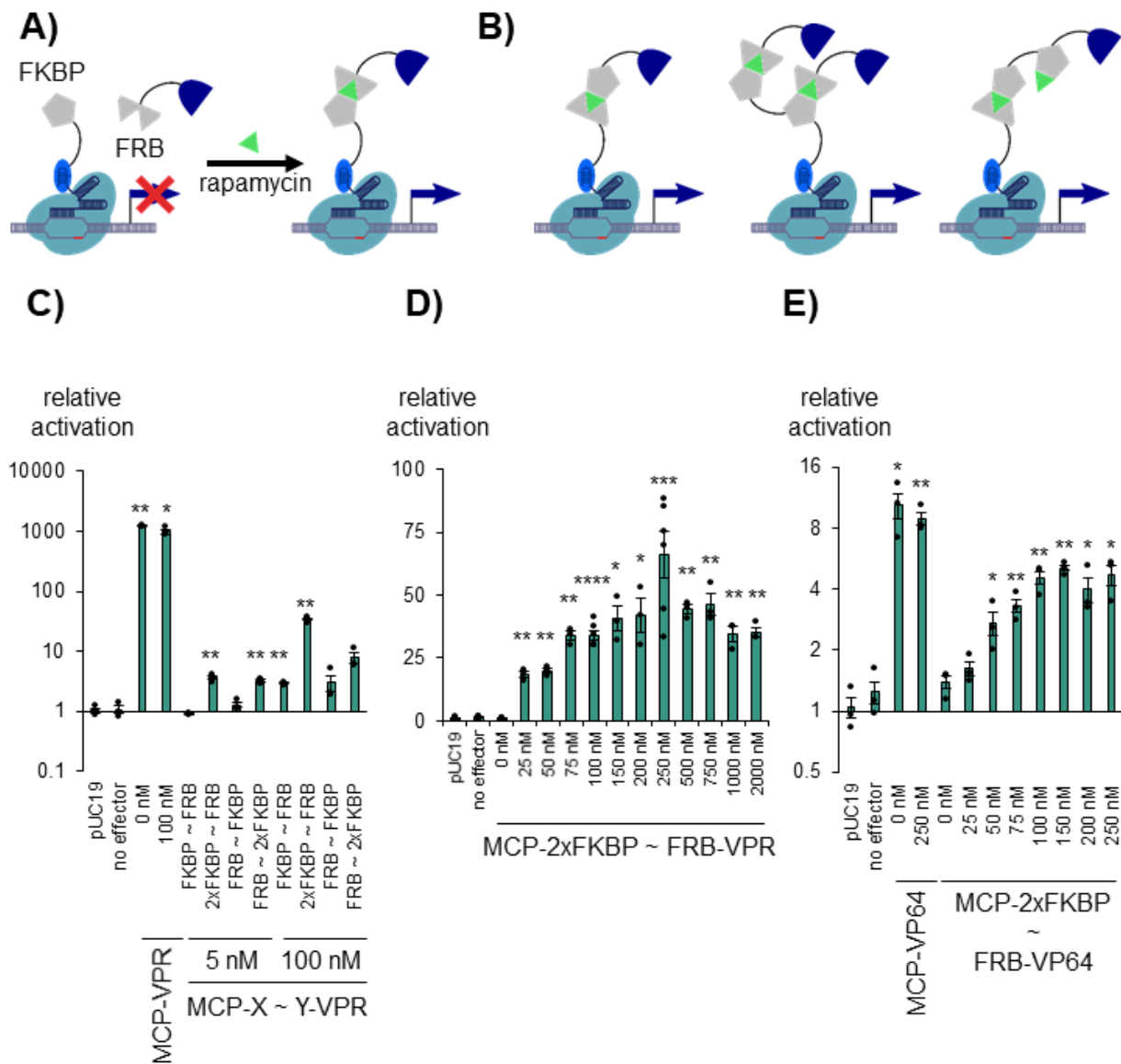


Figure 2-19: Tunable synthetic gene activation by rapamycin-induced recruitment of transcriptional activators. **A) and B)** Schematic representation of rapamycin-induced dimerization of FKBP and FRB (in multiple orientations) and the consequential recruitment of a transcriptional activator to a gene of interest. **C)** Relative activation of *ASCL1* by rapamycin-induced recruitment of the transcriptional activator VPR at rapamycin concentrations of 5 and 100 nM. **D)** Concentration-dependent increase of the achieved activational effects of MCP-2xFKBP~FRB-VPR compared to MCP-VPR. **E)** Concentration-dependent increase of the achieved activational effects of MCP-2xFKBP~FRB-VP64 compared to MCP-VP64. **C), D) and E)** Error bars denote the SEM from at least three biological repeats performed in technical triplicates. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, two-sided t-test.

2.4.2 Switch-OFF

Although the FKBP/FRB-system described in chapter 2.4.1 can reach up to $49\pm 1\%$ of the effect of an MCP-effector fusion construct, we wanted to cover the whole bandwidth of transcriptional alteration from 0 to 100%. To do so, I approached this challenge from a different perspective and developed an OFF-switch functionality. I employed another member of the Cas protein family: Csy4. As described in detail in chapter 1.2.3, Csy4 binds a small RNA stem-loop and exhibits an endonuclease activity. While this activity is inhibited by a point mutation of the H at position 29 to A, it can be recovered *in vitro* by adding imidazole to the reaction [210]. I utilized this effect and introduced the Csy4 stem-loop as a functional tag (CsyTag) into the EPIC'RISPR gRNA (**Figure 2-20A**, Supplementary Data **Table 0-3**). In the absence of imidazole, the system can recruit an ED to the gene of interest where it exhibits its function. After the addition of the chemical to the cell culture, the Csy4 mutant's endonucleolytic activity is revived, the MS2 stem-loop is released from the gRNA, prohibiting the recruitment of the effector and, consequently, stopping any further editing. Before I started the experiments described in this chapter, I tested the tolerance of HEK293 cells for imidazole. The detailed results of this test are described in chapter 2.6.

I first confirmed the system's feasibility by recruiting VPR to the promoter of *ASCL1* (**Figure 2-20B**). At 0 mM imidazole, a gRNA containing the CsyTag as well as an MS2 stem-loop (gRNA-CsyTag-MS2, Supplementary Data **Table 0-3**) can increase the transcriptional activity of the target gene by 526 ± 66 -fold. The addition of 32 mM imidazole to the cell culture reduced this effect to 26 ± 2 -fold, which represents a drop to $5.0\pm 0.4\%$ compared to the samples without imidazole. Other already published and unpublished CRISPR/Cas-based editing constructs achieved effects comparable to the set-up at 0 mM imidazole. While the direct fusion construct of dCas9-VPR achieved 593 ± 71 -fold activation, the EPIC'RISPR system increased the transcriptional activity of *ASCL1* by 321 ± 44 -fold (gRNA-CsyTag-MS2, w/o Csy4) and 337 ± 65 -fold (gRNA-MS2). Furthermore, the addition of wildtype Csy4 (Csy4 WT, Supplementary Data **Table 0-11**) reduced the activational effect of the system to 2.6 ± 1.8 -fold, which is not significantly higher than pUC19 transfected cells or cells transfected without MCP-VPR. These findings allow three conclusions: Firstly, the CsyTag does not impact the efficiency of the recruitment of MCP by the MS2 stem-loop, compared to the basic EPIC'RISPR system or a dCas9-effector fusion construct. Secondly, it shows that the

endonucleolytic activity of Csy4 H29A can be rescued *in vivo* by the simple addition of imidazole to the culture medium. Thirdly, the rescued activity of Csy4 H29A can reduce the exhibited transcriptional modulation of MCP-VPR by about 95%, which covers almost all the effect-range of the system.

Besides Csy4 H29A, I also tested the performance of an H29G mutant in this set-up, as H to G mutants have shown stronger rescue effects in HNH endonucleases in the past [256]. I hypothesized that such a mutant might be able to cover the remaining 5% compared to Csy4 WT or pUC19-transfected cells, respectively. For this, I repeated the experiment with Csy4 WT, Csy H29A and Csy4 H29G and quantified the total effect of the Csy4 cleavage on the transcriptional activity by normalising the transcriptional alteration effect on the effect achieved by the respective sample at 0 mM imidazole (**Figure 2-20C**). The strongest reduction was still achieved by Csy4 WT (91 ± 10 -fold), followed by Csy H29A (21 ± 0.6 -fold) and Csy H29G (10 ± 1.8 -fold). For all further experiments, I thus chose the H29A mutant over Csy4 H29G due to its stronger and more consistent effect.

Following up on these results, I tested the effect of different imidazole concentrations on the performance of the system (**Figure 2-20D**). I hypothesized that with lower concentrations of imidazole, the rescue effect should be reduced, allowing the system to exhibit a transcriptional activation effect closer to the completely uninduced state. The aim was to establish a mathematical model which can be used to calculate an exact concentration of imidazole necessary to achieve a specific, arbitrary alteration effect. In this experiment, I titrated the concentration of imidazole added to the culture medium from 0.5 to 32 mM. 1 mM imidazole was enough to induce a significant drop in transcriptional activity compared to untreated cells (241 ± 17 -fold). As predicted, the effects increase in a dose-dependent manner (26.4 ± 1.9 -fold at 32 mM), showing a strong negative correlation between relative activation and concentration of imidazole ($R^2=0.998$, **Figure 2-20E**). This permits the fitting of an exponential curve and the extraction of a formula for the calculation of transcriptional activation effect dependent on the concentration of imidazole.

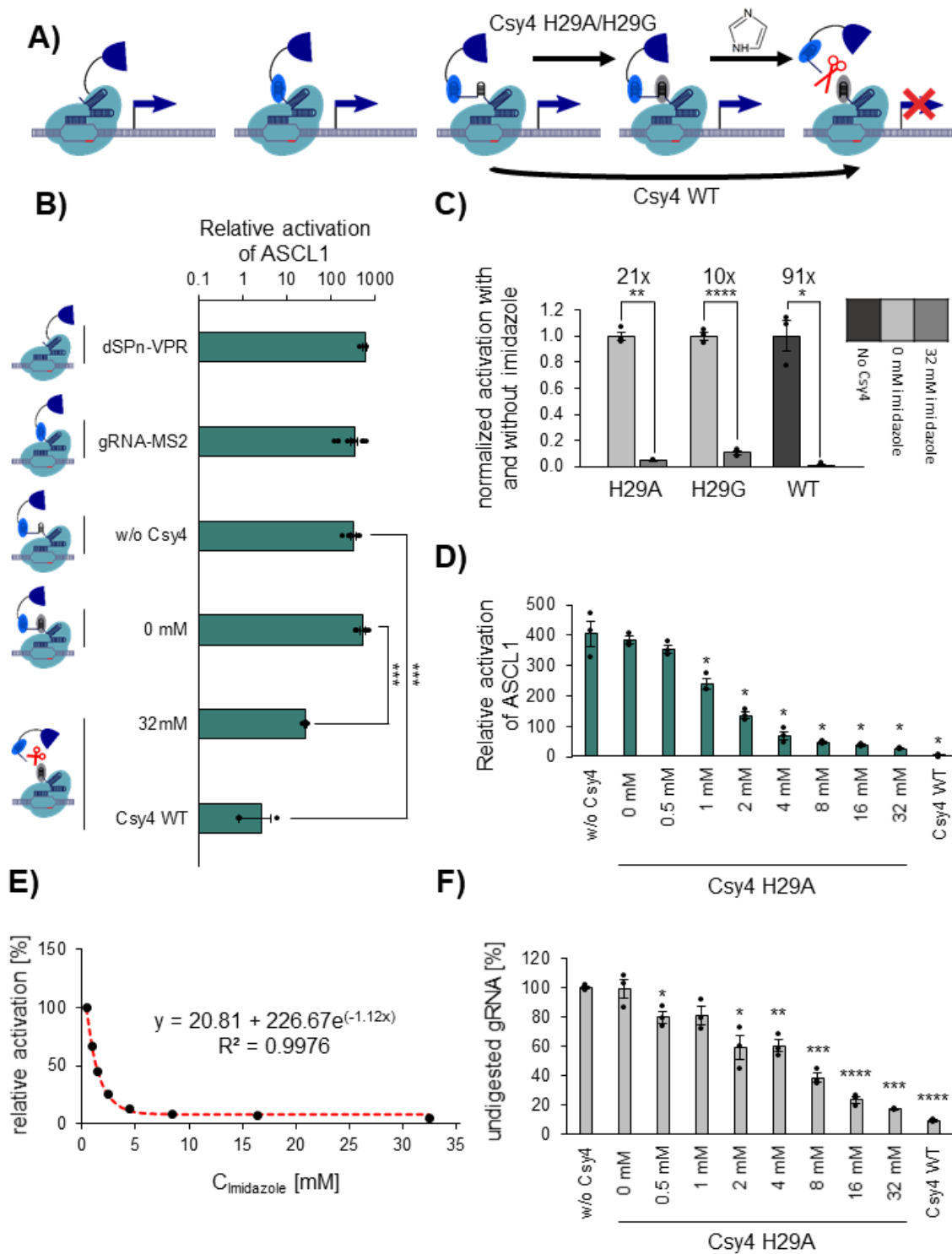


Figure 2-20: Tunable synthetic gene activation by imidazole-induced cleavage of gRNAs. Figures and figure text were taken from Oberacker et al. ([211] submitted) and have been slightly modified. **A)** CRISPR/Cas9 constructs for epigenetic editing. The model also describes the functionality of the CsyTag and Csy4 as a gRNA-degradation switch to remove the recruitment ability of the MS2 stem-loop. **B)** Relative activation of the ASCL1 target gene by VPR using various recruitment methods. Samples were normalized on mock-transfected cells. While there is no significant difference between samples without Csy4 and Csy4 H29A (0 mM imidazole), samples treated with 32 mM imidazole show a significantly reduced activation effect of VPR. Csy4 WT can effectively reduce the activation effects of

VPR down to base level. **C)** Rescue effects of 32 mM imidazole on different Csy4 mutants. **D)** Relative activation of the ASCL1 target gene by VPR using gRNA-CsyTag-MS2 and MCP-VPR in the presence of various concentrations of imidazole. Samples were normalized on mock-transfected cells. **E)** Relative activation of the ASCL1 target gene by VPR using gRNA-CsyTag-MS2 and MCP-VPR in the presence of various concentrations of imidazole. **F)** The portion of undigested gRNA over the concentration of imidazole [mM]. **B) to F)** Error bars denote the SEM from at least three biological repeats performed in technical duplicates. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, two-sided t-test relative to samples without Csy4.

I further determined the portion of digested to undigested gRNA to verify whether the signal reduction is indeed caused by MS2-release from the scaffold. I reverse-transcribed isolated RNA using random oligomers and analysed the cDNA with RT-qPCR (**Figure 2-20F**). For this, I used the primers Q_gRNA_f, Q_gRNA_cut_r and Q_gRNA_full_r and determined background signal with Q_BleoR_BG_f and _r (Supplementary Data **Table 0-12**). Adding imidazole dropped the percentage of undigested gRNA significantly already at concentrations of 0.5 mM (79.9%±4.0%). The lowest portion of undigested gRNA was detected at 32 mM (17.2%±0.2%); however, Csy4 WT achieved even stronger effects (9.4±0.6%), confirming our hypothesis of imidazole induced gRNA degradation.

Another factor that has not been addressed so far is the question of time. How long does it take for Csy4 to cleave the CsyTag and how quickly is the recruitment of the MCP-effector construct reversed? I answered this question by recruiting MCP-3xBFP to the telomeres of HEK293 cells (**Figure 2-21A**), as described before (see chapter 2.2.3) utilizing a CsyTag-modified gRNA (TELS-gRNA-CsyTag-MS2, Supplementary Data **Table 0-1**). I additionally transfected Csy4 H29A and monitored the fluorescence 1 dpt in live-cell imaging. Initially, the telomeres are visible as distinct dots in the nucleus of the transfected cells (**Figure 2-21B**). The addition of 32 mM imidazole had no immediate effect on the recruitment. However, after 1 min, the dots started to disappear and were completely undetectable within 5 min, demonstrating that the nucleolytic activity of Csy4 H29A can be recovered within a few minutes.

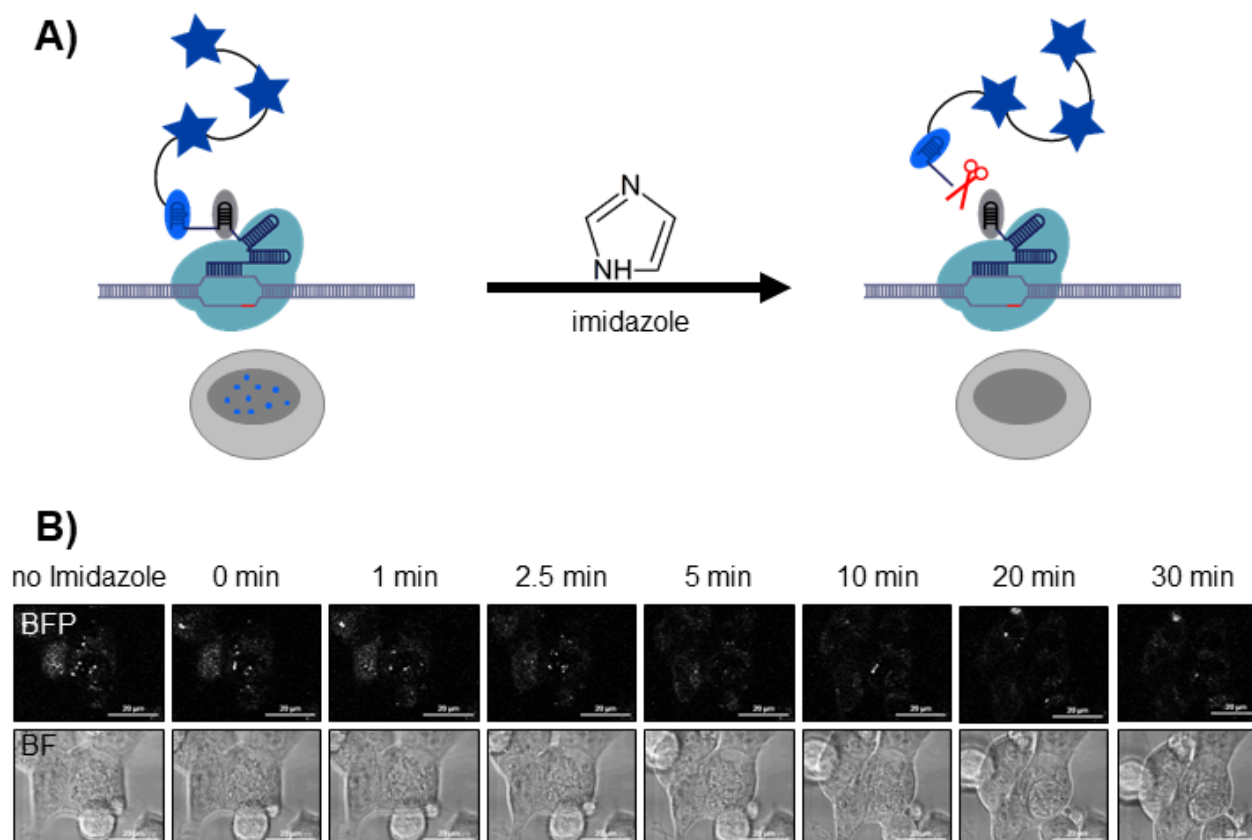


Figure 2-21: CRISPR/Cas-mediated labelling of telomers and imidazole-induced switch-OFF.

Figures and figure text were taken from Oberacker et al. ([211] submitted) and have been slightly modified. **A)** Schematic model of the imidazole-induced release of the MS2 stem-loop and the consequential inability of the system to recruit BFP to the telomers. **B)** Time course of MCP-3xBFP recruitment to the telomers in HEK293 cells in the presence of 32 mM imidazole. BFP has been depicted as white for better visibility on paper. Pictures were taken using an LSM 710 AxioObserver (Plan-Apochromat 63x/1.40 Oil DIC M27), detection for MCP-3xBFP with Ex405 and Em442.

I here developed two new, extremely powerful and compelling functionalities for the EPIC'RISPR platform. The FRB/FKBP-based ON-switch, as well as the CsyTag-based OFF-switch, allow not only the arbitrary induction of the transcriptional alteration effects but also their dose-dependent modulation within minutes. The achieved effects cover a range from 0 to 49% for the ON-switch and 5 to 100% for the OFF-switch. The Csy-tag system can further be applied for reversible chromatin labelling and the regulation of transgene expression (see chapter 2.6). The results of this chapter are part of a manuscript which is currently in submission [211].

2.5 EPIC'RISPR for large-scale parallel modulation of transcription

The vast majority of knowledge we currently have about the epigenetic network is based on associative data. This means we know that epigenetic modifications are linked to the transcriptional state of a gene and that they may appear at specific locations relative to the genetic environment. However, we do not know whether these marks cause the respective transcriptional state or if they are simply the product of it. The same is true for most epigenetic modifiers. From structural data, scientists extracted conserved sequences and demonstrated for some of them their specific enzymatic activities, substrate preferences and even products. However, we do not know much about how they interact with the epigenetic network, what their interaction partners are, how they find their targets or how they avoid off-target activity. To answer all these questions, associative data is simply not sufficient. Epigenetic editing, however, can specifically disturb the epigenetic network by recruiting EDs to a distinct gene and monitor the effects on transcriptional activity, changes in chromatin structure and the DNA methylation status. The EPIC'RISPR platform represents a particularly well-suited epigenetic editing tool for such a venture, as it enables high-throughput experiments for introducing and probing multiple epigenetic alterations simultaneously.

In this chapter, I describe the results of an experiment where I targeted more than 60 different epigenetic EDs to five differently expressed target genes. With the help of Sven Höhn and Luca Schelle, I fused over 100 different human genes or their catalytic domains (CD) with MCP as described in chapter 2.2.2. I then performed a large scale version of the experiments described in chapter 2.3.2 and recruited more than 60 of these constructs to five differently expressed genes. I harvested the cells 3 dpt, isolated their RNA, reverse transcribed it into cDNA and performed RT-qPCR analysis.

2.5.1 Controls

To verify the results of such a large scale experiment, I utilized the already verified constructs MCP-KRAB and MCP-VPR as controls (**Figure 2-22**). Both EDs performed similarly as observed before and achieved repression or activation, respectively, in all five target genes (compare **Figure 2-13**). As in all previous editing experiment, pUC19 transfected cells served as a basis for normalization. I also transfected samples only with dCas9 and the gRNA mix, as this combination has been shown to be sufficient for repressing certain genes if bound to the promoter [150]. In this set-up, however, the

gRNA/dCas9 complex did not change the expression level of any of the five target genes significantly. As described in chapter 2.3.2, the variation between the sample replicates is higher than observed for experiments with only one target.

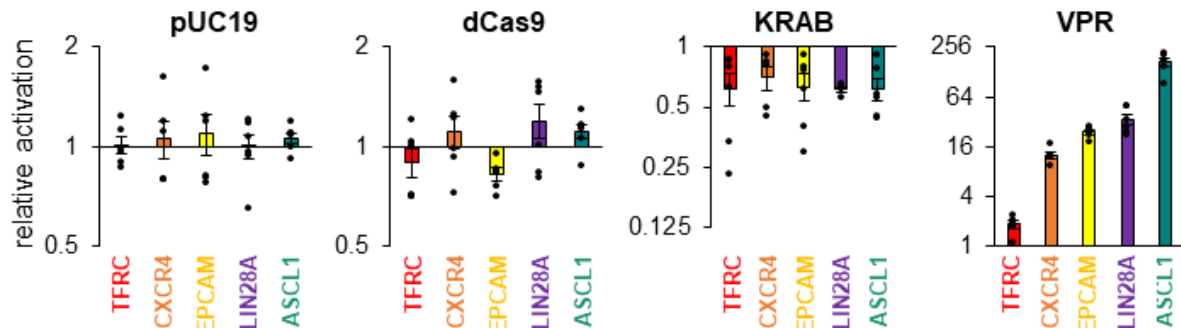


Figure 2-22: Control samples for multitarget perturbation. pUC19 transfected cells served as a non-targeting control, and all other samples were normalized to them. dCas9 and gRNA alone have been shown to repress certain targets. To determine their influence on the chosen targets, samples have been transfected without effectors. MCP-KRAB and MCP-VPR served as positive controls for effective transcriptional repression and activation. Error bars represent SEM of three to six biological replicates with two technical replicates each.

2.5.2 HKMTs

Methylation of histones is one of the most studied epigenetic modifications. To learn more about the effects of synthetically introduced methylation marks, I targeted 27 HKMTs of five different specificities (H3K4, H3K9, H3K27, H3K36 and H4K20) to five differently expressed genes and monitored their influence on transcriptional activity (Figure 2-23).

Overall, targeting any HKMT to *TFRC* and *EPCAM* does not seem to have a substantial effect on the transcriptional status of either gene, although *EPCAM* is slightly repressed in some cases. *CXCR4* and *LIN28A*, however, are marginally upregulated in most cases, and *ASCL1* shows both transcriptional activation and repression not following any specific pattern. As *CXCR4* is a monoexonic target, the results regarding this target gene must be evaluated carefully, as potentially co-purified genomic DNA could influence the results.

The effect on transcriptional activity of any ED depends on the positioning of the mark and the epigenetic surrounding, as some EDs need pre-methylated substrates or other modifications in the close proximity of the target to function. Therefore, it is crucial to look at each ED and target individually when evaluating this data.

H3K4 MTs (Figure 2-23A)

H3K4me3 is the main epigenetic modification associated with transcriptional activity. It has been shown to maintain transcription by interacting with the NURF complex [29], and it directly prevents DNA methylation [42]. In co-operation with H3K36 methylation, it even interacts with the HAT-recruiter complex NuA4, and it is therefore directly linked to actively promoting transcriptional activation [41]. I chose five known (KMT2B, KMT2C, KMT2D, SETD1A, SETD7) and one predicted (SET1B) H3K4-specific MTs, from which all but SETD7 are MLL family members, to specifically introduce H3K4 methylation in the promoter regions of the five target genes.

Almost no H3K4-MTs caused a significant alteration in *TFRC* activity, which is expected as the gene is actively transcribed, and the associated histones are most likely already methylated at their K4 residue. Interestingly, SETD1A induced a slight but significant reduction of transcription to 0.86 ± 0.02 -fold.

EPCAM and *LIN28A* both show similar low expression levels (Figure 2-10D), suggesting a lack of H3K4 methylation and potentially high H3K9me3 levels making them good potential targets for KMT2B, KMT2C, KMT2D and SETD7. None of these EDs was able to alter the transcription of *EPCAM*, but KMT2B, KMT2C and SETD7 increased *CXCR4* and *LIN28A* transcription significantly with KMT2B achieving the strongest effect of 2.31 ± 0.27 -fold relative expression on *LIN28A*. Interestingly, SETD1A was also able to significantly promote activation of *LIN28A* (1.96 ± 0.25 -fold) as well, although SETD1A and SETD1B have both shown not to catalyse H3K4 methylation if H3K9 is already methylated [257]. This might suggest different regulatory mechanisms of *EPCAM* and *LIN28A*, despite their similar expression levels.

ASCL1 is the gene with the lowest base expression of all five targets, and transcriptional activation with VPR achieves effects that are at least an order of magnitude higher than with any other of the five targets. Interestingly, none of the six H3K4-specific MTs was able to increase the transcriptional activity of *ASCL1* significantly. SETD1B surprisingly even strongly repressed the gene down to 0.42 ± 0.05 -fold transcriptional activity. As mentioned in the paragraph above, it is not likely that SETD1B causes activation of *ASCL1*, as the gene is strongly repressed and most likely contains H3K9me3 marks. However, the strong repressive effect of SET1B on the expression even suggesting more diverse roles of the ED under distinct circumstances.

H3K9 MTs (Figure 2-23B)

Although being a characteristic mark for heterochromatin, H3K9 methylation has also been located in the promoters of some active genes [28]. The synthetic introduction of H3K9 methylation is therefore especially interesting, as the influence on transcriptional activity is especially hard to predict. I targeted six members of the SUV39 family (EHMT1, EHMT2, SETDB1, SETDB2, SUV39H1, SUV39H2) as well as PR/SET domain 2 (PRDM2/RIZ), all of which are known as H3K9-specific MTs. While SUV39 family proteins are typically associated with transcriptional repression, PRDM2 is known as a positive regulator for some genes [258].

None of the SUV39 family members achieved any significant change in transcriptional activity in *TFRC* or *EPCAM*. For SUV39H1 and SUV39H2, this was expected as both proteins require H3K9me1 as a substrate [46], which should not be present in the promoter region of these genes. EHMT1 and EHMT2 mainly cause H3K9me2, which is known as a repressive mark in euchromatic regions. However, the two proteins facilitate the inheritance of this mark and might need H3K9me2 present on neighbouring nucleosomes to catalyse the methylation [52]. PRDM2, on the other hand, was able to increase *TFRC* activity slightly but significantly to 1.22 ± 0.06 -fold, which is rather surprising, as even VPR can increase the expression of this gene just about 1.85 ± 0.18 -fold.

Interestingly, all recruited EDs but SETDB1 activated *LIN28A* significantly. This is expected for PRDM2, which achieved the strongest activation of 4.15 ± 0.22 -fold, but less so for the SUV39 family. In comparison, *CXCR4* was also upregulated by all used EDs; however, only PRDM2, SETDB1 and 2 achieved significant effects. The changes observed in *ASCL1* transcription are rather ambiguous. The two strongest effects in either direction were achieved by EHMT 1, which significantly repressed the gene to 0.73 ± 0.08 -fold expression, and SUV39H2 activating *ASCL1* by 1.86 ± 0.21 -fold. This activation effect, however, is rather weak, considering the extreme potential of *ASCL1* for activation.

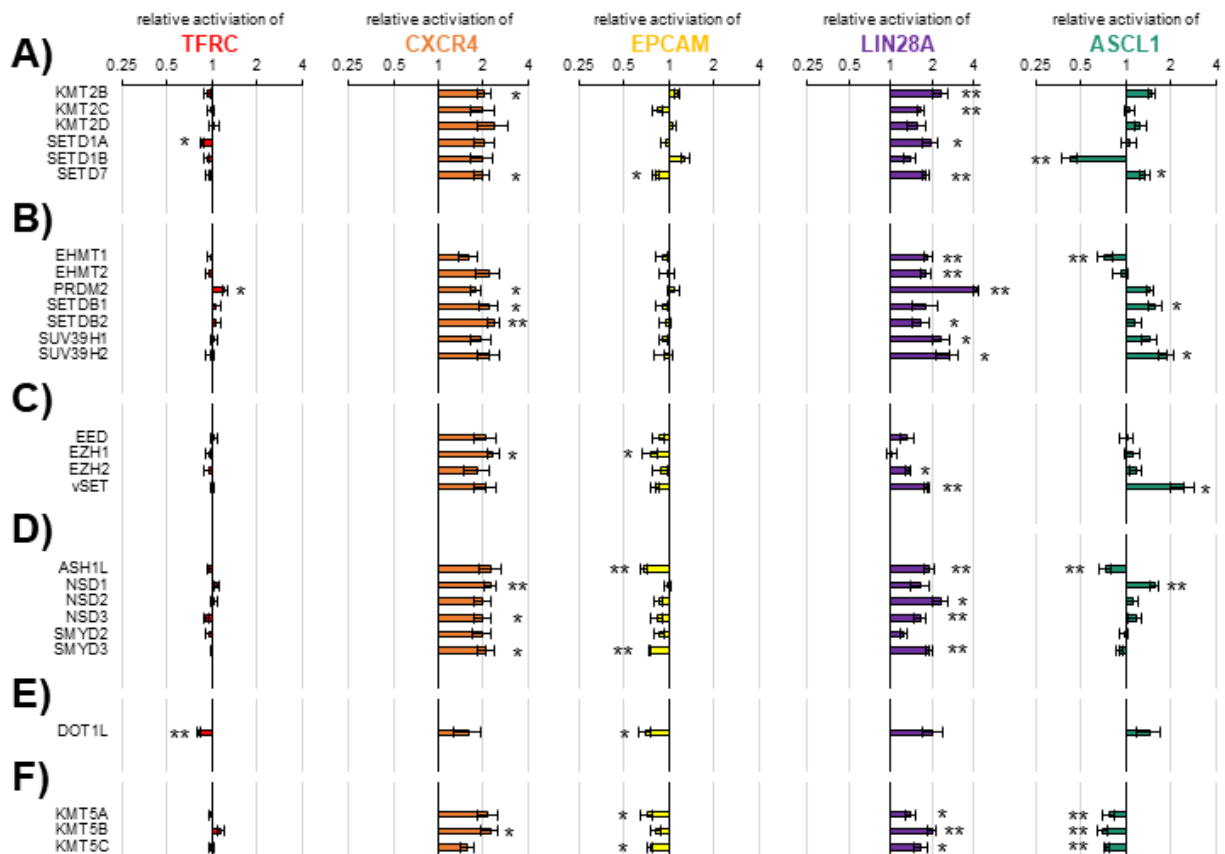


Figure 2-23: Simultaneous histone methylation of five differently expressed genes. A) H3K4 methyltransferases, mainly MLL family. B) H3K9 methyltransferases, mainly SUV38 family. C) H3K27 methyltransferases and facilitating proteins of the polycomb complex. D) H3K36 methyltransferases. E) H3K79 methyltransferase DOT1L. F) H4K20 methyltransferases. Error bars denote the SEM from three to six biological repeats performed in technical duplicates. *P < 0.05, **P < 0.01, two-sided t-test.

H3K27 MTs (Figure 2-23C)

H3K27 methylation is the hallmark of the PRC and therefore involved in transcriptional repression. However, the mark can also co-exist with H3K4me3 in bivalent promoters of important genes in ESC [65]. To elucidate the effects of H3K27me3, I recruited three PRC-associated proteins, namely the two main HKMTs, EZH1 and EZH2, as well as the Embryonic Ectoderm Development (EED) protein, which does not exhibit methylation activity itself but plays a crucial role in HKMT recruitment. Furthermore, I recruited the dimeric viral SET domain methyltransferase (vSET), the smallest known H3K27 MTs [259].

The addition of H3K27me3 on promoters of highly active genes is hard to predict, as they usually already contain H3K4me3 and the respective promoters would become bivalent. It hence comes to no surprise that *TFRC* displayed again a remarkable

resistance against epigenetic alteration and neither ED achieved any significant change in transcription. However, all four effectors were able to repress the expression of *EPCAM*, with EZH1 achieving significant repression down to 0.75 ± 0.08 -fold. In contrast to *TFRC*, *EPCAM* displays a lower transcriptional base activity reducing the chance of creating bivalent nucleosomes by adding H3K37me3 to already present H3K4me3.

EED, EZH1 and EZH2 showed no significant effects on neither *LIN28A* nor *ASCL1*. Both genes are already very weakly expressed and might already contain H3K27 methylation sites. And although *EPCAM* and *LIN28A* share similar expression levels, the same modifications might have a different impact on their transcriptional state, further showcasing the extreme complexity of how genes are controlled. Interestingly, vSET was able to significantly activate *LIN28A* and *ASCL1* by 1.83 ± 0.10 and 2.43 ± 0.49 -fold, respectively. However, as mentioned for SUV39H2, these effects are rather weak compared to the potential of both genes for transcriptional activation.

H3K36 MTs (Figure 2-23D)

Although H3K36 methylation is known to actively repress transcription by recruiting deacetylation complexes, it is usually located in the gene body where it prevents aberrant transcription initiation [72,73]. We, therefore, do not know much about how this mark will influence the transcription status of a gene, if introduced in the promoter region. To elucidate that, I recruited NSD1 and NSD2, which introduce H3K36me1 and me2, as well as NSD3, ASH1L, SMYD2 and SMYD3, which are known to trimethylated H3K36 based on premethylated substrate. As H3K36me3 is associated with DNA methylation [75], recruiting these modifiers to a promoter could potentially cause strong repressive effects.

As for most other HKMTs, neither H3K36-specific MT affected the transcription of *TFRC*. H3K36me1 is usually found close to the promoter of active genes, which means that additional marks might not have a big impact in the respective context. On the other hand, there is no naturally occurring H3K36me1 or me2 present in the promoter region of these genes, which lowers the probability for NSD3, ASH1L, SMYD2 and SMYD3 to find a substrate.

For *EPCAM*, neither of the me1-specific MTs achieved any significant drop in expression, however, recruiting ASH1L and SMYD3 repressed the gene significantly to 0.68 ± 0.04 -fold and 0.74 ± 0.02 -fold activity, respectively.

In the repression targets *LIN28A* and *ASCL1*, only ASH1L was able to achieve significant repression for *ASCL1* down to 0.73 ± 0.07 -fold. NSD1 however, weakly activated *ASCL1* to 1.55 ± 0.11 -fold expression. *LIN28A* was again upregulated by a majority of the EDs with NSD2 achieving the strongest effect (2.32 ± 0.32 -fold).

H3K79 MT (Figure 2-23E)

DOT1 like (DOT1L) HKMT is the only known enzymes to introduce H3K79me3 in the mammalian system. The modification is involved in many biological processes such as DNA damage response, polymerase elongation or cell cycle control, and it is strongly associated with active transcription [260]. Surprisingly, the introduction of H3K79me3 to *TFRC* and *EPCAM* represses both genes significantly down to 0.81 ± 0.03 -fold and 0.69 ± 0.08 -fold, respectively. Although causing slight activation of *CXCR4*, *LIN28A* and *ASCL1*, none of these effects is significant.

H4K20 MTs (Figure 2-23F)

Methylation of H4K20 correlates strongly with transcriptional repression. H3K4me1 is mainly introduced by KMT5A [80] and serves as a substrate for H3K4 trimethylation catalysed KMT5B and KMT5C [84]. When targeted to *EPCAM* and *ASCL1*, all enzymes achieved significant repression of both genes. *EPCAM* showed the strongest repression when modified by KMT5A (0.71 ± 0.07 -fold) and *ASCL1* when treated with KMT5B (0.70 ± 0.06 -fold); however, the effects of all three modifiers were very similar. Interestingly, neither enzyme was able to alter the transcriptional state of *TFRC* in this set-up, while *LIN28A* and *CXCR4* were upregulated by each of the HKMTs.

2.5.3 HKDMs

Removing an epigenetic modification can have the same impact on the transcriptional state of a gene as adding one. In the case of histone methylation, there are several examples of antagonism that have been described in chapter 1.1.3. The removal of a certain mark may, therefore, allow another to be introduced and hence alter gene expression. We know that several marks, such as H3K9me3 or H3K27me3, can interact with chromatin remodelers and removing them might alter chromatin structure

in some cases (see chapter 1.1.3). To study this, I recruited 23 known HKDMs to five target genes, with specificities for H3K4me, H3K9me, H3K27me, H3K36me from which many have more ambiguous roles (**Figure 2-24**). Although *CXCR4* shows a tendency to be activated, none of the observed effects is significant. *TFRC* and *EPCAM*, on the other hand, display an even stronger resilience towards transcriptional alteration by HKDMs with a slight tendency for being repressed. *LIN28A* and *ASCL1* display more variable responses and can be significantly activated or repressed, depending on the recruited ED.

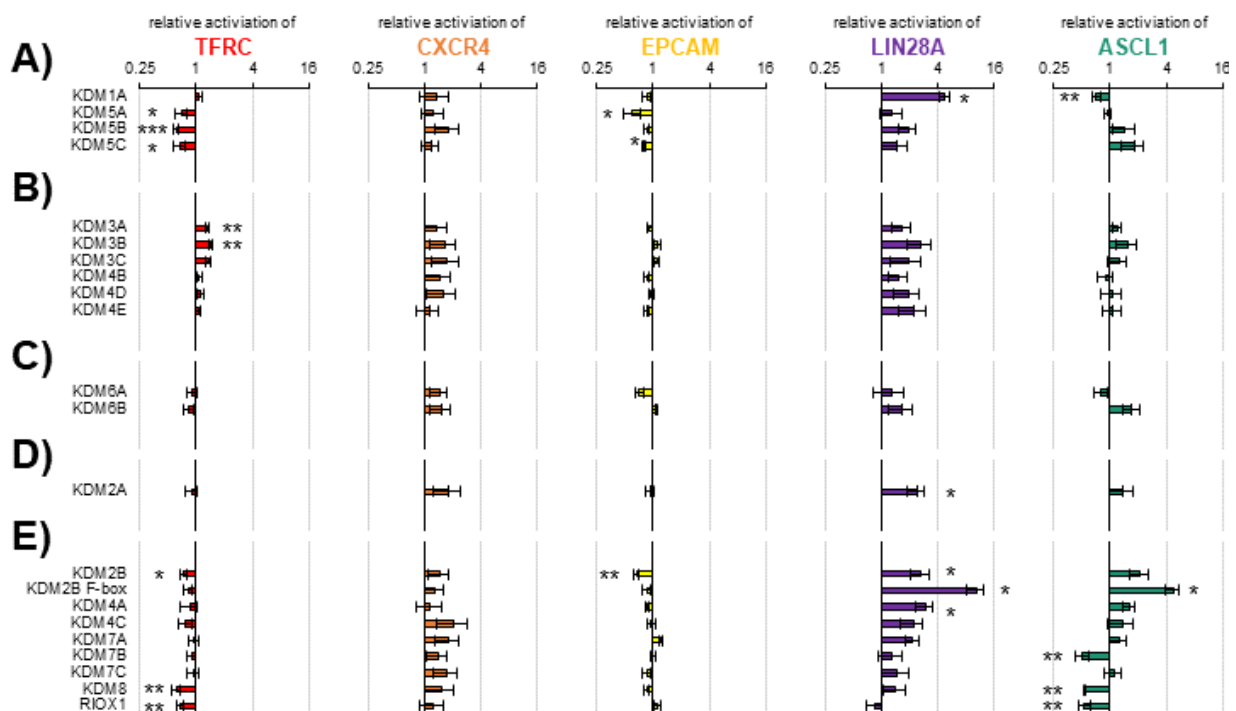


Figure 2-24: Simultaneous histone demethylation of five differently expressed genes. A) H3K4 demethylases **B)** H3K9 demethylases **C)** H3K27 demethylases **D)** H3K36 demethylases. **E)** Multipurpose HKDMs with a broader target specificity. Error bars denote the SEM from three to six biological repeats performed in technical duplicates. *P < 0.05, **P < 0.01, ***P < 0.001, two-sided t-test.

H3K4 DMs (Figure 2-24A)

One of the major purposes of H3K4me3 is the maintenance of transcriptional activity [36,42]. As such, its removal might cause transcriptional silencing by DNA methylation and histone deacetylation. To investigate the influence of H4K4me removal from the promoter region of genes, I recruited KDM1A (LSD1) as the most prominent H3K4 DM. As this enzyme only uses H3K4me2 as a substrate, I also tested three members of the Jumonji/ARID Domain (JARID) class HKDMs, namely KDM5A, KDM5B and KDM5C.

While KDM1A-recruitment had only weak to no effects on *TFRC* and *EPCAM*, all three JARID HKDMs repressed both genes significantly to 0.61 ± 0.04 -fold activity for *TFRC* (KDM5B) and 0.61 ± 0.15 -fold for *EPCAM* (KDM5A). As H3K4me3 is the main epigenetic mark for maintenance of transcriptional activity in the promoter, it makes sense that H3K4me3-specific HKDMs display stronger effects than KDM1A. None of the JARID family enzymes was able to change the expression levels of *LIN28A* and *ASCL1* significantly, which come to no surprise, as genes with a low base expression level should not contain much H3K4me3 in their promoters anyways, again demonstrating the different epigenetic regulatory mechanisms in similarly expressed genes such as *EPCAM* and *LIN28A*. KDM1A was able to surprisingly upregulate *LIN28A* to 4.75 ± 0.87 -fold and repress *ASCL1* significantly to 0.72 ± 0.07 -fold relative activity.

H3K9 DMs (Figure 2-24B)

The removal of H3K9me3 as one of the most well-characterized repressive marks may promote transcription and increase gene activity. To investigate this, I recruited six members of the Jumonji Domain (JMJD) containing HKDMs, with KDM3A, KDM3B and KDM3C being specific for H3K9me2 and KDM4B, KDM4D and KDM4E focusing on H3K9me3 as a target.

Two of the three H3K9me2-specific DMs were able to upregulate *TFRC* significantly up to 1.45 ± 0.07 -fold, but none of the H3K9me3 DMs induced any change in transcription. This is not surprising, as a gene as active as *TFRC* should not have any H3K9me3 present in its promoter region. None of the recruited enzymes achieved any significant change in the transcription status of either *EPCAM*, *LIN28A* or *ASCL1*, but all showed a tendency for activation in the latter two genes.

H3K27 DMs (Figure 2-24C)

Removing H3K27me3, the main mark of the PRC, suggests transcriptional activation. By recruiting KDM6A (UTX) and KDM6B (JMJD3), I wanted to investigate this possibility. Although showing tendencies for both activation and repression in different targets, neither enzyme could achieve a significant change in the expression of any of the five genes.

H3K36 DMs (Figure 2-24D)

H3K36 methylation is the main repressive histone mark in the gene body in which it is gradually distributed, from H3K36me1 close to the promoter to H3K36me3 at the 3'-end [70]. It is therefore unlikely that a gene, neither repressed nor active, has much H3K36 methylation in its promoter region. However, I targeted KDM2A as the main enzyme for removing H3K36me1 to the promoter of the five target genes to investigate its influence.

In four of five genes, there was no significant change in transcription. However, KDM2A was able to upregulate *LIN28A* significantly to 2.39 ± 0.52 -fold expression, which would be expected if H3K36me1 was present in the promoter region.

Multipurpose HKDMs (Figure 2-24E)

Interestingly, many HKDMs are rather ambiguous in their substrate selection and can remove methylation from multiple residues in the histone tails. While this makes the analysis of their functionality especially challenging, they might exhibit a stronger effect on the transcriptional state as they possibly introduce more than one epigenetic modification.

KDM2B or F-Box And Leucine-Rich Repeat Protein 10 (FBXL10) is a member of the JmjC Domain family and catalyses the removal of both H3K4me3 and H3K36me2, suggesting a role in the transcriptional repression, considering the functions of both modifications in active genes. I recruited both the full-length protein (KDM2B) as well as the F-box domain (KDM2B F-box) to the five target genes. KDM2B was able to significantly repress both *TFRC* (0.73 ± 0.08) and *EPCAM* (0.67 ± 0.06), while KDM2B F-box only showed a tendency for repression but no significant effects. Surprisingly both EDs were able to upregulate *LIN28A* significantly. KDM2B F-box achieved an increase in transcriptional activity to 10.3 ± 2.2 -fold, which represents the most potent activational effect of any ED on any gene used in this experiment, except for VPR. Although surprising, this effect agrees with the results of the H3K36-specific DMT KDM2A, which exhibited transcriptional activation on *LIN28A* as well. KDM2B F-box was also able to increase the transcription level of *ASCL1* to 4.7 ± 0.8 -fold, an effect that could only be matched by lysine acetyltransferase 8 (KAT8, see chapter 2.5.4).

KDM4A and KDM4C are both members of the JMJD family and catalyse the demethylation of H3K9me3 and H3K36me3. Both have a strong specificity for the

trimethylated state and do not exhibit demethylase activity on either me2 or me1. The only significant effect was achieved by KDM4A, which activated *LIN28A* to 2.9 ± 0.7 -fold, which agrees with the impact of other H3K9me3 specific HKDMs on this target.

KDM7A (KIAA1718) and KDM7B (PHF) both demethylate H3K9me2 and H3K27me2. They also bind H3K4 and recognize its methylation status, but react differently to what they detect [261]. In the presence of H3K4me3, KDM7A loses all H3K9me2-specific demethylase activity, while its activity toward H3K27me2 is strongly increased. KDM7B, however, cannot demethylate H3K9me2, unless H3K4me3 is present in the same nucleosome. This suggests an interesting relationship between the two HKDMs and vastly different roles in epigenetic programming. This opposing relationship is mirrored in their effects on *LIN28A* and *ASCL1*, respectively. KDM7A activates *LIN28A* (2.1 ± 0.5 -fold) but does not affect *ASCL1* transcription. KDM7B, on the other hand, does not influence *LIN28A* but represses *ASCL1* significantly to 0.51 ± 0.08 -fold.

KDM7C (PHF2) is enzymatically inactive but can be activated by protein kinase A-mediated phosphorylation. The enzyme then binds the AT-rich interactive domain 5B (ARID5B), with which it forms a complex [262]. If both requirements are complied with, KDM7C-ARID5B can bind and demethylate H3K9me2. The recruitment of KDM7C, however, did not affect the transcriptional state of any target, suggesting that phosphorylation or complex formation were somehow impaired.

KDM8 has a broad spectrum of possible substrates, as it generally demethylates arginine residues in H2, H3 and H4 as well as H3K9me1 and me2 [263–265]. Interestingly, after demethylation, it exhibits protease activity and cleaves off the respective histone tail, which facilitates transcriptional elongation. Surprisingly, recruiting KDM8 to *TFRC* (0.6 ± 0.1 -fold) and *ASCL1* (0.54 ± 0.02 -fold) leads to a significant drop in transcriptional activity in both genes. Other target genes were not affected.

RIOX1 (NO66) acts as both an HKDM and a ribosomal histidine hydroxylase [266]. Its primary substrates are H3K4me1, H3K4me3 and H3K36me3, suggesting repressive effects when targeted to the promoter region of actively transcribed genes. Recruitment of RIOX1 leads to significant repression of *TFRC* (0.68 ± 0.07 -fold) and *ASCL1* (0.54 ± 0.10 -fold), and even *LIN28A* is slightly downregulated (0.85 ± 0.22 -fold). *EPCAM* and *CXCR4* did not show a significant response to RIOX1-recruitment.

2.5.4 HATs and HDACs

Acetylation of histone lysine residues strongly correlates with transcriptional activation. The modification directly influences chromatin structure by covering the positive charge of the lysine, but it is also a potential binding partner for transcription factors that modulate gene activity [91]. Introducing and removing histone acetylation in the promoter region by recruiting HATs and HDACs promises direct and strong alteration effects on the transcriptional state of any gene.

Here, I recruited 8 known HATs, namely KAT2A (GNAT5), KAT3B (p300), KAT5, KAT6A and 6B, KAT7-9 (all MYST), as well as 14 HDACs of classes I (HDAC1-3, HDAC8), II (HDAC4-7) and III (SIRT1, 2, 6 and 7) (**Figure 2-25**). As HDACs tend to be extremely large proteins due to their multiple non-catalytic domains, I identified active CDs from published crystal structures [39,212] for HDAC2-6 and cloned them into the MCP vector as described in chapter 2.2.2. For HDAC3, I also tested the full-length version and compared it to the achieved effects of the CD. As HDAC 6 harbours two catalytically active acetylation domains, I cloned and tested them separately (annotated as CD1 and CD2).

Overall, the achieved effects on transcription were mostly expected. HATs generally upregulated several to all targeted genes with *ASCL1* showing the strongest responses. HDACs genuinely downregulated their targets but were rather particular about their targets, especially HDACs of class I.

HATs (Figure 2-25A)

The most responsive targets for the transfected HATs were *CXCR4* as well as the already repressed genes *LIN28A* and *ASCL1*. *LIN28A* was significantly upregulated by KAT2A, 3B, 6B and 9, with KAT6B achieving the strongest effect of 2.8 ± 0.5 -fold relative expression. KAT6B was also the enzyme that showed the biggest significant changes in *ASCL1* transcription by increasing its transcription to 2.35 ± 0.04 -fold. Although KAT8-recruitment lead to the highest activation of *ASCL1* in this experiment (5.4 ± 1.8 -fold), these effects were not significant due to high variability. All other HATs showed a certain tendency for activation in *ASCL1* as well, but the effects were also not significant. Furthermore, KAT6B was the only HAT to upregulate *EPCAM* (2.8 ± 0.5 -fold) significantly. *CXCR4*, on the other hand, was significantly upregulated by all HATs except for KAT2A and KAT9.

TFRC, as an already active gene, was rather resistant against transcriptional changes and did not display activation by any of the HATs. However, KAT2B and KAT8 slightly repressed the gene, which might be due to dCas9 binding, as the effects are not significantly stronger than in the respective control (**Figure 2-22**).

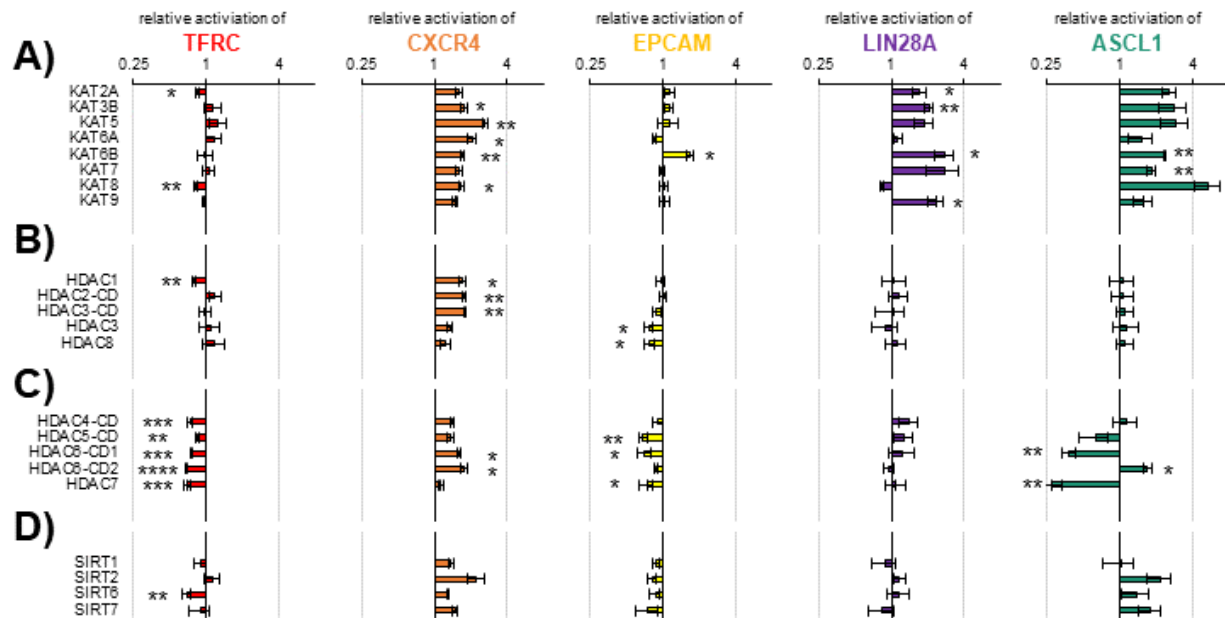


Figure 2-25: Simultaneous histone acetylation and deacetylation of five differently expressed genes. **A)** Histone acetyltransferases of the GNAT, p300 and MSYT families. **B)** Histone deacetylases of class I. **C)** Histone deacetylases of class II. **D)** Histone deacetylases of class III, members of the sirtuin family. Error bars denote the SEM from three to six biological repeats performed in technical duplicates. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, two-sided t-test.

HDACs Class I (Figure 2-24B)

HDAC1 was the only class I HDAC to repress *TFRC* significantly (0.79 ± 0.04 -fold), but it did not change the transcription status of *EPCAM*, *LIN28A* or *ASCL1*. *EPCAM*, on the other hand, was significantly downregulated by HDAC3 (full length, 0.77 ± 0.07 -fold) and HDAC8 (0.78 ± 0.08 -fold), which both did not show any significant effect on any other gene. Interestingly, none of the CDs was able to alter the transcription status of any of the five targets, suggesting that the interaction with other proteins is crucial for effective and rapid repression by deacetylation for class I HDACs. Interestingly, HDAC1 and the CDs of HDAC2 and 3 were all able to slightly upregulate *CXCR4*. Even though significant compared to pUC19-transfected cells, the achieved effects were rather small with 1.69 to 1.78-fold expression. If compared to cells transfected without ED, this increased transcription was not significant.

HDACs Class II (Figure 2-24C)

All five class II HDACs or their respective CDs significantly downregulated *TFRC* in this experiment, with CD2 of HDAC6 achieving the strongest effect of 0.69 ± 0.01 -fold transcription. For *EPCAM*, the recruitment of three out of five HDACs repressed the gene significantly, and the CD of HDAC5 resulted in the strongest downregulation to 0.69 ± 0.06 -fold. Interestingly, the expression of *LIN28A* could not be changed by any of the class II HDACs, while *ASCL1* was downregulated significantly by both CDs of HDAC6 as well as by HDAC7. Furthermore, HDAC6 and HDAC7 achieved the strongest repressive effects of any target in this experiment and downregulated *ASCL1* to 0.38 ± 0.05 -fold and 0.27 ± 0.06 -fold, respectively. As mentioned before, repression of *ASCL1* needs to be interpreted carefully, as the gene has very low base expression. However, these effects were even stronger than what I achieved by recruiting KRAB, suggesting a strong downregulating influence. *CXCR4* was slightly upregulated by both CDs of HDAC6. However, as explained for class I HDACs, these effects are not significant compared to cells transfected without an ED.

HDACs Class III (Figure 2-24D)

In this set-up, the sirtuins of class III did not display any significant impact on the transcription of any of the five target genes, except for SIRT6 which was able to repress *TFRC* significantly to 0.70 ± 0.06 -fold expression level.

2.6 CsyTag for transgene expression control

Considering the great potential of the CsyTag/Csy4 system, I wanted to extend its functionality from the arbitrary modulation of endogenous gene expression (see chapter 2.4.2) also to control the transcription of transgenes. I hypothesized that an mRNA modified with CsyTag sequences in its 5' and/or 3' UTR would be detected and digested by co-transfected Csy4, effectively releasing the 5' m7G-cap and/or the 3' poly-A tail. This would allow exonucleases to degrade the unprotected mRNA (**Figure 2-26A**). pCAG-GFPd2 (Gd2) represents an optimal reporter system for this task. GFP intensities can easily be measured by fluorescence microscopy or fluorescence-activated cell sorting (FACS) analysis, and the powerful synthetic promoter ensures a strong basal expression level. Additionally, the degradation domain (d2) at the C-terminus of GFP shortens the half-life of the synthesized protein, which makes the system more responsive to changes in mRNA titres. Based on this, I

designed several constructs with one or more CsyTag sequences in either UTR (**Figure 2-26B**, Supplementary Data **Table 0-11**) and further removed the NLS at the C-terminus of Csy4 (Csy4 WT Δ NLS, Csy4 H29A Δ NLS, Supplementary Data **Table 0-11**) to synchronize the locations of enzyme and mRNA.

To verify this hypothesis, I co-transfected HEK293 cells with the unmodified reporter vector (Gd2) or CsyTag-GFPd2 (C-Gd2) with and without the wildtype Csy4 (Csy4 WT) (**Figure 2-26C**) and monitored the results using fluorescence microscopy. As expected, cells transfected with C-Gd2 and Csy4 WT displayed no visible GFP signal, while samples without Csy4 showed a similar signal as Gd2 (with and without Csy4). These effects strongly support the suggested hypothesis and are similar to results obtained by others in comparable set-ups [205,206].

I then tested how the imidazole inducible Csy4 H29A Δ NLS mutant influences the GFP signal with and without 20 mM imidazole. For this, I cultivated transfected cells for three days and analysed them using FACS (**Figure 2-27A**). For a detailed description of the analysis, see chapter 4.3.4. In the presence of 20 mM imidazole, Csy4 H29A Δ NLS yields about 21.5% cells in Q2. This represents a reduction to 69% compared to cells transfected without Csy4 (31% in Q2) and down to 75% compared to cells with Csy4 but without imidazole treatment (28.7% in Q2). Csy4 WT transfected samples, on the other hand, contained only 0.5% of cells in Q2, a signal reduction down to 1.6% compared to cells without Csy4.

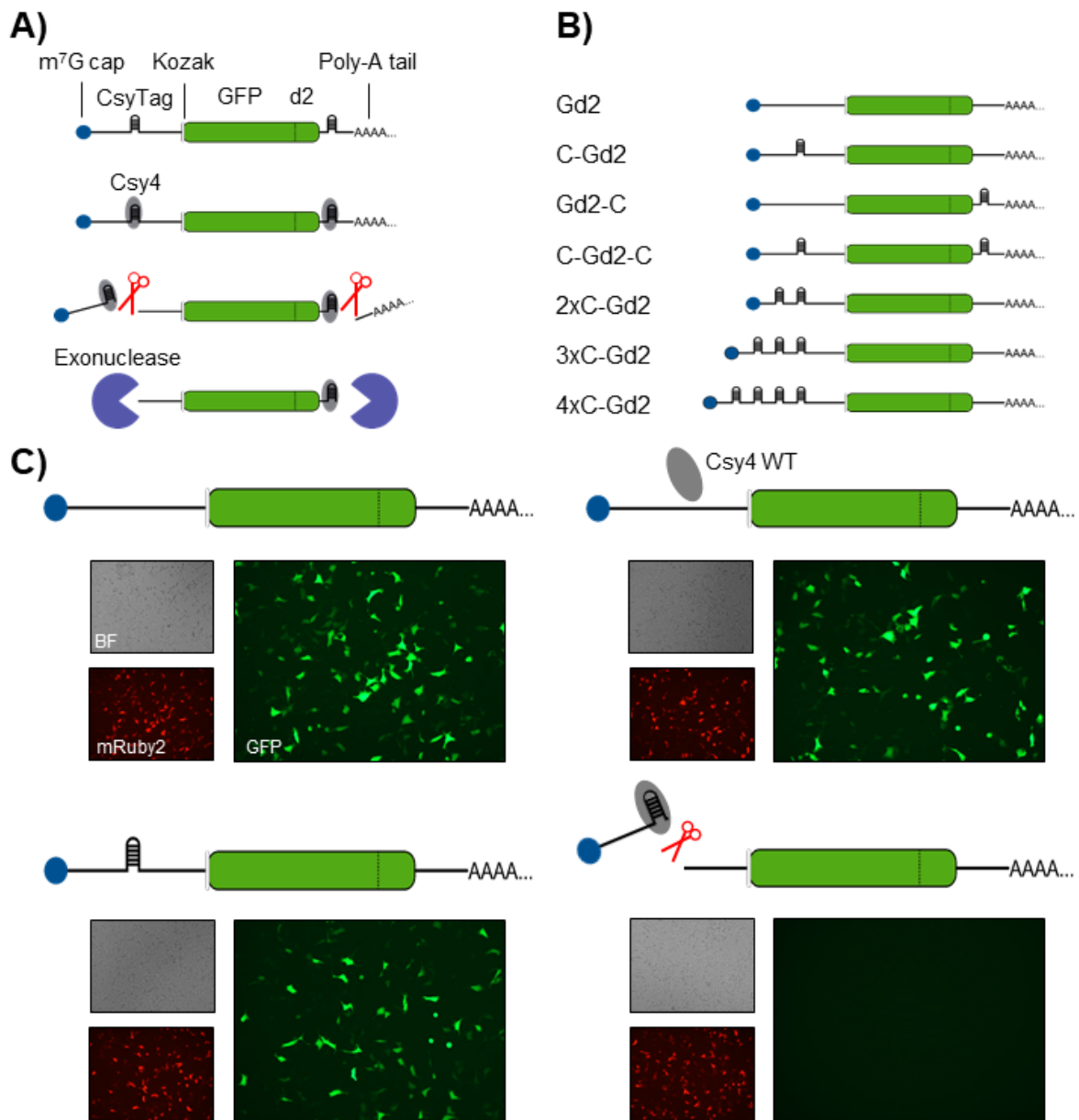


Figure 2-26: CsyTag for induced repression of transgene expression. Figures and figure text were taken from Oberacker et al. ([211] submitted) and have been slightly modified. **A)** Schematic representation of CsyTag's function as a degradation tag in mRNA. **B)** Schematic representation of CsyTag-GFPd2 constructs used in this thesis. **C)** Fluorescence Microscopy (EVOS) of HEK293 cells transfected with either GFPd2 or CsyTag-GFPd2, with or without Csy4 WT. mRuby2 functioned as a positive control for transfection.

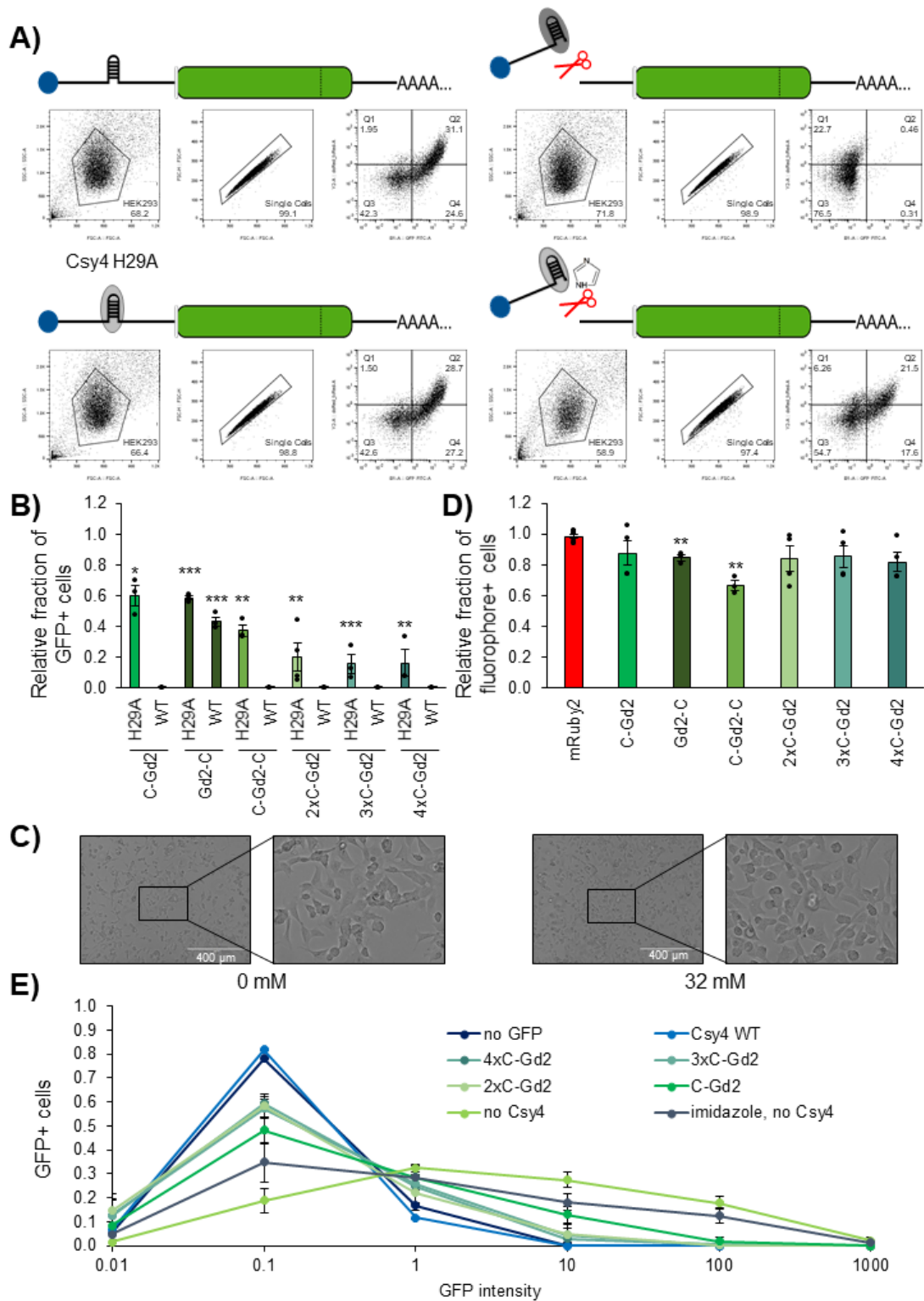


Figure 2-27: Imidazole rescues the nucleolytic activity of Csy4 H29A. **A)** FACS analysis of HEK293 cells transfected with C-Gd2 with and without Csy4 or Csy4 H29A. All samples were co-transfected with mRuby2. Lower left-hand panel: cells were additionally treated with 20 mM imidazole starting 1 dpt. The exemplary scatter plots show the selection process for the analysis from left to right. First, HEK293 cells were determined (left plot). From this fraction, aggregated cell clusters were excluded (middle plot), and the fluorescent single cells were validated in four sectors (Q1 to Q4, right plot). Only double-positive

cells (Q2) were used for analysis. **B)** 20 mM imidazole revives the nucleolytic activity of Csy4 H29A Δ NLS. Samples were normalized on the respective fluorescence signal of samples without imidazole treatment, Csy4 H29A Δ NLS positive samples afterwards on samples without Csy4. Csy4 H29A Δ NLS can effectively reduce the level of various CsyTag/GFPd2 constructs in the presence of 20 mM imidazole to different extend. Samples with multiple CsyTag in the 5' UTR display a stronger response to Csy4 activity. Error bars denote the SEM from at least one biological repeat performed in technical triplicates. **P < 0.01, ***P < 0.001, two-sided t-test. **C)** Influence of imidazole on the morphology of HEK293 cells. At concentrations at or below 32 mM cells show no change in morphology compared to untreated cells. The cells have been transfected with PuroR and treated with puromycin and the respective concentration of imidazole for 3 days. **D)** Influence of 32 mM imidazole on the fluorescence of mRuby2 and GFPd2 constructs. Samples were normalized on the respective fluorescence signal of samples without imidazole treatment. Error bars denote the SEM from at least one biological repeat performed in technical triplicates. **P < 0.01, ***P < 0.001, two-sided t-test. **E)** mRuby2 positive HEK293 cells (Q1+Q2) over the intensity of GFP as measured by FACS analysis. 1x to 4xC-Gd2 samples were treated with 32 mM imidazole.

To improve the effect of Csy4 H29A Δ NLS, I hypothesized that multiple CsyTags in the UTR would increase the chances of Csy4 binding and, therefore, compensate the compromised enzymatic activity of the mutant. I first transfected all designed constructs with and without Csy4 H29A Δ NLS, as well as with and without 20 mM imidazole (**Figure 2-27B**). The cells were cultured and analysed using FACS as described. I first normalised each Csy4 H29A Δ NLS-transfected sample (+/- imidazole) on the respective sample without Csy4, and I then normalised the imidazole-treated samples on samples cultivated without the chemical. As a positive control, I also transfected samples with Csy4 WT. In all cases, Csy4 H29A Δ NLS achieved a significant reduction of GFP signal. The signal of C-Gd2 dropped to 55 \pm 7%, which represents a more robust response than I achieved in the experiment before (75%, no replicates). Implementing the CsyTag in the 3' UTR (Gd2-C) yielded a similar response (59 \pm 1%). A CsyTag in both UTRs (C-Gd2-C) reduced the reporter signal even further down to 38 \pm 4%. I could also observe a tendency for a stronger effect when multiple CsyTags are present in the 5' UTR of the construct. However, there is no significant difference between 2xC-Gd2 (21 \pm 9%), 3xC-Gd2 (16 \pm 6%), and 4xC-Gd2 (16 \pm 9%). Csy4 WT achieved a significant reduction of reporter signal in all but one construct down to less than 1%. The wildtype enzyme reduces Gd2-C signal to 43 \pm 3%, suggesting that the removal of the 3' poly-A tail might be less distressing for mRNA stability than the 5' m7G-cap.

To further improve the effects, I wanted to increase the concentration of imidazole. For this, I first tested the tolerance of imidazole in our strain of HEK293 cells and found that concentrations up to 32 mM are well endured regarding cell morphology (**Figure 2-27C**). I also tested the influence of 32 mM imidazole on the fluorescence signal of different CsyTag/GFP constructs as well as on mRuby2 (**Figure 2-27D**). Most constructs did not show a significant drop in fluorescence when treated with imidazole, except for Gd2-C and C-Gd2-C. With $85\pm 2\%$ (Gd2-C) and $67\pm 3\%$ (C-Gd2-C) of the signal achieved by samples without imidazole, these two constructs displayed a certain sensitivity towards the chemical. Further considering the weak effect of Csy4 WT on Gd2-C, I, hence, decided not to use constructs with a CsyTag in the 3' UTR. For the constructs 1x to 4xC-Gd2, I repeated the experiment above with 32 mM imidazole and fractionated the mRuby2 positive cells (Q1+Q2) according to the measured GFP-intensity and displayed the values in a histogram (**Figure 2-27E**). In agreement with the results achieved so far, imidazole alone induces a slight shift of the histogram towards GFP negative cells. C-Gd2 moves this shift even further, followed by the constructs containing multiple CsyTags in their respective UTR. Again, 2x to 4xC-Gd2 showed similar effects with a tendency for a higher impact on signal reduction, if more CsyTags are present. Csy4 WT decreased the overall GFP intensity to levels comparable to GFP negative cells.

Having established the conditions for the CsyTag as a degradation tag for transgenic mRNA, I investigated the influence of different imidazole concentrations regarding the obtained GFP signal. For this, I incubated cells transfected with 1x to 4xC-Gd2 and Csy4 H29A Δ NLS with imidazole concentrations between 0 to 32 mM (**Figure 2-28A**). Again, constructs with multiple CsyTags (2x to 4xC-Gd2) responded the strongest to Csy4 cleavage. I observed a gradual drop of signal from 4 mM (2x: $89\pm 5\%$, 3x: $82\pm 4\%$, 4x: $92\pm 4\%$) to 28-32 mM (2x: $8\pm 1\%$, 3x: $9\pm 3\%$, 4x: $3.2\pm 0.2\%$), which is significantly lower than the signal measured at 20 mM (2x: $26\pm 3\%$, 3x: $13\pm 6\%$, 4x: $8\pm 2\%$). C-Gd2 was less sensitive and only started to display a significant GFP signal reduction at 16 mM ($80\pm 3\%$), which concurs with the results obtained so far. The construct achieved a maximum effect at 32 mM with $39\pm 1\%$ GFP positive cells, which is also significantly lower than the achieved effect at 20 mM in this experiment ($69\pm 3\%$). However, all constructs displayed a negative correlation of GFP signal with an increasing concentration of imidazole. This relationship can be fitted well with a polynomic function ($0.96 < R^2 < 0.99$, **Figure 2-28B**).

In a final experiment, I demonstrated that the reduced reporter signal correlates with the amount of cellular GFP protein (**Figure 2-28C**). For this, I extracted total protein (see chapter 4.3.6) from cells harvested in the titration experiment above (3xC-Gd2) and analysed 2 µg using SDS-PAGE and Western blotting (see chapter 4.3.7). The obtained signal of GFP protein decreased with increasing concentrations of imidazole. At 8 mM this drop in signal became significant, which concurs with the significant drop in GFP at 16 mM in **Figure 2-28A**.

In conclusion, I successfully established the CsyTag as a versatile degradation tag for transgenic mRNA. I can modulate and fine-tune the titre of mRNA by adding a specific amount of imidazole to the cell culture. I further showed that the drop of GFP signal correlates with the amount of protein present in the cells. The results of this chapter have been compiled in a manuscript which is currently in submission [211].

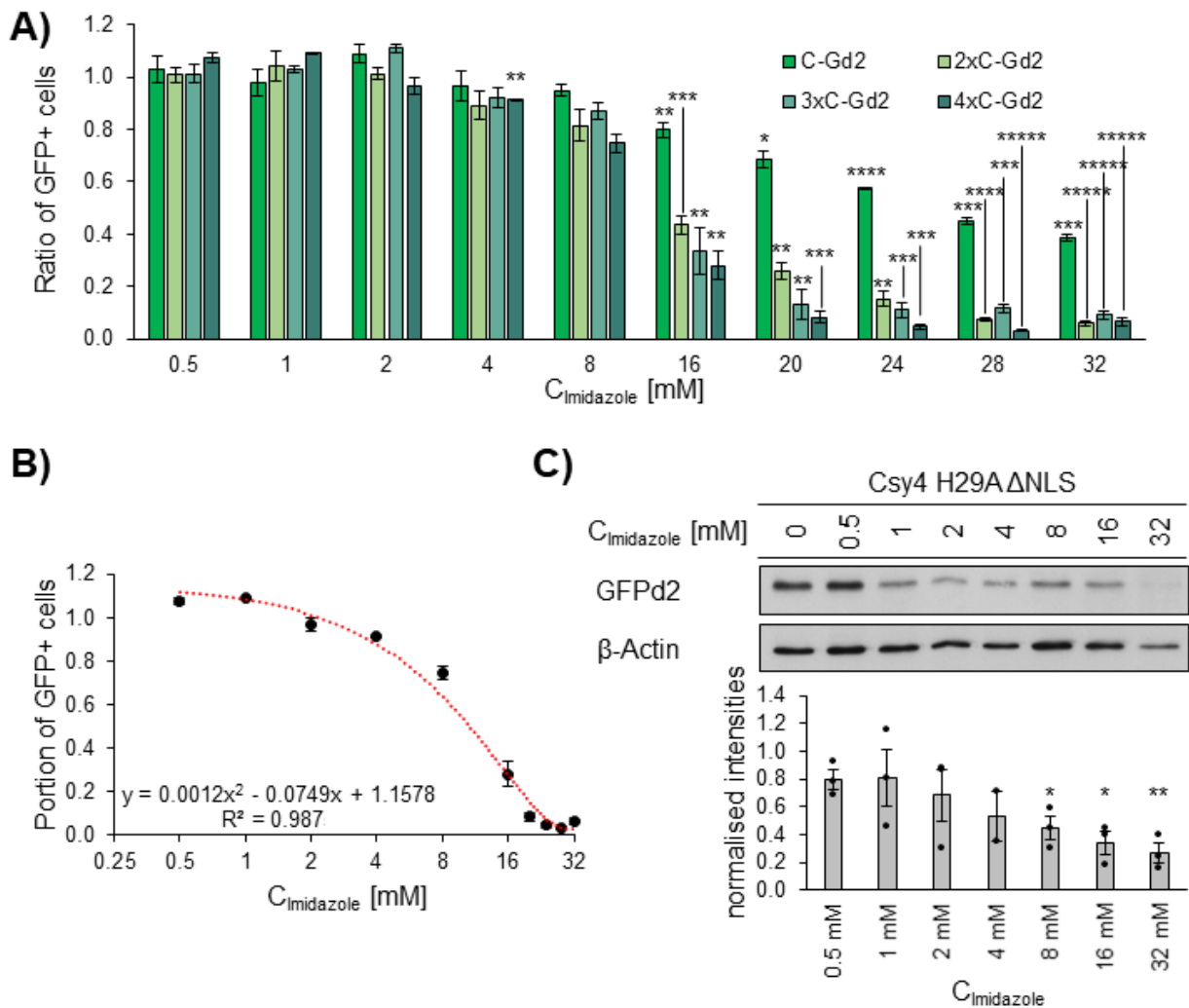


Figure 2-28: Titration of imidazole allows defined modulation of transgene expression. A) The fraction of GFP+ and mRuby2+ HEK293 cells (Q2) of all mRuby2+ cells (Q1+Q2) at different concentrations of imidazole. Error bars denote the SEM from at least three biological repeats performed in technical duplicates. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, *****P < 0.00001, two-sided t-test. **B)** Data for 4xC-Gd2 of A) exemplary plotted as a dot plot including a regression curve in the form of a polynomic function $y = 0.0012x^2 - 0.0749x + 1.1578$ $R^2 = 0.987$. **C)** Western blot analysis of total protein extracted from 3xC-Gd2/Csy H29A Δ NLS samples treated with various concentrations of imidazole. 1°AB: α -GFP IgG (from mouse, 1:2000), α - β -actin (from rabbit, 1:5000), 2°AB: ECL anti-mouse IgG-HRP (1:5000), ECL anti-rabbit IgG-HRP (1:5000). Signal was detected with Western Lightning Plus-ECL. Samples were normalized on imidazole untreated cells. Error bars denote the SEM from three biological repeats. *P < 0.05, **P < 0.01, two-sided t-test.

3 Discussion

3.1 EPIC'RISPR

In the past ten years, epigenetic research has made massive progress in elucidating the epigenetic network by profiling the transcriptional state of various genes, identifying epigenetic modifications and determine their localisation within a gene. However, the associative data provided by this research has not been able to determine the causality of epigenetic modifications and their functionality in the epigenetic system. The newly developed methods for epigenetic editing, especially those based on the CRISPR/Cas system, can introduce or remove single epigenetic marks and precisely perturb the epigenetic state [162]. The response of the epigenetic network can be analysed by measuring changes in the transcriptional state (e.g. by RT-qPCR) or the enrichment of specific modifications (e.g. by chromatin immunoprecipitation (ChIP) or bisulfite sequencing), which allows us to understand better the causalities behind distinct epigenetic marks and how they influence the chromatin structure. Shortly before I started working on this thesis, two groups published a version of the CRISPR/Cas editing system that utilizes short viral stem-loops in the gRNA scaffold, which can be recognized and bound by specific binding domains [175,176]. By fusing the desired EDs to the binding domains, this set-up unites the targeting and recruiting functions in one molecule, the gRNA. Inspired by this method, I developed the EPIC'RISPR platform, a toolbox-system for modular, parallel, and tuneable modulation of gene expression.

3.1.1 Utilizing a modular toolbox

As emphasised by the studies published prior to the start of this thesis [175,176], using aptamer/binders instead of direct dCas9-ED fusion constructs increases the modularity of the editing system by combining recruitment and targeting in the gRNA molecule. Like this, the system enables a flexible alteration of the target by simply exchanging the gRNA. However, each gRNA and AB-ED fusion construct still must be cloned individually, which consumes a lot of time. The EPIC'RISPR platform utilizes elegant cloning strategies for both components that allowed my students and me to create ready-to-use libraries of hundreds of differently modified gRNAs and aptamer binder-EDs within a few weeks (see chapters 2.2.1 and 2.2.2). From these libraries, one can

choose any combination of target, binding system and ED and directly transfect the desired components into the cells, vastly increasing the flexibility and modularity of the EPIC'RISPR system compared to other aptamer-based editing set-ups.

3.1.2 Parallel modulation of transcription and chromatin imaging

The true strength of aptamer-based CRISPR/Cas editing systems is the extreme flexibility in multiplexing experiments. Here, I successfully utilized the EPIC'RISPR system for all three variations of parallel targeting described in **Figure 2-11**. For modulating multiple target genes with one ED simultaneously, I first compared the efficiencies of single gRNAs versus a target-specific mixture of two to eight gRNAs. To do so, I transiently transfected dCas9, MCP-VPR and either a single MS2-modified gRNA or an equimolar mixture of all target-specific gRNAs into HEK293 cells. For *ASCL1*, *LIN28A* and *CXCR4* the gRNA mixture was at least as efficient as the best performing single gRNA or achieved even stronger transcriptional activation, suggesting additive effects when recruiting VPR to multiple positions in the promoter. In the case of *TFRC* and *EPCAM*, however, at least two of the single gRNA outperformed the respective mixture significantly ($P < 0.05$). Interestingly, for both targets, there is at least one of single gRNA that caused significant repression of the target gene ($P < 0.01$), which is rather counter-intuitive as the recruited VPR proved to be a powerful transcriptional activator as demonstrated for other gRNAs in this experiment. However, the respective gRNAs (*TFRC*: 2 and 5, *EPCAM*: 6) bind the target gene in the transcribed region of the promoter (**Figure 1-10**). Although this is true for other gRNAs as well, the precise positioning of these three gRNAs might cause CRISPR interference by sterically blocking the RNA polymerase [150]. Furthermore, using a mixture of up to eight gRNAs also reduces the total amount of every single gRNA and consequently, the impact it might have on recruiting the ED to the target.

To keep an optimal ratio between dCas9, AB-ED and aptamer-modified gRNA in a multiplexing experiment, I could not simply use a larger amount of the respective gRNAs. I hence chose the two to three most efficient gRNAs for each of the five target genes and combined them in an equimolar mixture. I transfected cells with this mixture as well as with dCas9 and the transcriptional modification domains MCP-VPR (**Figure 2-12** and **Figure 2-13A**). The observed efficacies varied strongly depending on the target gene. Firstly, the efficacies for *TFRC* and *EPCAM* were significantly better ($P < 0.0001$) in the multiplexing experiment than when targeting a single gene with a

target-specific mixture. For *EPCAM*, MCP-VPR achieved even stronger activation in the multiplex experiment than with the best performing single gRNA ($P < 0.05$). These findings support the hypothesis that CRISPR interference might have been indeed the reason for the weaker performance of gRNA mixtures containing “repressive” gRNAs, especially considering that the total amount of *EPCAM* specific gRNAs is lower in the multiplexing experiment compared to the set-up with only the best performing gRNA.

Activation of *ASCL1* and *CXCR4* in the multiplexing experiment was only insignificantly lower than upregulating both genes individually, although the total amount of target-specific gRNA was lowered by five-fold. Interestingly, *LIN28A* activation was about two-fold less efficient in the multiplexing experiment. Konermann et al. performed a similar multiplexing experiment for dCas9-ED fusion constructs combined with MS2-modified gRNAs and other MCP-ED components [175]. The group targeted ten genes simultaneously, including *ASCL1* and *LIN28A*, and observed a target-dependent reduction of efficiency when compared to individual editing experiments. While *ASCL1* was activated to similar levels as described here, the activation effect on *LIN28A* was reduced by approximately one order of magnitude. They further observed this reduction in seven more genes, while only one target (*IL1R2*) displayed similar efficiencies in both single-gene activation and multiplex activation experiment. To determine whether these effects are due to gRNA dilution or competition for binding the other two components (dCas9-ED and MCP-ED), the group performed the same experiment with ten-fold dilutions of the single gRNAs and observed strongly target-dependent results. Four of the ten target genes were upregulated compared to undiluted single gRNAs (with *ASCL1* displaying the most substantial effect), while six exhibited lower expression levels (including *LIN28A*). This suggests that the optimal concentration of gRNAs, the composition of gRNA mixtures as well as the ratio between gRNA, dCas9 and AB-ED need to be adjusted individually for each target to achieve the best effects. In the EPIC'RISPR multiplexing experiment, all five target genes were upregulated significantly, qualifying the utilized gRNA composition for parallel transcriptional activation of the five target genes *ASCL1*, *LIN28A*, *EPCAM*, *CXCR4* and *TFRC*.

Interestingly, I also observed a general correlation between activation potential by MCP-VPR and base expression level of the gene (**Figure 2-13A**). While actively transcribed genes such as *TFRC* were only weakly upregulated even by VPR, the more repressed genes such as *ASCL1* and *LIN28A* responded strongly to transcriptional

activation. While this is a very plausible finding, it also raises the question of how the basal expression of a gene influences repression experiments. To address this question, I repeated the same experiment but recruited the transcriptional repressor domain KRAB instead (**Figure 2-13B**). Although all genes were significantly downregulated ($P < 0.05$), the samples showed a higher deviation between biological and technical replicates compared to single-target experiments. However, similar as for activation, repression also seems to be influenced by the basal expression of the respective gene. While *TFRC* was downregulated as low as 0.23-fold for one biological replicate, *EPCAM* only displayed average repression of 0.7 ± 0.1 -fold. Interestingly, *ASCL1* and *LIN28A* were also strongly repressed in this experiment, although they generally show a weak base expression level in HEK293 cells. This low base expression together with the overall high variation in repression experiments suggests that any analysis of these two genes needs to be evaluated very carefully in the context of repression. Furthermore, the general efficacies dropped in each gene compared to the single-target experiments. *CXCR4* e.g. was downregulated to 0.7 ± 0.1 -fold relative activity, while in the single-target experiment, KRAB repressed the gene down to 0.41 ± 0.03 -fold (**Figure 2-10B**). In conclusion, the EPIC'RISPR platform is eligible for significant up- or downregulation of multiple target genes in parallel. It thereby yields efficiencies that are comparable to values reported in similar but less modular systems [175].

The next step in multiplexed epigenetic editing is the targeting of multiple genes with various EDs simultaneously. In this study, I used differently modified gRNAs (gRNA-MS2, gRNA-PP7, gRNA-5xPBSa) to successfully recruit three transcriptional modulator constructs (MCP-KRAB, PCP-VPR, PUFa-VPR) to three different genes (*CXCR4*, *LIN28A*, *ASCL1*) in parallel for effective up- or downregulation, respectively (**Figure 2-14B**). I further showed that each construct functions independently at their desired location without interfering with the transcriptional status of the other two genes. In the meantime, similar results were presented by three groups [176,179,267] who each, showed simultaneous activation and/or repression of two endogenous genes. Considering the extreme modularity of EPIC'RISPR, especially after including the PUF domains, which can be easily programmed to bind any desired nucleotide octamer sequence [269] specifically, the system can potentially be utilized to target dozens of target genes simultaneously, recruiting distinct epigenetic EDs to each location. This possibility represents a great advantage of aptamer-based systems over

classical dCas9-ED fusion constructs. Although it is generally possible to perform this kind of parallel targeting with dCas9-ED fusions, the respective set-ups are rather complicated and lack the modularity of EPIC'RISPR. There is a diverse selection of CRISPR/Cas systems with orthogonal Cas9 proteins which use different gRNA scaffolds and PAM sequences that can be used for such purposes [268–270]. However, there are two major limitations to this set-up. Firstly, each dCas9 variant requires the design of a fitting gRNA scaffold, which can be tedious and time-consuming and represents a strong contrast to the modularity of the EPIC'RISPR gRNA library cloning. Secondly, the PAM sequence recognition in mammalian cells, bacteria or *in vitro* might differ for each dCas9 protein and, therefore, has to be tested for each orthologue separately [271]. With the implementation of RNA binding domains and their respective binding sites in the CRISPR/Cas editing systems, such multiplexing experiments became far easier to conduct.

I further applied this version of EPIC'RISPR multiplexing for chromatin imaging by simultaneously recruiting multiple fluorophores to separate genetic loci (**Figure 2-14A**). I used differently modified gRNAs (gRNA-MS2, gRNA-PP7) to successfully direct the two fluorophore constructs MCP-3xBFP and PCP-3xGFP, respectively, to two repetitive chromosomal locations in parallel (chromosome 9, telomeres). This application of the EPIC'RISPR system will become incredibly valuable for studies that focus on the visualisation of chromatin organisation as well as the interaction and dynamics of different chromosomal regions in living cells. After I started this project, several studies showed the simultaneous recruitment of multiple fluorophores to different genetic loci as well [179,180,272], emphasising the great need for precise and parallel chromatin imaging.

One limitation of CRISPR/Cas-based chromatin labelling is signal strength, though. To visualise a non-repetitive locus in the genome, more than 30 gRNAs are necessary to be located close to each other [273]. Theoretically, this issue could be solved by adding multiple binding sites to the gRNA and consequently recruit more fluorophores per gRNA. In this thesis, I designed a gRNA with five consecutive MS2 aptamers at the 3' end and tested a potential signal amplification by transcriptional activation of *LIN28A*. In this experiment, the target gene was upregulated significantly using MCP-VPR (41±1-fold, P<0.001). However, the construct achieved only about 30% of the effect of the standard EPI'RISPR gRNA harbouring one MS2 aptamer. Konermann et al. tried in the past to utilize multiple binding domains in one gRNAs as well [175]. However,

they were only successful, when exchanging the internal stem-loop structures of the gRNA molecules with the respective aptamer. This limits the amplification to three aptamers in total, if another is added to the 3' end. Constructs with two aptamers at the 3' end did not achieve any upregulation at all in their experimental setup. Although Zalatan et al. successfully implemented such 2x-aptamer constructs with the two aptamers being located at the 3' end of the gRNA, they needed an additional short sequence at the far 3' end which folds back on the first aptamer to stabilize the construct [176]. Nevertheless, the respective constructs achieved a signal amplification when e.g. two copies of VP64 or p65 were recruited. Very recently, a research group around Zhang et al. implemented a gRNA with twelve successive MS2 stem-loops at the 3' end, trying to achieve signal amplification of an MCP-AID-UGI fusion construct [274]. Although they were able to achieve significant effects with the 12xMS2-construct, it always performed worse than the control 2xMS2-gRNA. Due to the self-complementary nature of the MS2 stem-loop, it might be that gRNAs containing multiple copies of one stem-loop at the 3' end have issues with proper folding and thus with stability.

The PBS/PUF system, on the other hand, does not require self-complementary sequences for site recognition and hence allows the addition of multiple PBS to the gRNA's 3' end. In this study, I even utilized 5xPBS repeats, as suggested in the original publication [179]. The respective group tested the effectiveness of up to 47xPBS-gRNAs regarding signal amplification but found the 5xPBS version to be the by far most effective construct in each application. While longer constructs were still able to induce significant effects, the signals were always lower than 5xPBS-gRNAs. Interestingly, the solution for this issue of signal amplification seems to lie in the use of dCas9-ED fusion constructs. Back in 2014, the group around Tanenbaum et al. described a technique for signal amplification in gene expression and fluorescence imaging [170]. By fusing multiple copies of epitopes (SunTag) to the C-terminus of dCas9, they were able to recruit up to 24 copies of scFv-fused fluorophores of EDs to one target and achieve significant signal amplification. A possible way of implementing this system to EPIC'RISPR would be to combine an AB domain with the SunTag. Like this, the established gRNA constructs could be utilized in combination with a system that has been demonstrated to amplify signal significantly.

The fact that I achieved a significant and robust upregulation with the 5xMS2-gRNA suggests that EPIC'RISPR gRNAs are still stable when harbouring multiple copies of

the same aptamer. The question remains why the achieved effects are weaker. A possible explanation could be that the proximity of the aptamers sterically hinders the proper binding of the large MCP-VPR molecule. In chapter 2.4.1, I already showed that the smaller VP64 was suited better for inducibility experiments where potentially multiple copies are recruited (compare **Figure 2-19C** and E). This hypothesis is also supported by the fact that all successful signal amplification studies have been performed with VP64 rather than with VPR [170,176]. A logical next step would, therefore, be to repeat the experiment described in **Figure 2-15B** with the MCP-VP64 construct, rather than with MCP-VPR.

In a fourth variation of multiplexing, multiple EDs with different functionalities are recruited to a single genetic location in order to achieve additive or synergistic effects. Several groups have shown that recruiting fusion constructs of dCas9 with more than one ED can lead to increased efficiencies in transcription and target gene methylation. VPR, for example, is a fusion of the three viral transcriptional activators VP64, p65 and Rta folds [109]. The construct can achieve transcriptional upregulation of several thousand folds in some genes, outperforming each of the individual effectors. Yeo et al. demonstrated similar effects for transcriptional repression when they fused the repression-domain KRAB with Methyl-CpG Binding Protein 2 and dCas9, resulting in significantly stronger downregulation of several target genes compared to dCas9-KRAB [169]. The groups around Stepper et al. and Amabile et al. further demonstrated that a synthetic fusion construct of dCas9 and the DNA methyltransferases DNMT3A and DNMT3L (dCas9-DNMT3A3L or co-targeting dCas9-DNMT3A and dCas9-DNMT3L) can introduce *de novo* methylation in the promoters of three target genes far more efficiently than dCas9-DNMT3A [103,168]. Furthermore, it has been shown that co-transfection of PRDM9 and DOT1L can have additive effects on transcriptional activation [166]. Konermann et al. combined dCas9-VP64 fusion constructs with MS2-modified gRNAs and co-recruited MCP-p65-HSF1, achieving stronger transcriptional activation than with either construct alone [175]. Surprisingly, however, very little research has been conducted that uses multiple aptamer structures or the PBS/PUF system for synergistic epigenetic editing and the respective groups only suggest this application as a possibility [179].

Interestingly, multiple studies utilized aptamer- or PBS-modified gRNAs for live-cell chromatin labelling to co-recruit up to three fluorophores to one locus in the genome [179,180,251]. To verify that the EPIC'RISPR system is capable of recruiting multiple

domains to one location, I designed three gRNAs with one to three copies of the following aptamer cassette: PP7-MS2-boxB. I quipped them with the spacer for chromosome 9 specific repeats and co-transfected the gRNAs with dCas9, PCP-3xGFP, MCP-3xBFP and N22p-3xRFP. All gRNAs could efficiently co-localise all three fluorophores, demonstrating that EPIC'RISPR gRNAs are stable, even with nine consecutive aptamers. Inspired by these results, I applied the system for epigenetic editing. By combining a transcriptional activator (VPR) with a very effective DNA demethylase (mTET3del1), I wanted to achieve a synergistic effect on the transcriptional status of the *EPCAM* target gene. However, the gene was only upregulated marginally higher than with VPR or mTET3del1 alone, and the difference was not significant. The samples were harvested 5 dpt, instead of 3 dpt, as studies by other members of our group suggested stronger effects of TET enzymes at later time points, even though VPR activity already drops 5 dpt. However, analysis of later time points of the studies conducted by my colleagues revealed that transcriptional activation by TET peaks much later about 10 dpt. Time-course experiments could hence potentially give a better insight into possible synergistic effects between co-recruited EDs. While VPR usually activates a gene quickly, demethylation is more of a long-term solution for upregulation. Therefore, combining both and analysing the effects over a longer period might result in interesting observations regarding epigenetic stability. Another reason for these rather low efficacies could be again the size of VPR in combination with another ED. Especially as with 339 aa in length, mTET3del1 is not a small partner. A possible solution for this would be to prepare a gRNA mix with two differently modified gRNAs which are specific for the same target but utilize different spacer sequences which are in proximity of each other. Possible combinations of spacers would be e.g. sg3 and sg4 for TFRC (**Figure 1-10**) or sg2 and sg3 for ASCL1 (**Figure 1-11**).

In conclusion, I here demonstrated that the EPIC'RISPR system is capable of highly parallel transcriptional modulation of numerous target genes with one or multiple EDs simultaneously. I showed that EDs targeted to one gene do not interfere with the up- or downregulation of other target genes. EPIC'RISPR can be used for simultaneous visualisation of multiple genomic locations, and I was able to co-localize up to three fluorophores to one target. I also presented evidence for stable and functional gRNAs modified with nine consecutive aptamers on their 3' end. I am confident that these

findings will help to elucidate further the epigenetic network and how various epigenetic enzymes and molecules synergize.

3.1.3 Inducibility and tuneability of transcription

It is the spatiotemporal coordination of multiple genes and their expression that defines and modulates cellular phenotypes. While gene expression of single or multiple genes can be regulated by CRISPR/Cas-based editing systems, the temporal control of these methods is limited. Unfortunately, this limitation restricts our ability to synthetically implement expression networks of a higher order which are necessary to fully mimic cellular pathways that depend on the consecutive and timely limited expression of specific genes. To address this issue, I developed two ligand-dependent switches for the EPIC'RISPR system that allow precise modulation of transcription regarding efficacy and time. The first of these two switches is based on the FRB/FKBP system and switches the EPIC'RSIRP system ON when rapamycin is added to the cell culture. The second switch is based on the catalytically inactive Csy4 mutant Csy4 H29A. Its activity can be rescued by imidazole, which allows to rapidly switch-OFF EPIC'RISPR. Both systems can be easily fine-tuned by titrating the respective chemical and can even be combined to create complex logic gate system.

Switch-ON

When we started working on the use of FKBP and FRB for inducible modulation of endogenous gene expression, no other publication was available that utilized this system in combination with CRISPR/Cas9. However, during the experimental work, three other groups published similar systems in HEK293 and HEK293T cells, respectively. Braun et al. recruited MCP-FKBP and FRB-HP1 to the *CXCR4* promoter using two MS2 stem-loops inside the gRNA scaffold [200]. Bao et al. fused the dimerization domains to the C-terminus of dCas9 and recruited VPR to four different genes, including *ASCL1* [197]. In a similar approach, Tak et al. combined the FKBP/FRB system with the CRISPR-associated endonuclease in *Prevotella* and *Francisella 1* (Cpf1 or Cas12a) and recruited VPR and p65 to the target genes *HBB*, *AR* and *NPY1R* [198].

To find the optimal set-up for the FKBP/FRB system, we designed fusion constructs of both dimerization domains with MCP and VPR, respectively, in different orientations (**Figure 2-19A** and B). MCP-X and Y-VPR (X and Y being either FKBP, 2xFKBP or

FRB) were transfected in an equimolar ratio, together with dCas9 and the *ASCL1* specific gRNA-MS2 mix described in chapter 2.2.3. The cells were treated with A/C Heterodimerizer for two days starting 1 dpt, harvested 3 dpt, RNA was isolated, reverse-transcribed into cDNA and analysed via RT-qPCR. In this experiment, the combination of MCP-2xFKBP and FRB-VPR achieved the strongest activation of the *ASCL1* target gene (34.1 ± 1.7 -fold), which has been observed by Braun et al. as well for the repression of *CXCR4*. In the study of Bao et al. combinations using repeating FKBP molecules (dCas9-2xFKBP and FRB-VPR) also achieved the strongest activational effects on *ASCL1*, *TTN*, *RHOXF2* and *IL1RN*. Tak et al. even utilized four repeats of FKBP and demonstrated increasing effectivity with each additional FKBP molecule. These results are supported by a study from 2012 in which the authors suggested that such an orientation would be beneficial [275], as one molecule MCP-2xFKBP can potentially recruit two molecules of FRB-VPR, while the reverse orientation can only recruit one molecule of 2xFKBP-VPR. Recruiting multiple EDs could possibly lead to an amplification of the introduced effects. I illustrated this model in **Figure 2-19B**.

Another explanation for stronger effects with 2xFKBP could be that the additional binding site simply increases the chances of recruiting any VPR. This seems especially plausible when considering the low maximal activation effect when compared to the direct MCP-VPR construct ($2.8 \pm 0.1\%$). To address this, we titrated the amount of A/C Heterodimerizer added to the culture medium from 25 to 2000 nM and found 250 nM to be the optimal concentration, with weaker effects when using higher (>750 nM) or lower concentrations (<200 nM) of the chemical. This suggests the following model: At higher concentrations of the chemical, both FKBP and FRB might bind separate rapamycin molecules which blocks their binding site for the respective another rapamycin-dimerization domain complex, and consequently leads to inefficient recruitment of VPR. At low concentrations, on the other hand, there are free, unbound dimerization domains which deteriorate VPR-recruitment, as well. At the right concentration, however, all dimerization domains are fully assembled, and optimal recruitment of VPR can be achieved. Furthermore, the supplier of the A/C Heterodimerizer states that concentrations of more than 1 μ M can be toxic for cultured

cells². Lower effects at 2000 nM might, therefore, be explained with toxic effects of the chemical. However, the optimisation of the A/C Heterodimerizer concentration only increased the achieved effects to 66±9-fold activation, which represents 3.2±0.5% of the effect introduced by MCP-VPR. Bao et al. observed comparable peaks with ~50-fold activation of *ASCL1* at 75 nM rapamycin. Interestingly, Braun et al. did not observe efficiencies that low and even achieved stronger effects with the dimerization system than with direct MCP-ED fusions when repressing *CXCR4*. This is a rather counter-intuitive result considering they detected these strong effects with each permutation of FKBP and FRB orientation, and even when not using 2xFKBP. The study by Tak et al. further observed dependency on the selected target gene. When recruiting FRB-VPR with Cpf1-4xFKBP, the research group was able to upregulate *HBB* and *NPY1R* to about 50% of the effect of Cpf1-VPR, but not the *AR* target gene, which did not show any significant activation. Interestingly, when recruiting FRB-p65 in the same manner, they achieved a strong activation of *AR*, while observing far weaker effects on *HBB* and *NPY1R*.

As VPR is a rather sizeable multidomain protein and HP1 recruited by Braun et al. is relatively small, we hypothesized that the size of VPR might sterically hinder the dimerization process. This would also explain, why Tak et al. saw better effects compared to the direct fusion for VPR, as the used 4xFKBP offers more space for multiple VPR molecules. To test this, we recruited the smaller activating subunit VP64 (consisting of 4xVP16) in the same manner to the target gene and achieved up to 49±0.1% of the effect of MCP-VP64. As the exact binding characteristics of the FKBP/FRB system are not fully understood yet [201], we cannot be sure that it is indeed the size of the ED that determines the efficiency of complex assembly or if some other characteristics of the recruited molecules contribute in some way. However, the results obtained by Braun et al., Bao et al. and we suggest that there is at least a general dependency on the used ED.

Interestingly, the strongest activation with FRB-VP64 was achieved at 150 nM A/C Heterodimerizer, with similar efficacies ranging from 100 nM to 250 nM, while

² <https://www.takarabio.com/products/gene-function/idimerize-inducible-protein-interaction-systems/dimerizer-ligands>

FRB-VPR displayed the strongest activation at higher concentrations (200 to 750 nM) with a peak at 250 nM. Tak et al. utilized the same analogue at concentrations of 500 μ M. Although generally achieving more potent effects, the supplier states that this chemical is only non-toxic to cells at concentrations below 1 μ M and recommends a range of 0.05 to 500 nM for cell culture use. Using the chemical in 1000-fold higher concentrations than recommended makes the obtained results unreliable, especially as the group failed to address toxicity at all. Bao et al. used similar concentrations (100 nM) as we, but of another rapamycin analogue (rapalog, Clontech Laboratories) with comparable effects on FRB-VPR recruitment. Braun et al., however, treated their cells with only 3 nM rapamycin (RAP, Selleckchem), suggesting a higher assembly efficiency when using the natural ligand rapamycin rather than its analogues.

In this thesis, I further demonstrate that the FKBP/FRB-modified EPIC'RISPR system can be fine-tuned by altering the concentration of A/C Heterodimerizer. This tuning effect ranges from 13 to 49% relative activation compared to the direct fusion (MCP-VP64), allowing tight control of transcriptional modulation. This effect has also been shown by Tak et al. for Cpf1-based constructs for concentrations of 1 μ M to 1 mM.

A more recent study created so-called chemical epigenetic modifiers (CEMs) consisting of a rapamycin analogue (FK506) linked to a molecule that interacts with the endogenous epigenetic machinery [199]. The CEM can be recruited to a target gene by dCas9-FKBP and recruit various epigenetic players, depending on the molecule attached to FK506. Although the system shows a certain elegance, the achieved effects were relatively small compared to e.g. the direct dCas9-VPR approach (~6%).

Besides the FKBP/FRB system, there are several other possible candidates for drug-induced dimerization. Bao et al. [197] and Gao et al. [196] recently utilized the gibberellin (GA)-inducible GID1-GAI system, to recruit GID1-VPR to multiple target genes by fusing GAI with dCas9 with effects of 1.5% to 37% compared to the direct dCas9-VPR fusion construct when targeting *ASCL1* or *CD95*, respectively. Gao et al. further implemented the abscisic acid (ABA)-inducible ABI-PYL1 system in the same way and activated *CXCR4* from ~50% to almost 100% of the effect with dCas9-VPR. Using both systems GA and ABA, they were able to create multiple logic gates including AND, OR, NAND, and NOR gates, allowing a specific control over transcriptional modulation.

Besides chemically induced dimerization, the field of optogenetics offers several other interesting systems for induced dimerization that could potentially be used for induced epigenetic editing [188]. The respective dimerization partners assemble upon exposure to a certain wavelength, allowing a technically elegant modulation of gene expression. The system most commonly used in the context of epigenetic editing is cryptochrome 2 (Cry2)-cryptochrome-interacting bHLH (CIB). The system assembles in the presence of strong blue light (450-488 nm) and disassembles in the dark. It was already employed for modulation of gene expression with TALE arrays [193,194] when the CRISPR/Cas9 system was not yet utilized for such applications. Later, Polstein and Gersbach combined the system with CRISPR/Cas9 and applied it for transcriptional activation [195]. By fusing one or multiple CIB molecules to the dCas9 N- and C-termini, the group was able to recruit Cry2-VP64 to three target genes, including *HBG1/2*, *ASCL1* and *IL1RN*, achieving strongly target-dependent effects. While *HBG1/2* and *IL1RN* could be activated almost as well as with a direct dCas9-VP64 fusion construct, the impact on *ASCL1* was very poor (~10%). Besides this variability in achievable effects, the system requires highly energetic light for activation. This can cause DNA damage and produces free radicals, making the system not ideal for long term experiments [276]. We tried to include this system in the EPIC'RISPR platform and used LED boxes with Cry2-CIB-specialised light-cycle settings for exposure [277]. However, the obtained results were inconclusive, and we observed severe morphological changes and cell death in our culture (results not shown).

An alternative to Cry2-CIB is the phytochrome B (PhyB)-phytochrome-interacting factor (PIF) system, which assembles under red light (650 nm), reducing the chance of light-induced cell damage dramatically [188,190,191]. Interestingly, the heterodimer stays assembled, even after light exposure, which reduces possible light damage even further. Upon exposure to far-red light (750 nm), however, PhyB and PIF disassemble into monomers, allowing an elegant switch from ON to OFF and vice versa with short exposure times. The system has one disadvantage, though. It also requires phycocyanobilin as a co-factor for proper assembly that has been shown to modulate gene expression, limiting the application of the PhyB-PIF system in mammalian cell culture drastically [278].

Switch-OFF

While the above-described functional upgrades for EPIC'RISPR revolve around an offline *status quo* of the system which can be switched on by adding the trigger (e.g. light or a specific chemical), there are relatively limited options for a reverse inducibility system, which is online until the trigger is added. Here, I established the recognition site of Csy4 (CsyTag) for the EPIC'RISPR system to manipulate endogenous gene expression arbitrarily. In the absence of the trigger chemical imidazole, the system achieved activation effects comparable to already published constructs when recruiting VPR (**Figure 2-20B**) [109]. Co-transfecting Csy4 WT reduced the transcriptional activation effect to 0.5% (2.6 ± 1.8 -fold) compared to the uninduced state. This already represents a NOT gate-like switch-OFF functionality for CRISPR/Cas9-mediated epigenetic editing. Furthermore, comparable results can be obtained by co-transfecting Csy4 H29A and imidazole (4.7% compared to uninduced, 18 ± 0.5 -fold) in a dose-dependent manner (**Figure 2-20D**). This enables arbitrary and highly predictable manipulation of transcription within five minutes by the addition of a simple chemical to the cell culture. However, the system has to be benchmarked for each desired ED and cell type.

The catalytic activity of HNH family enzymes in which the H of the active centre is altered to G can be rescued in the same way as H to A mutants by adding imidazole, with even higher efficiencies [256]. I hence also tested a Csy4 H29G mutant which displayed a ~2-fold lower sensitivity towards imidazole than Csy4 H29A, yet still reduced the activational effect to 10.4% compared to uninduced (31 ± 5 -fold). The catalytic centres of Csy4 and enzymes of the HNH family are not identical. The dissolved imidazole might be more likely to assume the position of the imidazole ring from H when the additional methyl group of A is present in the context of the Csy4 catalytic centre. Both Csy4 mutants displayed similar transfection efficiencies, and they are regulated by the same promoter. Different expression levels can hence be excluded as a potential reason for the weaker effects of Csy4 H29G.

Variations and combination

During the experimental work on this project, a group around Ferry et al. published a study in which they employ CsyTag for a switch-ON functionality for epigenetic editing with CRISPR/Cas9 [279]. In this Study, CsyTag is utilized as a linker that connects the 5' end of the gRNA with a sequence, complementary to the spacer. During the initial

status quo, the spacer-complementary sequence can fold back onto the spacer and form a double-stranded RNA by Watson-Crick-base pairing. This conformation physically blocks the assembly of the gRNA/dCas9 complex. After cleavage of CsyTag with Csy4 WT, the spacer-complementary RNA strand is released. The matured gRNA can then bind dCas9 and recruit the complex to the target location in the genome. Theoretically, this system should respond similarly to imidazole-induced cleavage with Csy4 H29A. If this is true, this ON-switch could be combined with the switch-OFF functionality for CsyTag presented here in this thesis to simultaneously alter the transcriptional state of multiple target genes in different directions (**Figure 3-1A**). Gene A could be activated by MCP-VPR recruitment using A-specific gRNA that contains CsyTag as described in this thesis (AgRNA-CsyTag-MS2, **Figure 3-1A** upper panel), while gene B would not be targeted by a B-specific gRNA built as described by Ferry et al. [279] with a CsyTag-linked B-complementary (cB) sequence blocking the spacer (cB-CsyTag-BgRNA-MS2, **Figure 3-1A** lower panel). Upon imidazole induced cleavage of the CsyTag sequences, the effects would be reversed: AgRNA would still target gene A but not recruit MCP-VPR anymore, removing the initial activation effect. BgRNA-MS2, on the other hand, can assemble with dCas9 and recruit the effector to gene B.

In another variation, Ferry's CsyTag application could be combined with the FRB/FKBP system described above to create an AND gate functionality, where epigenetic editing only occurs in the presence of rapamycin and imidazole (**Figure 3-1B**). This system could be used in parallel to the application described in **Figure 3-1A**, or the basic FRB/FKBP EPIC'RISPR system to induce editing at different time points. Furthermore, combining the EPIC'RISPR switch-ON (FRB/FKBP) and -OFF (CsyTag) functionalities in one molecule could allow the precise start and shut-down of editing by adding either chemical (**Figure 3-1C**). Like this, complex AND/NOT gate-networks and other logic gate systems could be employed to modulate the transcriptional landscape of a target cell in novel ways. This could be especially interesting for stimulating the transcription of genes which are consecutively expressed in nature like certain developmental genes. Simulating native master triggers like this could encourage the cells to pursue large-scale biological pathways such as cell differentiation.

To go even further, dCas9 could also be modified with a self-excising degron structure [280]. In such a scenario, this small molecule-assisted shutoff (SMASh) called protein tag would be fused to the dCas9 C-terminus via an HCV NS3 protease-recognition site. The SMASh-tag harbours an internal protease which cleaves off the linker after

folding. The tag is then processed by the proteasome due to its degron structure. Adding a protease inhibitor can avoid SMASh-tag removal and hence facilitates the degradation of the fused protein as well. Like this, the amount of present dCas9 in the cell could be easily controlled and would allow the construction of even more complex logic gate arrangements. During my master thesis, I developed weaker degron variations to reduce the premature degradation of the fusion protein before the SMASh-tag is released [281]. Although these variants were initially designed to stabilise influenza viruses carrying SMASh-tag modified polymerases, they could be especially crucial for EPIC'RISPR based editing, where the ratio between the different components, including dCas9, is crucial for the efficiency of the system.

Furthermore, the here presented system could potentially be utilized for visual tracking and reversion of epigenetic editing. If used in a gRNA with two or more aptamers, a fluorophore, as well as an ED, can be recruited to distinct genetic locations simultaneously and monitored in live-cell imaging. Upon imidazole addition, the CsyTag could be cleaved, and the recruitment complex would disassemble. If requested, even the reassembly time of the complexes could theoretically be determined after washing out remaining imidazole.

The main advantage of CsyTag is its versatility. It has many more possible applications from which some were investigated in this thesis, and they will be discussed in chapter 3.2. In conclusion, the results shown in this study demonstrate that the catalytic activity of Csy4 H mutants can be rescued by imidazole in an *in vivo* set-up, and more importantly, that it can be fine-tuned and adjusted in a dose-dependent manner. CsyTag adds new functionalities to the EPIC'RISPR platform and theoretically other CRISPR/Ca9 based editing system. Utilizing CsyTag as an OFF-switch in this context offers a fresh perspective compared to other inducible systems that focus on switch-ON functionalities, which all usually work on the protein level. CsyTag hence opens up a new way of approaching dose-sensitive scientific problems. The system allows the design of epigenetic modulation networks for potentially modulating complete biochemical pathways or even cell differentiation. The results discussed in this chapter are part of a written manuscript which is currently in submission [211].

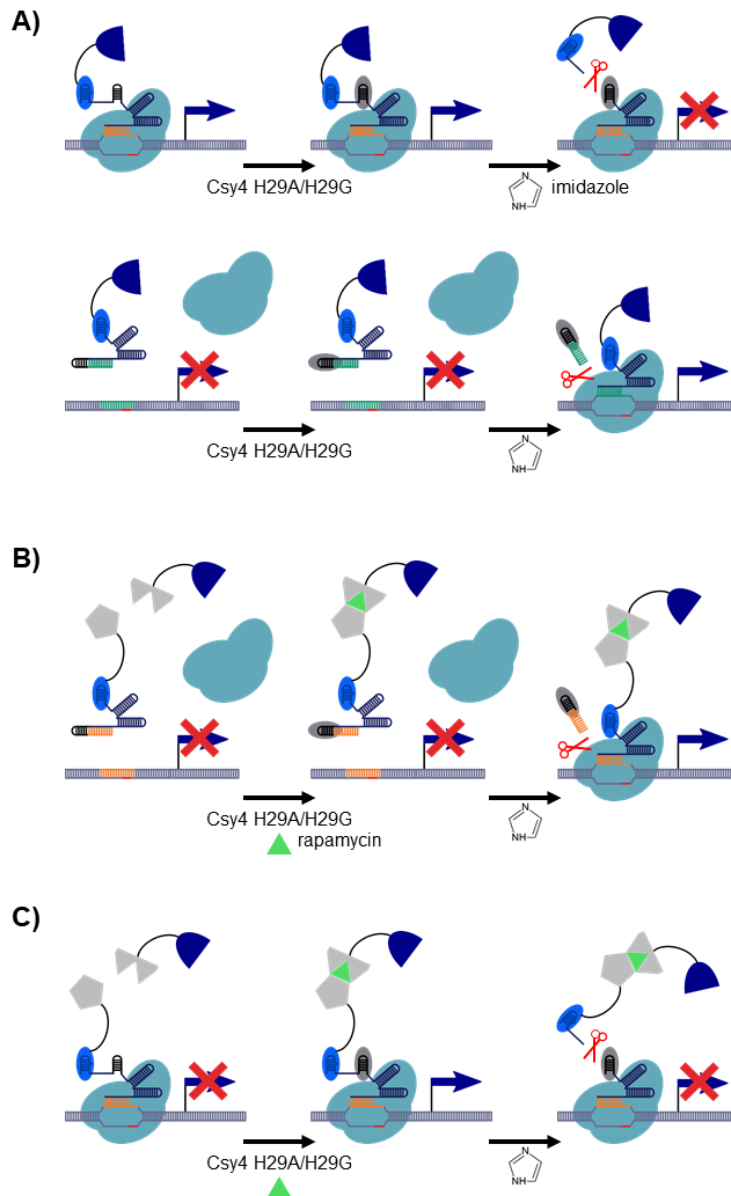


Figure 3-1: Schematic representation of ligand controlled epigenetic logic gates. A) Upper panel: switch-OFF functionality (NOR gate) of the EPIC'RISPR system allows epigenetic modification until imidazole is added. Lower panel: switch-ON functionality employed by Ferry et al. [279] can be used to start epigenetic alteration upon imidazole addition. **B)** Combining the FRB/FKBP EPIC'RISPR switch-ON functionality with the switch-ON by Ferry et al. allows transcriptional modulation only in the presence of both rapamycin and imidazole. Depending on the chosen ED this system creates AND or NOR gates **C)** EPIC'RISPR switch-ON (FRB/FKBP) and -OFF (CsyTag) in one gRNA allows the combination of AND and NOT operations.

3.1.4 Probing the epigenetic code

To investigate the functionalities of a broad range of epigenetic molecules, I utilized the EPIC'RISPR system to recruit more than 60 different EDs to the promoter of five genes simultaneously and then monitored the changes in their transcriptional states.

In this chapter, I will discuss the most striking results of this experiment in regards to transcriptional activation and repression and, based on these results, suggest potential candidates for synergistic co-targeting.

Transcriptional activation

One of the most interesting effects in this experiment was achieved by the HKDM KDM2B and its F-box domain. The full-length KDM2B protein is a known demethylase which demethylates both H3K4me3 and H3K36me3, respectively, and is therefore associated with transcriptional repression. Additionally, KDM2B has been shown to interact with the PRC1 complex [282,283] and H3K79 demethylation [284], further suggesting a role in downregulating transcription. Here, both EDs were able to weakly repress *TFRC* and *EPCAM*, while *LIN28A* and *ASCL1* were both upregulated significantly. KDM2B F-box increased *LIN28A* transcription by 10.3 ± 2.2 -fold, which represents the strongest effect on transcription modulation of any ED in this experiment except for VPR. *ASCL1*, on the other hand, was upregulated 4.7 ± 0.8 -fold. This increase is still among the strongest effects observed in this experiment, yet far lower compared to VPR in the same set-up (169 ± 19 -fold). The F-box motif itself does not exhibit any demethylation activity. However, it is present in a broad range of proteins where it usually facilitates complex formation by protein-protein interaction. These complexes can be involved in ubiquitination (SCF E3 complex), translational repression (FOG-2/GLD-1) but also transcription elongation (Elongin SIII complex) [285]. It is, therefore, hard to predict how F-box recruitment influences the transcriptional state of a gene and it might be dependent on other factors present in proximity to the target regions. As the upregulation only occurred in two genes with the lowest expression levels, the recruitment of any larger protein complex by the F-box domain might open the chromatin structure to an extent that the transcription machinery gains access to the otherwise densely packaged area. Alternatively, this result could suggest an unknown role of H3K36 methylation in the repression of *LIN28A*. This hypothesis is supported by the observation that other H3K36 DMTs, namely KDM2A and KDM4A, achieve similar results on this target gene. Interestingly, *LIN28A* was the only target whose transcription these two enzymes were able to significantly influence.

Another relatively strong activator in this experiment was KAT8 (MYST1), which upregulated *ASCL1* by 5.4 ± 1.2 -fold. In contrast to histone methylation, which usually

interacts with a row of other factors to promote or repress transcription, acetylation directly influences the chromatin structure and enables the transcription machinery to access the promoter. It, therefore, comes to no surprise that an acetyltransferase is amongst the strongest activators of the investigated EDs. Interestingly, KAT8 was not able to upregulate *TFRC*, *EPCAM* or *LIN28A*, suggesting some sort of crosstalk between KAT8 and local modifications. Considering *ASCL1* is by far the most repressed gene, KAT8 might be involved in the early stages of transcriptional activation of strongly repressed genes. However, further experiments are necessary to evaluate this possibility.

KAT3B or p300 is one of the few EDs which showed a relatively consistent effect on transcription, independent of the target it was recruited to. The known HAT significantly upregulated *CXCR4*, *LIN28A* and *ASCL1* to 1.8 ± 0.1 - to 2.8 ± 1.0 -fold and mildly increased the transcription of *TFRC* and *EPCAM*. In contrast to most other EDs in this experiment, KAT3B has already been used for transcriptional modulation in a study by Hilton et al. utilizing a dCas9 fusion construct [163]. In accordance with the results in my experiment, the group reported transcriptional activation in each target gene (*IL1RN*, *MYOD*, *OCT4*), suggesting a relatively unspecific acetylation functionality of KAT3B.

PRDM2 is another example of a relatively consistent transcriptional activator. When recruited, it upregulated all targets at least to some degree, except for *EPCAM* on which it did not have any significant effect. The strongest activation was achieved with *LIN28A*, which was upregulated by 4.2 ± 0.2 -fold. Although being an H3K9-specific MT, PRDM2 is suspected of acting as a transcription factor. By binding to the macrophage-specific TPA-responsive element of the heme oxygenase 1 gene, it might play a role in transcriptional activation, which supports the findings in this thesis [286]. However, to elucidate the exact mechanisms of PRDM2 mediated transcriptional upregulation, more research has to be conducted.

Transcriptional repression

SETD1B achieved one of the strongest repressive effects in this experiment. The ED downregulated *ASCL1* to 0.4 ± 0.1 -fold transcriptional activity, which is similar to the effects achieved by the transcriptional repressor KRAB. Although SETD1B is a known H3K4-specific MT which suggests a role in promoting or maintaining transcriptional activation, the enzyme can only perform H4K4 methylation if the corresponding H3K9

residue is unmethylated [257]. Due to the major role of H3K9me3 in maintaining low transcription profiles, it is therefore unlikely for SETD1B to induce transcription in a gene with such a low base expression level as *ASCL1*. This is supported by the fact that SETD1B was able to upregulate the expression of genes with a higher base expression such as *CXCR4*, *EPCAM* and *LIN28A*. However, it is not clear how the ED can repress the already weakly transcribed target *ASCL1* even further. Although the deviation among the analysed samples was rather low (SEM = 0.05, n = 6), the low base expression of *ASCL1* complicates the analysis of repression experiments. Especially in cases like this, more experiments are required to determine the modification state of the gene regarding typical repressive histone modifications as well as DNA methylation profiles. Nevertheless, the repression effect exhibited by SETD1B is even stronger than what KRAB was able to achieve on the same gene (0.6 ± 0.1 -fold).

KDM5A, 5B and 5C are JARID family HKDMs and catalyse the removal of H3K4 methylation [287]. As such, these enzymes are usually associated with transcriptional repression, which concurs with the results obtained in this experiment. All three significantly repressed *TFRC* (0.61 ± 0.03 - to 0.7 ± 0.1 -fold) and KDM5A, as well as 5C, downregulated *EPCAM*, too (0.6 ± 0.2 - to 0.79 ± 0.04 -fold). The already very weakly expressed *LIN28A* and *ASCL1*, however, were not affected by any of the HKDMs, which is not surprising, as both genes should have low or no levels of H3K4me3 in their associated nucleosomes. As mentioned in chapter 2.5, it is an interesting observation that despite their similar base expression levels, *EPCAM* and *LIN28A* show very different behaviour towards certain EDs, suggesting separate underlying programming mechanisms for both genes. To define these programs, it would be necessary to determine the methylation state of the gene as well as present histone modifications.

As the acetylation of histone plays a crucial role in chromatin structure, I also investigated the effects of HDACs on transcription. As discussed for KATs, by shrouding the positive charge of the lysine residue histone acetylation represents a far simpler mechanism than most methylation marks, which often depend on whole protein complexes to influence transcription. It hence comes to no surprise that most of the HDACs are among the strongest repressive EDs in this set-up. Both catalytic domains of HDAC6 significantly downregulated *TFRC* (CD1: 0.76 ± 0.01 -fold, CD2: 0.69 ± 0.01 -fold) and *EPCAM* (CD1: 0.7 ± 0.1 -fold, CD2: 0.87 ± 0.04 -fold). CD1 even repressed

ASCL1 to 0.38 ± 0.07 -fold, which represents one of the strongest repressor domains for this gene. The strongest repressive effect of any ED in this experiment, however, was achieved by HDAC7, which reduced *ASCL1* expression to 0.27 ± 0.08 -fold. Although repression by deacetylation is plausible, it is not expected from histones located in the *ASCL1* gene to contain many acetyl-marks. ChIP analysis of the respective promoter region would be of great benefit to explain these effects. Nevertheless, the observations regarding transcriptional repression by HDACs correlates with studies presented by other groups [288,289].

Variable effects

KDM1A or LSD1 can demethylate H3K4 and H3K9 and as such exhibits repressive or activational function, depending on the context [43,290,291]. In this experiment, it was able to upregulate *LIN28A* significantly by 4.8 ± 0.6 -fold and simultaneously repress *ASCL1* down to 0.7 ± 0.1 -fold activity, demonstrating the extreme versatility of this epigenetic effector molecule. The repressive effect on *ASCL1* is comparable to what I achieved by recruiting KRAB with the EPIC'RSIPR system (0.6 ± 0.1 -fold), and the activation of *LIN28A* is among the strongest upregulating effects in this experiment. A study performed by Kearns et al. utilized a dCas9-LSD1 fusion construct and recruited it to the promoter of *Tbx3* [167]. In their set-up, the ED achieved strong repression of the target gene, which was also comparable to their dCas9-KRAB construct, supporting the observations I made in my experiment. Although many investigations on the function of KDM1A have been conducted, the complexity of this ED is still not fully understood, and I could not find another study showing transcriptional activation by epigenetic editing with KDM1A.

DOT1L (KMT4) is the only known HKMT to methylate H3K79, as *DOT1L* knock-out cells show global loss of H3K79me1, me2 and me3 [292,293]. DOT1L deficient mouse embryos are not able to survive, and they show symptoms of aneuploidy, telomere elongation and alterations in heterochromatin formation [293]. Aberrant H3K79 methylation in known tumour suppressor genes further plays an important role in leukaemia formation [294,295]. Additionally, converting the methylation states from me1 to me3 positively correlates with increasing transcriptional activity [292], further associating DOT1L with active transcription. Interestingly, in this thesis, DOT1L recruitment displayed rather ambiguous effects in the targeted loci. While it only slightly upregulated *LIN28A* (2.0 ± 0.4 -fold), *CXCR4* (1.6 ± 0.3 -fold) and *ASCL1* (1.4 ± 0.3 -fold),

the enzyme repressed the transcription of *TFRC* (0.81 ± 0.03 -fold) and *EPCAM* (0.69 ± 0.07 -fold) significantly. In accordance with these results, a study by Steger et al. demonstrated that increased H3K79 methylation does not always correlate with increased mRNA levels [292]. However, researchers around Cano-Rodriguez et al. recruited dCas9-DOT1L to the promoters of *EPCAM* and *PLOD2* and achieved weak upregulation of both genes by about 1.5- to 3-fold 2 dpt [166], which contradicts the results in my experiment. The repression of *EPCAM* with DOT1L is as strong as with the KRAB control (0.63 ± 0.10 -fold) in the EPIC'RSIPR set-up. Six independent samples varied from 0.48 and 0.87, suggesting good reliability of the results. One major difference between Cano-Rodriguez's and my experiments are the used gRNAs. Although both set-ups target the promoter region, different spacer sequences were utilized (compare Supplementary Data **Table 0-1** and [219]), which might influence the results.

3.1.5 Outlook

While the parallel modulation of endogenous gene expression provides valuable knowledge about the functionality of the recruited EDs, analysing the transcriptional state of a gene before and after perturbation does not answer the question of how these EDs achieve their respective effects. Unfortunately, I was not able to obtain more data with other techniques due to time limitations caused by the recent COVID-19 pandemic. I initially planned to employ ChIP to determine the composition and alteration of histone marks. Analysing which modifications are present at a specific locus and how effector recruitment changes their composition would be invaluable for understanding how a gene is regulated and how the respective effector influences this regulation. This would be especially interesting for gene pairs like *LIN28A* and *EPCAM*, which show similar expression levels yet different behaviour towards perturbation. Besides the analysis of the histone code, it would be very beneficial to investigate the DNA methylation state as this mark is essential for transcriptional regulation. Bisulfite sequencing is a state of the art technology to determine the methylation status of a DNA sequence efficiently and if or how it changes after perturbation. The combined data from transcription states, histone modifications and DNA methylation would allow us to understand the regulatory program behind the expression of an individual gene. This would potentially enable us to find the best ED or ED combinations to efficiently

reprogram the respective gene and, using multiplexed epigenetic editing, and maybe even whole cells.

Open questions

Some EDs tested in the large scale experiment in chapter 2.5 did not affect target expression significantly, which can have several possible reasons. Firstly, the respective ED might be functionally inactive. All constructs used in this thesis are either based on full-length proteins or functional subunits for which crystal structures are provided online [39,212]. Although these domains should be functional, it might be that fusing them to the AB causes sterical issues or problems with proper folding. To determine that all constructs are expressed properly, total protein could be extracted from transfected cells which can then be validated by SDS-PAGE and Western blotting (see chapter 4.3.7). This was initially done for the controls (VPR, KRAB, dCas9), but not yet for all of the EDs. A second explanation would be, that the respective domains are active and functional, but the introduced mark does not cause a change in transcription, a hypothesis which was recently published for histone methylation [296]. To investigate this possibility, ChIP data could help clarify if the respective enzyme even modifies the histone code as suspected. A third explanation was discussed in detail for specific domains already (see chapter 2.5). Some epigenetic molecules have a strong substrate specificity and might not modify the respective histones if said substrate is not present. Other EDs might be blocked by antagonistic marks, preventing them from introducing any effect. ChIP data would here as well offer a valuable insight into how these molecules function and help to determine optimal ED/target combinations. Additionally, as discussed in chapter 3.1.2, depending on the respective functionality, each ED might require different amounts of time to achieve significant effects on transcription (e.g. VPR versus mTET3del1). It would hence be necessary to perform time-course experiments for each recruited ED to determine the exact extent of their capabilities and functionalities.

Epigenetic stability and additive or synergistic effects

Altering the epigenetic state of a gene does not necessarily mean that this change is stable. I demonstrated in this thesis for *EPCAM* that VPR-induced activation quickly decreases between 3 dpt (24 ± 2 -fold) and 5 dpt (3.3 ± 0.4 -fold). Observing only a single time point does hence not yield any information about the stability of the underlying marks. Several groups have investigated the stability of epigenetic modifications with

contradicting results [166,168,171,172]. It would be invaluable for our understanding of the epigenetic code to determine under which conditions and what context a specific epigenetic modification is stable and how it can be altered. To do so, large-scale time-course experiments would be necessary to test the influence of dozens of EDs or ED combinations at several time points. EPIC-RISPR would be a perfectly suited tool for such experiments.

Recruiting VPR to a repressed gene, completely overrides its transcriptional programming and forces the target to be strongly overexpressed. Interestingly, none of the more than 60 recruited EDs achieved similar results in this experiment, although almost all of them have been shown to exhibit epigenetic editing capabilities. Even the strongest effect achieved by any of the recruited EDs in this experiment was at least about three times lower than the effect of VPR on the respective target gene (10 ± 2 -fold compared to 35 ± 5 -fold by KDM2B F-box on *LIN28A*). This finding suggests certain robustness or stability of the epigenetic program. It would indeed be fatal for a cell if an epigenetic modifier could alter the transcriptional programming of a nearby gene simply by being accidentally mistargeted. This hypothesis is supported by a study which demonstrates that at least histone methylation is not sufficient for target gene repression [296]. VPR, on the other hand, is a synthetic construct which consists of three independent transcriptional activator domains (VP64, p65 and Rta). It achieves far stronger effects than any of the singular domains it consists of by completely overwriting the epigenetic code [109]. This kind of synergistic behaviour has been observed in several other cases by various fusion constructs [103,168,169] or co-recruitment of EDs [166,175] and has been discussed in detail in chapter 3.1.2. The data obtained in this thesis revealed efficient transcriptional modifiers from each group of the investigated epigenetic modifier families: HKMTs, HKDMs, HATs and HDACs. This data set constitutes a strong basis of information which can be used to identify the most interesting and promising candidates for co-recruitment experiments. The co-recruitment of EDs that showed similar effects on the transcriptional state of the same target could potentially result in additive activational or repressive effects.

For more effective activation, the surprising and strong activation effects of the KDM2B F-box domain on *LIN28A* could be combined with KAT3B, which also activated the gene significantly. Both EDs follow completely different mechanisms to achieve this effect. While KAT3B acetylates lysine residues and hence opens up the chromatin, the F-box domain recruits large complexes which might be involved in transcriptional

elongation. These two mechanisms display great potential for combined effects. Depending on the context, KDM1A could further be utilized as a third domain to demethylate H3K9me2, potentially allowing the activation of heterochromatic regions. Even though KDM6A and KDM6B did not show significant effects on their own, removing H3K27me3 and hence polycomb repression could lead to interesting combinatorial effects with the other mentioned domains. Furthermore, PRDM2 displayed transcriptional activation potential for four of the five targeted genes. Its possible function as a transcription factor makes it an interesting candidate to be combined with the other positive regulators described in this paragraph. Finally, combining KDM2B F-box with KDM2A and KDM4A on *LIN28A* could lead to an even stronger activation, if H3K36me was indeed involved in the repression of this gene. While KDM4A catalyses the removal H3K36me3, KDM2B is specific for H3K36me2 and KDM2A for H3K36me1. Combined, the three HKDMs could consecutively prepare the optimal substrate for each other.

Another interesting ED-combination would be KDM5A-5C with HDAC7. Simultaneously removing the H3K4me3 mark and acetyl groups from the promoter region of an active target gene could potentially be far more effective than eliminating only one type of modification. These effects might be enhanced even further by also adding KDM8, which can remove H3K9me1, a mark associated with active transcription.

Epigenetic research and medicine

Modern epigenetic research usually utilizes cell- or tissue culture experiments to determine epigenetic changes in a cell population. Although this approach offers great reliability as it includes stochastic differences in the respective population, recent advancements in single-cell analysis could potentially allow us to investigate the distinct effects of epigenetic editing from a new perspective [297].

Dysregulation of the epigenetic code and hence gene expression can lead to pathological phenotypes, potentially causing a great variety of diseases, including various cancer types [298]. Due to our lack of knowledge about the exact causalities, epigenetic drugs usually target a broad range of effectors, typically inhibiting whole enzyme classes or domains associated with the respective dysregulated mark [299]. Additionally, these drugs are typically not localized but rather systemic and therefore influence the whole organism. Logically, this can lead to a broad range of severe side-

effects. A deeper understanding of the epigenetic system will hence be crucial for developing specific epigenetic drugs.

3.1.6 Conclusion

EPIC'RSIPR represents a powerful tool for functional synthetic epigenetics. Utilizing the platform's incredible modularity, I created libraries of more than one hundred modified gRNAs and EDs. I then used the components of these libraries for highly parallel epigenetic editing and chromatin imaging in several distinct ways. Firstly, I successfully recruited one ED to five genetic promoter regions simultaneously and significantly up- or downregulated the transcription of the respective genes. Secondly, I demonstrated the modulation of the transcriptional states of three genes in parallel using different EDs. I further showed that the employed EDs are only modulating their assigned target gene. Thirdly, I efficiently recruited three different fluorophores to the same genomic location using three differently modified gRNAs. Although I was not able to successfully modify transcription with these constructs, I confirmed that EPIC'RSIPR gRNAs are stable with up to nine aptamers at their 3' end, while other groups struggle to stabilize gRNAs with more than two. I further introduced the functionality of inducibility to the EPIC'RSIPR platform. Utilizing ligand-dependent dimerization of FKBP and FRB, I created a tuneable ON-switch which allows modulation of transcription between 0% and 45% compared to the respective uninducible EPIC'RSIPR construct. I also designed an OFF-switch by employing the nuclease deficient Csy4 H29A mutant, which can be reactivated using imidazole. This switch allows transcriptional modulation between 5% and 100% compared to the uninduced state in a matter of minutes. All these functionalities make EPIC'RSIPR extremely versatile in its applications, which distinguishes the platform from other published epigenetic editing tools. To demonstrate how easily the system can be applied, I performed a large-scale experiment in which I perturbed the transcription of five differently expressed genes with more than 60 different epigenetic enzymes or domains. This experiment yielded valuable knowledge about the functionality of each of the recruited EDs and suggested interesting hypotheses about the functionality of some epigenetic marks and regulatory mechanisms. Besides the effects discussed in detail in chapter 3.1.4, the experiment allowed four major conclusions:

1. The majority of the single EDs tested in this thesis can significantly change the transcriptional state of a gene 3 dpt, however, not to an extent as achieved by synthetic transcriptional modulators which completely overwrite the epigenetic program (e.g. VPR).
2. The transcriptional state of a gene negatively correlates with its potential for being activated by transcriptional modulators (e.g. VPR). This means that a repressed gene can be activated far stronger compared to its base expression level than an already active gene.
3. The base expression level of a gene does not predetermine the general effect of an epigenetic ED on its transcriptional state, suggesting different mechanisms of epigenetic regulation that cause a similar transcriptional state (*EPCAM* versus *LIN28A*).
4. One epigenetic ED might facilitate transcriptional activation and repression on different targets simultaneously, depending on the context, even though the respective domain has so far only been associated with either up- or downregulation (e.g. KDM2B F-box).

As many of the recruited EDs exhibited other effects on transcription than anticipated, it is very hard to draw direct conclusions for each effector molecule individually without more information. Our libraries contain several enzyme classes including arginine methyltransferases, arginine deaminases, poly(ADP-ribose)transferases, kinases, phosphatases and ubiquitinases, which yet need to be analysed in a similar experiment. I am confident that another large-scale set-up to investigate the functionalities of these epigenetic modifiers will yield interesting observations and allow conclusions about the regulatory mechanisms behind the respective transcriptional states. The here performed large-scale experiment demonstrated the impressive potential of the EPIC'RISPR platform, especially considering that this is only one of its possible applications.

3.2 CsyTag for transgene expression control

3.2.1 Scientific context

The applications of CsyTag and the imidazole inducible H29A mutant of Csy4 are not limited to controlling endogenous gene expression via the EPIC'RISPR system or other

CRISPR/Cas9-based methods [279]. The system has also been established for the analysis of *in vitro* RNA-protein interaction [210], and, in accordance with the results in this thesis, as a degradation tag for transgene control [205,206,300]. For the latter, CsyTag is implemented in the 5' UTR of transiently transfected reporter constructs, which allows the Csy4-mediated cleavage of the 5' m7G-cap. Most similar to the set-up in this thesis is hereby the study conducted by Borchardt et al. [206]. In contrast to their results, I was able to reduce the reporter signal to undetectable levels using confocal fluorescence imaging, when cleaving CsyTag with Csy4 WT. Borchardt et al. achieved similar results only when implementing CsyTag in-frame right after the AUG start codon, but not when it was placed in the 5' UTR where it does not interfere with the coding part of the transgene. However, both studies use different reporter systems to quantify their effects. While Borchardt et al. utilize a *Gaussia* luciferase reporter construct, I determined GFP signals using FACS analysis. When using Csy4 WT, I was able to reduce the reporter signal to 0.5% GFP-positive cells, which corresponds to a reduction down to 1.6% compared to Csy4-untreated cells. In comparison, Borchardt et al. achieved signal reductions to 2.7% to 4.2% (~24-fold to ~37-fold decrease), respectively, which are relatively similar results considering that both studies were conducted using different reporter systems under the control of different promoters (CBA and pCAG) that might cause different expression levels. Furthermore, both Borchardt et al. and I demonstrated that removing the 3' poly-A tail by introducing CsyTag in the 3' UTR is less effective and yields a lower signal reduction. As the 5' m7G-cap is essential for binding the translation initiation factor eIF4E to the mRNA [207], removing the cap might actively prohibit an effective translation, while mRNA without 3' poly-A tail might still be translated. Furthermore, the orientation of CsyTag might play a crucial role. The cleavage of CsyTag occurs on the lower 3' end of the stem-loop (**Figure 2-26A**), allowing Csy4 to stay bound to the upstream loop-sequence. This means that the 3' end of a reporter lacking the 3' poly-A might still harbour Csy4 which could sterically prevent degradation by exonucleases.

In contrast to other studies that use CsyTag as a tag for transgene expression control, I was further able to modulate and fine-tune the achieved signal reduction effects by rescuing the activity of a Csy4 H29A mutant with imidazole. Using various concentrations of this chemical, I decreased the reporter signal down to $39 \pm 1\%$ (32 mM) of the initial signal in a dose-dependent manner. I also demonstrate that multiple CsyTag sequences in the 5' UTR amplify the achieved reduction. Using four

CsyTags, I reduced the GFP signal to $3.2 \pm 0.2\%$ (~31-fold decrease) compared to untreated cells, which is similar to the effects I observed with Csy4 WT (1.6%) and what Borchardt et al. reported for their Csy4 wildtype construct. Additionally, I could modulate expression levels in a dose-dependent manner in four different constructs (1x to 4xC-Gd2).

3.2.2 Outlook

Theoretically, CsyTag could also be applied to the EPIC'RISPR system. If integrated into the UTR of the plasmids of dCas9, gRNA and/or ED, the respective components could be turned-off on the mRNA level. Consequently, the expression of the endogenous target would be influenced, as well. However, digesting freshly synthesized mRNA does not shut down the already expressed components. The system would still be functional until the respective molecules degrade over time. To control endogenous gene expression, it is hence more efficient to utilize CsyTag directly in the gRNA, where the EPIC'RISPR system can be turned off within 5 min (see chapter 2.4.2).

Expression control of transgenes is usually conducted with inducible promoter systems [301–303]. Nature developed a great variety of such regulatory systems that allow positive and negative feedback loops as well as logic gate systems *in vivo* [2,304–306]. These systems, however, typically require multiple and sometimes large components while CsyTag is a small, 28 nt long sequence that can be easily inserted in most RNA molecules without interfering with protein expression before induction. Furthermore, promoter-based control of gene expression cannot remove already transcribed mRNA, which leads to a slower trigger response of the system. One other regulatory system that utilizes RNA degradation is RNA interference [307], which could be a viable alternative to the CsyTag system as it does not require any alteration of the desired construct, and ligand-regulated versions that can be tuned exist as well [308]. Another alternative which works on the protein level, however, could be a furin cleavage site [309]. It allows an amino acid sequence-specific cleavage of the reporter protein by the endoprotease furin.

3.2.3 Conclusion

The major advantage of CsyTag over most control systems described in chapter 3.2.2 is its versatility. In this thesis, I demonstrate its applicability in expression control of transgenes and endogenous genes and show dose-dependent fine-tuning in both cases. Furthermore, I apply the system for reversible chromatin labelling and show the quick response time of fewer than five minutes. In conclusion, CsyTag and the imidazole-inducible Csy4 H29A can be considered as a valuable addition to already existing techniques which allows dose-dependent effect modulation. The results of this project are part of a publication which is currently in submission [211].

3.3 Bio-On-Magnetic-Beads (BOMB)

Epigenetic research is notorious for its high sample count, which makes it challenging for small groups with low funding to compete in this field. The costs and time that need to be invested are simply not feasible for many laboratories. With the development of the BOMB platform, we negated these disadvantages and made any nucleic acid-based research affordable for anyone. In this chapter, I will discuss in detail the benefits that BOMB provides, present the research it has already impacted and debate possible future improvements and possibilities for the platform.

3.3.1 Benefits of BOMB

Time

The main reason for us to develop the BOMB platform was the immediate need to process huge numbers of samples. For the experiments described in chapter 2.5, Sven Höhn, Luca Schelle, and I isolated RNA from HEK293 cells, prepared cDNA, and then amplified and cloned over 100 human genes in less than a month. The big benefit of magnetic beads over column-based systems is the lack of any centrifugation step, which makes the BOMB protocols suitable for most common liquid handling systems. Even if such a machine is not available, the processing of multiple batches of 96-well plates at a time by a single researcher is easily doable. Especially with large sample numbers, these protocols are much faster than commercial column-based kits.

Community and collaborations

Besides the various protocols, the BOMB platform also offers a web forum³ for the discussion and exchange of open-source protocols with a continuously growing number of members. Up to today, several community members contributed with protocols for alternative bead systems and new designs for magnetic racks. The need for such a community, and open-science in general, has become obvious in the light of the recent COVID-19 pandemic and the accompanying issues with the supply chains. Since then, the BOMB community has become a hub for collaborations⁴ that work on alternatives for the commercial kits that are used in the official testing protocols. In collaboration with the Canterbury Health Laboratories, our colleagues in New Zealand were able to modify the original extraction protocols to reliably test for COVID-19 [310], an astonishing achievement with a most likely huge impact on one of the most challenging problems of our time. Since then, other labs, too, have established our beads for testing protocols [311–315]. In another COVID-19 related collaboration with the start-up OpenCell, in which I was involved in myself, we developed mobile emergency testing facilities with a theoretical capacity of over 2,400 tests per day [315]. In this project, we participated in the adaptation of the BOMB protocols for the automated extraction of viral RNA using liquid handling robots.

Costs

The most obvious advantage of the BOMB platform is the financial aspect for the user. Depending on the type of sample and the supplier, commercial purification methods range from 1 to 14 € per sample (**Table 3-1**). Considering that the above-described synthesis protocol in chapter 2.1.1 yields MNPs for more than 40,000 isolations of DNA, RNA or TNA, the price for a BOMB nucleic acid extraction ranges from 0.05 € to 0.32 € per sample. This represents a price reduction compared to commercial alternatives from 10- to 120-fold.

The costs in **Table 3-1** have been calculated based on the high-yield protocols described in the original publication [237]. However, these extractions were performed

³ <https://bomb.bio/forums>

⁴ <https://bomb.bio/forums/forum/protocols/covid-19-rna-extraction/>

with a sample volume of 1.2 to 2.2 ml (deep-well plates). Thanks to the modularity of the BOMB protocols, scaling down the sample sizes to e.g. 0.2 ml microtiter plates is easily doable and reduces the costs per sample even further. Additionally, depending on the desired downstream processing (e.g. PCR) it might not be necessary to perform a digest with RNase A or DNase I, eliminating another huge cost factor. Even though obtaining a magnetic rack is a one-time investment, several hundred € for one rack can be financially challenging for some institutions. The protocols for self-crafting magnetic racks we provide, reduce these costs to negligible levels.

Table 3-1: Cost comparison of BOMB protocols and commercial solutions. The table and table description were taken from Oberacker et al. [237] and have been slightly modified. Total costs per 96 samples were calculated, considering plastics, solvents, and enzymes like DNase I. These costs were omitted for the kit content replacement cost column. The price range for the commercial kits depicts list prices offered by three major distributors.

Protocol	BOMB	BOMB	Commercial	Price
	complete (96 samples)	kit content (96 samples)	kit (96 samples)	advantage (kit/BOMB total)
Clean-up	5 €	0.34 €	155 – 218 €	31 – 43x
Gel extraction	11 €	1.5 €	155 – 228 €	14 – 21x
Plasmid DNA extraction	14 €	5 €	146 – 191 €	10 – 14x
TNA extraction from cells	11 €	2.2 €	349 – 1332 €	32 – 121x
TNA extraction from tissues	23 €	6.5 €	349 – 471 €	15 – 20x
gDNA extraction	12 €	2.2 €	183 – 354 €	15 – 30x
Total RNA extraction	32 €	2.5 €	440 – 592 €	14 – 18x
Bisulfite conversion	16 €	5 €	207 – 655 €	13 – 41x

The financial aspect of any experiment is an important issue for many institutions and research groups that run on a tight budget. This is especially true for any research done in developing countries where funding is generally limited, as well as for educational institutions like schools or universities. BOMB allows these users to follow up on ideas and projects that were unthinkable with commercial solutions.

3.3.2 Outlook

So far, we developed a range of protocols for the purification and manipulation of nucleic acids using bead systems with a ferrite core encapsulated in either silica- or

carboxyl-groups. However, many other variations for the preparation of MNPs exist that exhibit different properties, and there is also the possibility to further functionalize the beads with chemical groups. To tap the full potential of the magnetic bead system, other ways of bead-synthesis and modification need to be considered.

Alternative ferrites, such as cobalt ferrite (CoFe_2O_4) can be used for core particle synthesis and silica-coating, too [316], which leads to more flexibility in the choice of raw material. Additionally, the ferrite core can be covered with a layer of PS [231], which is also the basis of commercially used Sera-Mag SpeedBeads (Merck)⁵. The main advantage of these cores compared to exposed ferrites is that the PS is not prone to oxidation, leading to more stable beads.

Other coating reactions might also be useful for certain applications. A modified version of the Stöber method for silica-coating uses water as a solvent, yielding an extremely thin silica-layer (< 5 nm), which changes the behaviour of the resulting MNPs in solution [317]. The respective beads have a bigger surface to volume ratio, and therefore a higher capacity per mass. However, this comes at the cost of a longer settling time when exposed to a magnet. Additionally, water is not only a regenerative source of material, but it is also less expensive than alcohol, making protocols like this a perfect expansion to the BOMB platform. The coat-layer itself can also be further modified, adding new and specific chemical properties to the MNPs. A silica-surface could e.g. be equipped with highly reactive sulfonic acid [318] or with a layer of carboxyl-groups [316]. A Carboxyl-layer, either on the ferrite core or a layer of silica, can be further coupled with nitrilotriacetic acid. These magnetic His-traps can be used for affinity purification of biotin-tagged proteins [316,319]. Similarly, streptavidin can be added to a carboxyl-layer via tetrazole chemistry, allowing the isolation of biotin-tagged molecules [320]. Theoretically, such a biotin-modified molecule could be the imidazole inducible Csy4 H29A (see chapter 2.4.2). The enzyme could capture CsyTag-modified RNA molecules, a technique which has been demonstrated on an avidin-agarose matrix for the capture of ribonucleoprotein complexes [210]. Potentially, imidazole could also be used to rescue the catalytic activity of a Cas9 nicking mutant as well.

⁵ <https://www.sigmaaldrich.com/technical-documents/articles/biology/sera-mag-and-sera-mag-speedbeads-magnetic-particles.html>

One of the two endonucleolytic centres of Cas9 is an HNH domain. It has been shown before, that the catalytic activity H to A and H to G mutants can be rescued with imidazole in such domains [256]. Under the right conditions, a dCas9 H840A mutant could theoretically bind double-stranded DNA in a sequence-specific manner and release the captured molecule after unspecific nucleic acids have been washed away. In another study, PS-based MNPs have been functionalized with azide-groups [321], allowing the modulation of the surface with a variety of alkyne-linked molecules by copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC). Such an azide-functionalisation can also be achieved on silica-surfaces [322] that could potentially allow sequence-specific isolation of nucleic acids using alkyne-tagged probes [323].

Besides the intensively described silica- and carboxyl-layers, there are several other possibilities for coating MNPs, such as gold-surfaces on PS-based beads [324] and the silver-coating of ferrite cores [325] using a microwave for bead synthesis. Each functionalization method provides unique chemical properties for specific applications, displaying that we only scratched the surface of MNP-biochemistry.

3.3.3 Conclusion

In conclusion, BOMB is an open-source platform for the isolation, purification, and manipulation of nucleic acids, which shows great potential for a leading role in open-science. We already built a community which is joined by more and more scientists every day. The collaboration group is successfully involved in multiple highly relevant projects, including COVID-19 related research, and plans to further extend its scientific horizon by adding new functionalities and possible applications to the magnetic bead system. The BOMB protocols have already been adopted by numerous laboratories [311–315,326–345]. The fact that some of these are highly renowned research groups with an abundance of funding [327,332] demonstrates that BOMB is not simply cheaper than commercially offered products but even delivers better quality.

4 Material and Methods

4.1 Plasmids

4.1.1 Cloning

For any cloning process in this thesis, the desired sequences were amplified in a PCR reaction using 2 U Q5 High-Fidelity DNA Polymerase, 1x Q5 Reaction Buffer, 0.2 mM dNTPs (all NEB), 1 ng template DNA and 0.5 μ M of forward and reverse primer in a 50 μ l reaction. The reaction was then incubated, as described in **Table 4-1**.

Table 4-1: Incubation step for PCR.

Temperature [°C]	Time [s]	
95	60	
95	15	}30x
60-68	15	
72	10	
72	60	
4	hold	

To analyse PCR amplification products, 5 μ l sample was combined with 1x loading dye (**Table 4-2**) and loaded on 1% Biozym LE Agarose gels (Biozym) with 1x Tris-phosphate-EDTA (TPE) buffer (**Table 4-2**). The gel was exposed to 5 V/cm in electrophoresis chambers (Serva Electrophoresis) for 20 to 40 minutes, and DNA bands were visualized with the Quantum system (PEQLAB Biotechnologie GmbH). 2 μ l GeneRuler DNA Ladder (Thermo Scientific) served as a reference for fragment size.

After PCR amplification, samples were treated with 40 U DpnI (NEB, 37 °C, 2 h) and purified using either the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) or BOMB protocols #4.1 or #4.4 (depending on the fragment size, see chapter 2.1.2) with home-made silica-beads (see chapter 2.1.1). Concentrations were determined by UV-spectroscopy (NanoDrop 1000 spectrophotometer, Thermo Scientific). For ligation cloning, the PCR product was phosphorylated using T4 polynucleotide kinase and ligated into the desired vector backbone with T4 ligase (both NEB) according to the manufacturer's instructions. Alternatively, the purified PCR product was combined with

the vector in a Gibson assembly reaction (NEBuilder HiFi DNA Assembly Master Mix, NEB) in a molar ratio of 3:1. If not stated otherwise, ligation- or Gibson-products were transformed into electrocompetent TOP 10 *E. coli* and plasmids were isolated, validated and quantified as described in chapter 4.3.1.

Table 4-2: Buffer recipes 1. 6x DNA loading dye and 10xTPE buffer.

6x DNA loading dye		10x TPE	
Tris-HCl	10 mM	Tris base	900 mM
Bromphenol blue	0.03%	EDTA	20 mM
Xylene blue	0.03%	H ₃ PO ₄	to pH8.0
Glycerol	60%		
EDTA	60 mM		
GelRed	4%		

4.1.2 gRNAs and EDs

To produce the gRNA- and ED-plasmids used in this thesis, the hierarchical cloning strategies developed in chapters 2.2.1 and 2.2.2 were utilized. For detailed instructions, see the respective chapters.

4.1.3 Csy4 plasmids

The ORF of wild-type Csy4 was obtained from pHMGWA-Pa14Csy4 (Addgene #41091) [204] using pHAGE-EFS_Csy4_NcoI_f and pHAGE_EFS_Csy4_SpeI_r (Supplementary Data **Table 0-10**) as primers. The PCR product was then cloned into pHAGE-EFS-PCP-GGS-FLAG-nls (derived from pHAGE-PCP-EFS-3xGFPnls, Addgene #75385) [180] by replacing the PCP domain, yielding pHAGE-EFS-Csy4-GGS-FLAG-nls (Csy4 WT, Supplementary Data **Table 0-11**) in a Gibson assembly reaction.

The mutations H29A and H29G were introduced by amplifying Csy4 WT with Csy4_H29A_HindIII_f and Csy4_H29A_HindIII_r or Csy4_H29G_HindIII_f and Csy4_H29G_HindIII_r (Supplementary Data **Table 0-10**) and using the construct in a Gibson assembly reaction. The resulting vectors pHAGE-EFS-Csy4H29A-GGS-

FLAG-nls and pHAGE-EFS-Csy4H29G-GGS-FLAG-nls are further referred to as Csy4 H29A and Csy4 H29G, respectively.

To remove the NLS together with the GGS-linker and the FLAG-tag, Csy4 WT and Csy4 H29A were amplified using NLS_out_f and NLS_out_r (Supplementary Data **Table 0-10** and **Table 0-10**) and re-ligated.

4.1.4 CsyTag-modified plasmids

CsyTag was introduced in either 5'- and/or 3'-UTR of pCAG-GFPd2 (Addgene #14760) [302]. For this, unique restriction sites in either of the two regions were chosen. Pairs of overlapping oligonucleotides containing the Csy4 restriction site from *Streptococcus thermophilus* (CsyTag_NotI_f and _r, CsyTag_KpnI_f and _r, or CsyTag_EcoRI_f and _r, Supplementary Data **Table 0-10**) were annealed, elongated using Q5 DNA Polymerase (NEB) and purified. The vector was linearized with the respective restriction enzyme and combined via Gibson assembly. In this manner the constructs pCAG-CsyTag-GFPd2 (C-Gd2) and pCAG-GFPd2-CsyTag (Gd2-C) were assembled, and subsequently pCAG-2xCsyTag-GFPd2 (2xC-Gd2) and pCAG-CsyTag-GFPd2-CsyTag (C-Gd2-C).

For pCAG-3xCsyTag-GFPd2 (3xC-Gd2) and pCAG-4xCsyTag-GFPd2 (4xC-Gd2) the oligonucleotides multiCsyTag_I1_f and _r, multiCsyTag_I2_f and _r, multiCsyTag_I2_f and multiCsyTag_end, as well as multiCsyTag_I3_f and multiCsyTag_end (Supplementary Data **Table 0-10**) were annealed, elongated and purified as described before, leading to the inserts I1, I2, I2_end and I3_end. I1 and I2_end (for 3xC-Gd2) or I1, I2 and I3_end (for 4xC-Gd2) were combined in an equimolar ratio and amplified in a PCR reaction. The resulting products were separated and extracted from an agarose gel (BOMB protocol #4.3) [237] and inserted into linearized C-Gd2 via Gibson assembly. The exact position of each CsyTag in the UTR has been summarized in **Table 0-11** of the Supplementary Data section.

4.1.5 PuroR

For experiments using GFP as part of the analysis, the pEGFP-puro (Addgene #45561) [346] plasmid could not be utilized for the enrichment of transfected cells. The GFP ORF was hence removed by PCR amplification using PuroR_GFPout_f and _r

(Supplementary Data **Table 0-9**). The product was digested with DpnI, ligated, transformed and verified as described in chapter 4.1.1.

4.2 Cell culture

4.2.1 HEK293 culture conditions

Basic culturing conditions

HEK293 cells were cultured at 37 °C and 5% CO₂ using Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 1x Penicillin-Streptomycin (200 U/ml penicillin, 0.2 mg/ml streptomycin) and 8 mM L-glutamine (all Sigma-Aldrich). For passaging, cells were washed with Dulbecco's Phosphate Buffered Saline (DPBS) without (w/o) calcium chloride (CaCl₂) and magnesium chloride (MgCl₂), detached with Trypsin-EDTA (all Sigma-Aldrich) and seeded at dilutions of 1:2 to 1:12 in T175 flasks (Sarstedt).

Harvesting

To harvest cells after an experiment, the wells were washed twice with DPBS (w/o MgCl₂ and CaCl₂) and detached with Trypsin-EDTA. To stop the digest, trypsin was diluted 1:3 with DMEM supplemented as described above. The cells were dispersed by gentle pipetting and transferred to 1.2 ml MegaBlock 96 deep well plates (Sarstedt). The plates were centrifuged (500 g, 5 min), the supernatant was removed, and the pellets were frozen at -80 °C.

4.2.2 Transfections

Basic procedure

One day before a transfection experiment, cells were detached and seeded at 40% density in 6-well plates or 8-well tissue culture chambers (Sarstedt). DNA (**Table 4-3**) and PEI MAX (40 kDa, Polysciences) were combined in a mass ratio of 1:3 with DMEM (w/o supplements) as a basis. The amount of hereby used plasmid DNA varied between 400 ng (Chromatin labelling, 8-well) and 2 µg (other experiments, 6-well). The mixture was incubated at room temperature (RT) for 20 min and added dropwise to the culture. 15 hours post-transfection (hpt), the cells were washed with DPBS (with MgCl₂ and CaCl₂, Sigma-Aldrich) and fresh DMEM (supplemented) was added. The medium was exchanged every day until the cells were harvested. To verify the expression of the transfected plasmids and to enrich transfected cells, 5% of each transfection mix consisted of either pEGFP-puro (Addgene #45561) [346] or pPuroR (see chapter 4.1.5). By adding 2.5 µg/ml puromycin (InvivoGen) to the culture medium, starting 1 dpt, untransfected cells die and are removed during washing. In experiments where samples expressed EGFP, pcDNA3-mRuby2 (Addgene #40260) [347] was used instead of pEGFP-puro to ensure a visual expression control that can be observed using microscopy (EVOS FLoid Cell Imaging Station, ThermoFisher). In all controls, the remaining DNA was topped-up with pUC19 plasmid (Addgene #50005) [348].

Table 4-3: Transfection schematics with the fraction of plasmids in % of total DNA. Brackets in plasmid names indicate that the respective feature is optional. The exact construct is described in each of the corresponding chapters.

		Plasmids								
		Chapter	pCMB-target-gRNA-(Aptamers) (equimolar mix)	pHAGE-TO-dCas9-(ED)-nls	pHAGE-EFS-Aptamerbinder- ED/Fluorophore/Dimerizer-nls	Dimerizer-ED-nls	pCAG-(1-4xCsyTag)-GFPd2-(CsyTag)	pHAGE-EFS-Csy4-(nls) (WT, H29A, H29G)	pEGFP-puro or pcDNA3-mRuby2	pPuroR
EPIC'RISPR	samples	2.3 2.5	15.0%	15.0%	65.0%				5.0%	
	Controls (dCas9-ED)	2.3	75.0%	20.0%					5.0%	
	Inducible (FKBP/FRB)	2.4.1	15.0%	15.0%	32.5%	32.5%			5.0%	
	Inducible (Csy4)	2.4.2	15.0%	15.0%	32.5%			32.5%	5.0%	
	Chromatin labelling	2.3	74.0%	15.0%	6.0%					5.0%
	Chromatin labelling (Csy4)	2.4.2	74.0%	15.0%	3.0%			3.0%		5.0%
	Transgene expression	2.6					10.0%	80.0%	5.0%	5.0%

Small molecules

In the experiments described in chapter 2.4, small molecules were added to the culture medium. If not stated differently, the respective chemical was added to the culture medium with the first medium change after transfection (together with puromycin) and at each subsequent day of the experiment until the cells were harvested.

For switch-ON experiments featuring FKBP and FRB (see chapter 2.4.1), the rapamycin analogue A/C Heterodimerizer (TaKaRa) was used in different concentrations. Depending on the experiment, 5 to 2000 nM of the chemical was dissolved in supplemented DMEM (including puromycin) and added to the cells.

In Csy4 related experiments (see chapters 2.4.2 and 2.6) a stock solution of 1 M imidazole was prepared. For this, the required amount of crystalline imidazole (Roth) was dissolved in DPBS (w/o MgCl₂ and CaCl₂) supplemented with 40 µM Phenol Red (Honeywell), the pH was adjusted to 7.0, and the solution was sterile filtered. As for A/C Heterodimerizer, dilutions in DMEM were prepared before adding the medium to the cells. These dilution contained concentrations between 0.5 and 32 mM of imidazole.

4.3 Downstream processing and analysis

4.3.1 DNA

Plasmid DNA was isolated from 2 ml TOP 10 *E. coli* cultures using either Plasmid Purification Kit (Macherey-Nagel) or BOMB protocols #5.1 or #5.2 (see chapter 2.1.3). For the BOMB protocols, the cultures were pelleted, and the supernatant was discarded. The cells were resuspended in a Tris-EDTA buffer and lysed by alkaline lysis with sodium hydroxide. After the lysis was complete, the pH was reduced by adding a potassium acetate-buffer with (#5.1) or without guanidine-hydrochloride (#5.2). Cellular debris and gDNA were then removed either by centrifugation (#5.1) or by a physical pulldown using silica-coated magnetic beads (#5.2). In either case, the supernatant was transferred to a fresh vial. To bind plasmid DNA, 1 V of 96% EtOH and a fresh load of magnetic beads were added to the samples. For #5.2 the EtOH was exchanged for an EtOH-based buffer which also contained guanidine-hydrochloride. Remaining contaminants were subsequently removed by two to four wash steps with a Tris-HCl/80% EtOH-based buffer, and plasmid DNA was eluted

using ddH_2O and quantified by UV-spectroscopy (NanoDrop 1000 spectrophotometer). After the clean-up, DNA was analysed by restriction digest. For this, 100 ng plasmid DNA was combined with 0.2 μl of the respective restriction enzyme (NEB) and a fitting buffer in 1x concentration. The reaction was incubated for 15 min to 1 h at the recommended temperature and analysed using agarose gel electrophoresis as described above (see chapter 4.1.1). Plasmids were further sent for Sanger sequencing (LGC Genomics) to ensure sequence integrity.

4.3.2 RNA

RNA was isolated from frozen HEK293 cell pellets utilizing the TRIZOL-based BOMB protocol #8.1 (see chapter 2.1.4). For this, harvested HEK293 cell pellets were dissolved in TRI reagent (BOMB protocol #B.1) [237] by vigorous shaking (Microtiter plate orbital shaker, IKA MS 3 basic). To capture the nucleic acids, 1 V of RNA binding buffer was added as well as silica-coated magnetic beads (1:10 dilution). After binding the supernatant was removed and the samples were washed four times with 90% EtOH. The beads were dried (50 °C, 5 min) to evaporate remaining EtOH, and DNA was digested under constant shaking using a DNaseI reaction mix (37 °C, 30-60 min). 4 V of RNA binding buffer were added to bind the intact RNA to the beads, and the digested DNA was removed by four additional wash steps with 90% EtOH. The beads were dried (50 °C, 30 min) and RNA was eluted with RNase free ddH_2O . To ensure a high quality of the extracted RNA, each sample was analysed by UV-spectroscopy. Additionally, I verified the RNA integrity of approximately 10% of samples from each batch by agarose gel electrophoresis. For this, 1x Tris-Acetate-EDTA (TAE, **Table 4-4**) buffer was used instead of TPE. To denature the RNA, the samples were combined with 1x RNA loading dye (**Table 4-4**) and incubated at 65 °C for 1 min.

Table 4-4: Buffer recipes 2. 50x TAE buffer and 2x RNA loading dye.

50x TAE buffer		2x RNA loading dye	
Tris base	40 mM	Formamide	95%
EDTA	2 mM	SDS	0.02%
Acetic acid	to pH8.5	Bromphenol blue	0.02%
		Xylene blue	0.01%
		EDTA	1 mM

4.3.3 RT-qPCR

Reverse Transcription

For analysis of expression levels with RT-qPCR, the extracted RNA had to be reverse transcribed into cDNA. For this, 200-500 ng RNA was combined on ice with random octamers (0.25 μ M, Thermo), dT₁₈-oligonucleotides (0.75 μ M, Thermo, not for the results presented in **Figure 2-20F**), dNTPs (0.5 μ M each, Thermo), RNasin (0.2 U/ μ l, Promega), 1x RT buffer and 10 U/ μ l M-MuLV-RT (both Enzymatics) in 20 or 40 μ l reactions. The mixture was then incubated, as described in **Table 4-5**.

RT-qPCR

After the reverse transcription, the cDNA was diluted (1:2). 2 μ l of this dilution was used as a template for RT-qPCR and combined with the respective primers (0.25 μ M each, Supplementary Data **Table 0-12**) as well as 1x ORA sSEE q-PCR Green Rox L Mix (highQu). If not stated otherwise, biological triplicates and two or three technical replicates were prepared and analysed for each sample. Furthermore, each sample plate contained a non-template control as well as a non-reverse transcriptase control for quality control.

Table 4-5: Incubation steps and programs for reverse transcription and RT-qPCR.

Reverse transcription		RT-qPCR	
Temperature [°C]	Time [min]	Temperature [°C]	Time [s]
42	1	95	180
25	5	95	10
37	5	61	30
42	60	measure }30x	
90	10	95	10
4	hold	65	
		↓ (Δ 0.5 °C), measure (5 s)	
		95	

The reactions were then incubated in a LightCycler 96 Instrument (Roche) using the program described in **Table 4-5** and Cq values were extracted via the LC96-software (Roche). To analyse relative transcription levels, the human housekeeping gene *REEP5* was chosen as a reference due to its consistent and strong expression in HEK293 cells [349]. Utilizing the Pfaffl-method [350], the surplus of target transcript compared to the REEP5 reference was determined by equation (1).

$$\text{relative expression} = \frac{E_{\text{target}}^{\Delta Cq_{\text{target}}(\text{control-sample})}}{E_{\text{reference}}^{\Delta Cq_{\text{reference}}(\text{control-sample})}} \quad (1)$$

E hereby represents the PCR efficiency of the primer pairs, which were determined by 10-fold dilution series, while ΔCq describes the difference between the control (pUC19-transfected cells) and the sample of interest. Using this equation, I calculated the relative expression for each replicate relative to the other primer pair of the same sample as well as the averages of both primer pairs for the control. I then took the mean of the four to six values obtained for each biological replicate. The relative expression of the three biological replicates was used to calculate a mean value and an SEM for each sample.

4.3.4 FACS analysis

Cells were prepared for FACS analysis 3 dpt as described in chapter 4.2.1. After centrifugation, the pellet was washed twice with DPBS (w/o MgCl₂ and CaCl₂) and gently resuspended until no visible clusters could be observed. The cell suspension was seeded in 96-well flat-bottom cell culture plates (Sigma-Aldrich) and analysed using the MACSQuant VYB instrument (MACS Miltenyi Biotec). After identifying and gating HEK293 cells (**Figure 2-27A**, left scatter plot) by forward- and side-scatter, aggregated cell clusters from the analysis (**Figure 2-27A**, middle scatter plot). Like this, at least 10.000 living cells were identified. For these cells, the fluorescence signal was determined for GFP and mRuby2 (control for transfection efficiency), and four sectors (Q1 to Q4, **Figure 2-27A**, right scatter plot) were defined. To determine the changes in fluorescence, the signal of double-positive cells (Q2) between the samples was compared.

4.3.5 LSM

For live-cell imaging, cells were either seeded and transfected in 8-well tissue culture chambers which are suitable for microscopy of living cells in a culture medium, or in normal 6-well plates. In the case of the latter, cells were transferred to 8-well tissue culture chambers 15 hpt. 24 hpt the fluorescence signal was detected with an LSM 710 Zeiss confocal microscope (Zeiss, Plan-Apochromat 63 × /1.40 Oil DIC M27 objective, XL-LSM 710 S1 chamber for controlling temperature as well as the CO₂ level). For the experiments described in chapter 2.4.2, 1 M imidazole stock solution was added dropwise to a final concentration of 32 mM, and the signal was monitored after 0, 1, 2.5, 5, 10, 20 and 30 min.

4.3.6 Total protein extraction

To extract total protein from transfected cells, the harvested pellets were washed twice with precooled DPBS (-20°C, w/o MgCl₂ and CaCl₂, centrifugation: 2500g, 5 min, 4°C) and resuspended by gentle pipetting in 150 µl Radioimmunoprecipitation assay (RIPA) buffer (150 mM NaCl, 5 mM EDTA pH8.0, 50 mM TRIS pH8.0, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 1x protease inhibitor cocktail). To increase the yield, the suspension was sonicated (50% pulse, 30 sec) and shaken on a microtiter plate orbital shaker (Merck, 15 min, 250 rpm). The cellular debris was pulled down by centrifugation

(2500 g, 15 min, 4 °C) and the supernatant was transferred to a new vial. Protein concentration was determined by utilizing the Pierce BCA Protein Assay Kit (Thermo Scientific) and an EnSpire multimode plate reader (PerkinElmer).

4.3.7 SDS-PAGE and Western blot

2 µg total protein was separated by SDS-PAGE (12% PAA) and electro-blotted onto a nitrocellulose membrane (GE Healthcare) [351]. Under constant, gentle shaking, the membrane was first blocked in PBS-T (50 mM K₂HPO₄, 150 mM NaCl, and 0.1% Tween 20, pH 7.4) with 5% milk powder for 1 h and then incubated over-night at 4 °C in either Anti-GFP IgG (from mouse, 1:2000 in PBS-T, Sigma-Aldrich) or Anti-beta Actin IgG (from rabbit, 1:5000 in PBS-T, abcam). After three more washes with PBS-T, the membrane was incubated for 1 h at RT with the secondary antibody, which was either ECL anti-mouse IgG-HRP (1:5000, GE Healthcare) or ECL anti-rabbit IgG-HRP (1:5000, GE Healthcare). The membrane was washed three more times with PBS-T and eventually covered with Western Lightning Plus-ECL (PerkinElmer) substrate. A High-performance chemiluminescence film (GE Healthcare) was exposed for 10 sec to 1 min to the membrane, and the signal was analysed using the ImageQuant TL software (GE Healthcare).

4.3.8 *in vitro* transcription of gRNA

To obtain the *in vitro* transcribed RNAs used to validate the BOMB clean-up protocol (**Figure 2-3F**) for small ncRNA, dsDNA templates of gRNAs were synthesized by PCR with a forward primer coding for a T7 promoter and the desired spacer sequence and reverse primer matching the template plasmid. The PCR was performed according to the user manual utilizing Taq Polymerase (NEB) and the program shown in **Table 4-6**. The product was purified (see chapter 2.1.2), and the success of the reaction was controlled via agarose gel electrophoresis (see chapter 4.1.1).

For the *in vitro* transcription reaction, 20 µl purified PCR-product was combined with 1x Transcription Buffer, 0.6 U/µl T7 RNA Polymerase and 2 mM of each ATP, UTP, GTP and CTP (all Thermo Scientific) as well as 0.2 U/µl RNasin (Promega) in a reaction volume between 100 µl and 400 µl. After the reaction (37 °C, 4 h), the DNA was degraded by DNaseI (NEB, 2 U/µl, 37 °C, 2 h). The enzyme was deactivated using EDTA (2.5 µM, pH8.0) at high temperatures (65 °C, 15 min). gRNA was precipitated

with 0.5 M ammonium acetate and 3 V 96% EtOH (-20 °C, overnight). The solution was pelleted, washed twice with 80% EtOH (21500 g, 4 °C, 15 min) and eluted in RNase free water. Quantification was performed by via UV-spectroscopy (NanoDrop 1000 spectrophotometer), and the integrity of the purified gRNA was determined by SDS-PAGE (see chapter 4.3.7) using a 12% denaturing urea gel.

Table 4-6: PCR program for gRNA template synthesis.

Temperature [°C]	Time [s]	
95	60	
95	30	}30x
60	25	
68	30	
68	60	
4	hold	

Bibliography

1. Waddington CH. *The Strategy of the Genes: A Discussion of some Aspects of theoretical Biology*. George Allen and Unwin; 1957.
2. Müller-Hill B. *The lac operon: a short history of a genetic paradigm*. de Gruyter; 1996.
3. Huang S. Inverse relationship between genetic diversity and epigenetic complexity. *Nat Preced*. 2009; 1–15. doi:10.1038/npre.2009.1751.2
4. Berger SL, Kouzarides T, Shiekhata R, Shilatifard A. An operational definition of epigenetics. *Genes Dev*. 2009;23: 781–783. doi:10.1101/gad.1787609
5. Heard E, Martienssen RA. Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell*. 2014;157: 95–109. doi:10.1016/j.cell.2014.02.045
6. Miska EA, Ferguson-Smith AC. Transgenerational inheritance: Models and mechanisms of non-DNA sequence-based inheritance. *Science* (80-). 2016;354: 59–63. doi:10.1126/science.aaf4945
7. Horsthemke B. A critical view on transgenerational epigenetic inheritance in humans. *Nat Commun*. 2018;9: 1–4. doi:10.1038/s41467-018-05445-5
8. Peterson CL, Laniel MA. Histones and histone modifications. *Curr Biol*. 2004;14: 546–551. doi:10.1016/j.cub.2004.07.007
9. Berger SL. Histone modifications in transcriptional regulation. *Curr Opin Genet Dev*. 2002;12: 142–148. doi:10.1016/S0959-437X(02)00279-4
10. Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13: 484–492. doi:10.1038/nrg3230
11. Weichenhan D, Plass C. The evolving epigenome. *Hum Mol Genet*. 2013;22: 1–6. doi:10.1093/hmg/ddt348
12. Roberts TC, Morris K V., Weinberg MS. Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics*. 2014;9: 13–20. doi:10.4161/epi.26700
13. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997;389: 251–260. doi:10.1038/38444
14. Schones DE, Zhao K. Genome-wide approaches to studying chromatin modifications. *Nat Rev Genet*. 2008;9: 179–191. doi:10.1038/nrg2270
15. Henneman B, van Emmerik C, van Ingen H, Dame RT. Structure and function of archaeal histones. *PLoS Genet*. 2018;14: 1–21. doi:10.1371/journal.pgen.1007582
16. Rosa S, Shaw P. Insights into chromatin structure and dynamics in plants. *Biology (Basel)*. 2013;2: 1378–1410. doi:10.3390/biology2041378
17. Khorasanizadeh S. The Nucleosome: From Genomic Organization to Genomic Regulation. *Cell*. 2004;116: 259–272. doi:10.1016/S0092-8674(04)00044-3
18. Chen T, Dent SYR. Chromatin modifiers and remodellers: Regulators of cellular differentiation. *Nat Rev Genet*. 2014;15: 93–106. doi:10.1038/nrg3607
19. Lee BM, Mahadevan LC. Stability of histone modifications across mammalian

- genomes: Implications for “epigenetic” marking. *J Cell Biochem.* 2009;108: 22–34. doi:10.1002/jcb.22250
20. Jenuwein T, Allis CD. Translating the histone code. *Science* (80-). 2001;293: 1074–1080. doi:10.1126/science.1063127
 21. Alam H, Gu B, Lee MG. Histone methylation modifiers in cellular signaling pathways. *Cell Mol Life Sci.* 2015;72: 4577–4592. doi:10.1007/s00018-015-2023-y
 22. Nimura K, Ura K, Kaneda Y. Histone methyltransferases: Regulation of transcription and contribution to human disease. *J Mol Med.* 2010;88: 1213–1220. doi:10.1007/s00109-010-0668-4
 23. Mosammamarast N, Shi Y. Reversal of Histone Methylation: Biochemical and Molecular Mechanisms of Histone Demethylases. *Annu Rev Biochem.* 2010;79: 155–179. doi:10.1146/annurev.biochem.78.070907.103946
 24. Nakatani Y. Histone acetylases—versatile players. *Genes to Cells.* 2001;6: 79–86. doi:10.1046/j.1365-2443.2001.00411.x
 25. Kouzarides T. Histone acetylases and deacetylases in cell proliferation. *Curr Opin Genet Dev.* 1999;9: 40–48. doi:10.1016/S0959-437X(99)80006-9
 26. Davie JR, Spencer VA. Control of histone modifications. *J Cell Biochem.* 1999;32/33: 141–148. doi:10.1002/(sici)1097-4644(1999)75:32+<141::aid-jcb17>3.0.co;2-a
 27. Kooistra SM, Helin K. Post-translational modifications: Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Biol.* 2012;13: 297–311. doi:10.1038/nrm3327
 28. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell.* 2007;129: 823–837. doi:10.1016/j.cell.2007.05.009
 29. Ruthenburg AJ, Allis CD, Wysocka J. Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. *Mol Cell.* 2007;25: 15–30. doi:10.1016/j.molcel.2006.12.014
 30. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell.* 2005;120: 169–181. doi:10.1016/j.cell.2005.01.001
 31. Guenther MG, Jenner RG, Chevalier B, Nakamura T, Croce CM, Canaani E, et al. Global and Hox-specific roles for the MLL1 methyltransferase. *Proc Natl Acad Sci.* 2005;102: 8603–8608. doi:10.1073/pnas.0503072102
 32. Herz HM, Mohan M, Garruss AS, Liang K, Takahashi Y, Mickey K, et al. Enhancer-associated H3K4 monomethylation by trithorax-related, the drosophila homolog of mammalian MLL3/MLL4. *Genes Dev.* 2012;26: 2604–2620. doi:10.1101/gad.201327.112
 33. Cheng J, Blum R, Bowman C, Hu D, Shilatifard A, Shen S, et al. A role for H3K4 mono-methylation in gene repression and partitioning of chromatin readers. *Mol Cell.* 2014;53: 979–992. doi:10.1016/j.molcel.2014.02.032
 34. Glaser S, Schaft J, Lubitz S, Vintersten K, van der Hoeven F, Tuftteland KR, et al. Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development.* 2006;133: 1423–1432.

doi:10.1242/dev.02302

35. van Nuland R, Smits AH, Pallaki P, Jansen PWTC, Vermeulen M, Timmers HTM. Quantitative Dissection and Stoichiometry Determination of the Human SET1/MLL Histone Methyltransferase Complexes. *Mol Cell Biol.* 2013;33: 2067–2077. doi:10.1128/mcb.01742-12
36. Mizuguchi G, Tsukiyama T, Wisniewski J, Wu C. Role of nucleosome remodeling factor NURF in transcriptional activation of chromatin. *Mol Cell.* 1997;1: 141–150. doi:10.1016/S1097-2765(00)80015-5
37. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards suite: From gene data mining to disease genome sequence analyses. *Curr Protoc Bioinforma.* 2016;54: 1.30.1-1.30.33. doi:10.1002/cpbi.5
38. Braschi B, Denny P, Gray K, Jones T, Seal R, Tweedie S, et al. Genenames.org: The HGNC and VGNC resources in 2019. *Nucleic Acids Res.* 2019;47: D786–D792. doi:10.1093/nar/gky930
39. Consortium TU. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* 2019;47: D506–D515. doi:10.1093/nar/gky1049
40. Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, et al. Yng1 PHD Finger Binding to H3 Trimethylated at K4 Promotes NuA3 HAT Activity at K14 of H3 and Transcription at a Subset of Targeted ORFs. *Mol Cell.* 2006;24: 785–796. doi:10.1016/j.molcel.2006.10.026
41. Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, et al. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature.* 2006;442: 96–99. doi:10.1038/nature04835
42. Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* 2010;38: 4246–4253. doi:10.1093/nar/gkq147
43. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell.* 2004;119: 941–953. doi:10.1016/j.cell.2004.12.012
44. Di Stefano L, Walker JA, Burgio G, Corona DFV, Mulligan P, Näär AM, et al. Functional antagonism between histone H3K4 demethylases in vivo. *Genes Dev.* 2011;25: 17–28. doi:10.1101/gad.1983711
45. Karmodiya K, Krebs AR, Oulad-Abdelghani M, Kimura H, Tora L. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics.* 2012;13: 1–18. doi:10.1186/1471-2164-13-424
46. Rea S, Eisenhaber F, O’Carroll D, Strahl BD, Sun ZW, Schmid M, et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature.* 2000;406: 593–599. doi:10.1038/35020506
47. Lehnertz B, Ueda Y, Derijck AAHA, Braunschweig U, Perez-Burgos L, Kubicek S, et al. Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. *Curr Biol.* 2003;13: 1192–1200. doi:10.1016/S0960-982(03)00432-9
48. Hashimoto H, Vertino PM, Cheng X. Molecular coupling of DNA methylation and

- histone methylation. *Epigenomics*. 2010;2: 657–669. doi:10.2217/epi.10.44
49. Fujita N, Watanabe S, Ichimura T, Tsuruzoe S, Shinkai Y, Tachibana M, et al. Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. *J Biol Chem*. 2003;278: 24132–24138. doi:10.1074/jbc.M302283200
 50. Rougeulle C, Chaumeil J, Sarma K, Allis CD, Reinberg D, Avner P, et al. Differential Histone H3 Lys-9 and Lys-27 Methylation Profiles on the X Chromosome. *Mol Cell Biol*. 2004;24: 5475–5484. doi:10.1128/mcb.24.12.5475-5484.2004
 51. Escamilla-Del-Arenal M, da Rocha ST, Spruijt CG, Masui O, Renaud O, Smits AH, et al. Cdy1, a New Partner of the Inactive X Chromosome and Potential Reader of H3K27me3 and H3K9me2. *Mol Cell Biol*. 2013;33: 5005–5020. doi:10.1128/mcb.00866-13
 52. Shinkai Y, Tachibana M. H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev*. 2011;25: 781–788. doi:10.1101/gad.2027411
 53. Zhang T, Termanis A, Özkan B, Bao XX, Culley J, de Lima Alves F, et al. G9a/GLP Complex Maintains Imprinted DNA Methylation in Embryonic Stem Cells. *Cell Rep*. 2016;15: 77–85. doi:10.1016/j.celrep.2016.03.007
 54. Yamane K, Toumazou C, Tsukada Y, Erdjument-Bromage H, Tempst P, Wong J, et al. JHDM2A, a JmJc-Containing H3K9 Demethylase, Facilitates Transcription Activation by Androgen Receptor. *Cell*. 2006;125: 483–495. doi:10.1016/j.cell.2006.03.027
 55. Qi HH, Sarkissian M, Hu G, Wang Z, Bhattacharjee A, Gordon DB, et al. The XLMR gene PHF8 encodes a histone H4K20/H3K9 demethylase and regulates zebrafish brain and craniofacial development. *Nature*. 2012;466: 503–507. doi:10.1038/nature09261
 56. Kim J-Y, Kim K-B, Eom GH, Choe N, Kee HJ, Son H-J, et al. KDM3B Is the H3K9 Demethylase Involved in Transcriptional Activation of *lmo2* in Leukemia. *Mol Cell Biol*. 2012;32: 2917–2933. doi:10.1128/mcb.00133-12
 57. Laurent B, Ruitu L, Murn J, Hempel K, Ferrao R, Xiang Y, et al. A Specific LSD1/KDM1A Isoform Regulates Neuronal Differentiation through H3K9 Demethylation. *Mol Cell*. 2015;57: 957–970. doi:10.1016/j.molcel.2015.01.010
 58. Couture JF, Collazo E, Ortiz-Tello PA, Brunzelle JS, Trievel RC. Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nat Struct Mol Biol*. 2007;14: 689–695. doi:10.1038/nsmb1273
 59. Shin S, Janknecht R. Diversity within the JMJD2 histone demethylase family. *Biochem Biophys Res Commun*. 2007;353: 973–977. doi:10.1016/j.bbrc.2006.12.147
 60. Aldiri I, Vetter ML. PRC2 during vertebrate organogenesis: A complex in transition. *Dev Biol*. 2012;367: 91–99. doi:10.1016/j.ydbio.2012.04.030
 61. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev*. 2002;16: 2893–2905. doi:10.1101/gad.1035902
 62. Völkel P, Angrand PO. The control of histone lysine methylation in epigenetic

- regulation. *Biochimie*. 2007;89: 1–20. doi:10.1016/j.biochi.2006.07.009
63. Dietrich N, Lerdrup M, Landt E, Agrawal-Singh S, Bak M, Tommerup N, et al. REST-mediated recruitment of polycomb repressor complexes in mammalian cells. *PLoS Genet*. 2012;8. doi:10.1371/journal.pgen.1002494
 64. Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome Regulation by Polycomb and Trithorax Proteins. *Cell*. 2007;128: 735–745. doi:10.1016/j.cell.2007.02.009
 65. Voigt P, LeRoy G, Drury WJ, Zee BM, Son J, Beck DB, et al. Asymmetrically modified nucleosomes. *Cell*. 2012;151: 181–193. doi:10.1016/j.cell.2012.09.002
 66. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell*. 2006;125: 315–326. doi:10.1016/j.cell.2006.02.041
 67. Ferrari KJ, Scelfo A, Jammula SG, Cuomo A, Barozzi I, Stützer A, et al. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. *Mol Cell*. 2014;53: 49–62. doi:10.1016/j.molcel.2013.10.030
 68. Lee MG, Villa R, Trojer P, Norman J, Yan KP, Reinberg D, et al. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* (80-). 2007;318: 447–450. doi:10.1126/science.1149042
 69. Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, et al. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development*. 2009;136: 3131–3141. doi:10.1242/dev.037127
 70. Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J Biol Chem*. 2005;280: 17732–17736. doi:10.1074/jbc.M500796200
 71. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol*. 2014;13: 115–126. doi:10.1038/nrm3274
 72. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*. 2005;123: 581–592. doi:10.1016/j.cell.2005.10.023
 73. Joshi AA, Struhl K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to pol II elongation. *Mol Cell*. 2005;20: 971–978. doi:10.1016/j.molcel.2005.11.021
 74. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. Regulation of alternative splicing by histone modifications. *Science* (80-). 2010;327: 996–1000. doi:10.1126/science.1184208
 75. Ball MP, Li JB, Gao Y, Lee JH, Leproust EM, Park IH, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol*. 2009;27: 361–368. doi:10.1038/nbt.1533
 76. Lin CH, Li B, Swanson S, Zhang Y, Florens L, Washburn MP, et al. Heterochromatin Protein 1a Stimulates Histone H3 Lysine 36 Demethylation by the Drosophila KDM4A Demethylase. *Mol Cell*. 2008;32: 696–706. doi:10.1016/j.molcel.2008.11.008

77. He J, Kallin EM, Tsukada YI, Zhang Y. The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15Ink4b. *Nat Struct Mol Biol.* 2008;15: 1169–1175. doi:10.1038/nsmb.1499
78. Klose RJ, Zhang Y. Regulation of histone methylation by demethylimination and demethylation. *Nat Rev Mol Cell Biol.* 2007;8: 307–318. doi:10.1038/nrm2143
79. Morris SA, Rao B, Garcia BA, Hake SB, Diaz RL, Shabanowitz J, et al. Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification. *J Biol Chem.* 2007;282: 7631–7640. doi:10.1074/jbc.M607909200
80. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet.* 2008;40: 897–903. doi:10.1038/ng.154.
81. Kuo AJ, Song J, Cheung P, Ishibe-Murakami S, Yamazoe S, Chen JK, et al. ORC1 BAH domain links H4K20 to DNA replication licensing and Meier-Gorlin syndrome. *Nature.* 2005;484: 115–119. doi:10.1038/nature10956
82. Botuyan M V, Lee J, Ward IM, Kim J-E, Thompson JR, Chen J, et al. Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair. *Cell.* 2006;127: 1361–1373. doi:10.1016/j.cell.2006.10.043
83. Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, et al. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* 2004;18: 1251–1262. doi:10.1101/gad.300704.
84. Yang H, Pesavento JJ, Starnes TW, Cryderman DE, Wallrath LL, Kelleher NL, et al. Preferential dimethylation of histone H4 lysine 20 by Suv4-20. *J Biol Chem.* 2008;283: 12085–12092. doi:10.1074/jbc.M707974200
85. Wyrick JJ, Parra MA. The role of histone H2A and H2B post-translational modifications in transcription: A genomic perspective. *Biochim Biophys Acta - Gene Regul Mech.* 2009;1789: 37–44. doi:10.1016/j.bbaggm.2008.07.001
86. Osley MA. Regulation of histone H2A and H2B ubiquitylation. *Briefings Funct Genomics Proteomics.* 2006;5: 179–189. doi:10.1093/bfpg/ell022
87. Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, et al. Histone Deimination Antagonizes Arginine Methylation. *Cell.* 2004;118: 545–553. doi:10.1016/j.cell.2004.08.020
88. Di Lorenzo A, Bedford MT. Histone arginine methylation. *FEBS Lett.* 2011;585: 2024–2031. doi:10.1016/j.febslet.2010.11.010
89. Crump NT, Hazzalin CA, Bowers EM, Alani RM, Cole PA, Mahadevan LC. Dynamic acetylation of all lysine-4 trimethylated histone H3 is evolutionarily conserved and mediated by p300/CBP. *Proc Natl Acad Sci.* 2011;108: 7814–7819. doi:10.1073/pnas.1100099108
90. Lee DY, Hayes JJ, Pruss D, Wolffe AP. A Positive Role for Histone in Acetylation in Transcription Factor Access to Nucleosomal DNA. *Cell.* 1993;72: 73–84. doi:10.1016/0092-8674(93)90051-Q
91. Marmorstein R, Zhou M-M. Writers and Readers of Histone Acetylation: Structure, Mechanism and Inhibition. In: Allis CD, Caparros M-L, Jenuwein T, Reinberg D, editors. *Epigenetics.* 2015. pp. 117–141. doi:10.1101/cshperspect.a018762

92. Zhang W, Bone JR, Edmondson DG, Turner BM, Roth SY. Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of Gcn5p acetyltransferase. *EMBO J.* 1998;12: 3155–3167. doi:10.1093/emboj/17.11.3155
93. Avvakumov N, Côté J. The MYST family of histone acetyltransferases and their intimate links to cancer. *Oncogene.* 2007;26: 5395–5407. doi:10.1038/sj.onc.1210608
94. Torchia J, Glass C, Rosenfeld MG. Co-activators and co-repressors in the integration of transcriptional responses. *Curr Opin Cell Biol.* 1998;10: 373–383. doi:10.1016/S0955-0674(98)80014-8
95. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res.* 2011;21: 564–578. doi:10.1038/cr.2011.42
96. Seto E, Yoshida M. Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. In: Allis CD, Caparros M-L, Jenuwein T, Reinberg D, editors. *Epigenetics.* 2014. pp. 143–168. doi:10.1101/cshperspect.a018713
97. Razin A. CpG methylation, chromatin structure and gene silencing - A three-way connection. *EMBO J.* 1998;17: 4905–4908. doi:10.1093/emboj/17.17.4905
98. Glozak MA, Seto E. Acetylation/deacetylation modulates the stability of DNA replication licensing factor Cdt1. *J Biol Chem.* 2009;284: 11446–11453. doi:10.1074/jbc.M809394200
99. Smith ZD, Meissner A. DNA methylation: Roles in mammalian development. *Nat Rev Genet.* 2013;14: 204–220. doi:10.1038/nrg3354
100. Shen J cheng, Rideout WM, Jones PA. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.* 1994;22: 972–976. doi:10.1093/nar/22.6.972
101. Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr ARW, James KD, Turner DJ, et al. Orphan CpG Islands Identify numerous conserved promoters in the mammalian genome. *PLoS Genet.* 2010;6. doi:10.1371/journal.pgen.1001134
102. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci.* 2006;103: 1412–1417. doi:10.1073/pnas.0510310103
103. Stepper P, Kungulovski G, Jurkowska RZ, Chandra T, Krueger F, Reinhardt R, et al. Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. *Nucleic Acids Res.* 2017;45: 1703–1713. doi:10.1093/nar/gkw1112
104. Hahn MA, Wu X, Li AX, Hahn T, Pfeifer GP. Relationship between gene body DNA methylation and intragenic H3K9ME3 and H3K36ME3 chromatin marks. *PLoS One.* 2011;6. doi:10.1371/journal.pone.0018844
105. Jeltsch A, Jurkowska RZ. New concepts in DNA methylation. *Trends Biochem Sci.* 2014;39: 310–318. doi:10.1016/j.tibs.2014.05.002
106. Ravichandran M, Jurkowska RZ, Jurkowski TP. Target specificity of mammalian DNA methylation and demethylation machinery. *Org Biomol Chem.* 2018;16: 1419–1435. doi:10.1039/c7ob02574b
107. Triezenberg SJ, Kingsbury RC, McKnight SL. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes*

- Dev. 1988;2: 718–729. doi:10.1101/gad.2.6.718
108. Beerli RR, Segal DJ, Dreier B, Barbas CF. Toward controlling gene expression at will: Specific regulation of the *erbB-2/HER-2* promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc Natl Acad Sci.* 1998;95: 14628–14633. doi:10.1073/pnas.95.25.14628
 109. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, Iyer E, et al. Highly-efficient Cas9-mediated transcriptional programming. *Nat Methods.* 2015;12: 326–328. doi:10.1038/nmeth.3312.
 110. Ragoczy T, Miller G. Role of the Epstein-Barr Virus Rta Protein in Activation of Distinct Classes of Viral Lytic Cycle Genes. *J Virol.* 1999;73: 9858–9866. doi:10.1128/jvi.73.12.9858-9866.1999
 111. Mukerjee R, Sawaya BE, Khalili K, Amini S. Association of p65 and C/EBP β with HIV-1 LTR modulates transcription of the viral promoter. *J Cell Biochem.* 2007;100: 1210–1216. doi:10.1002/jcb.21109
 112. Witzgall R, O’Leary E, Leaf A, Onaldi D, Bonventre J V. The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc Natl Acad Sci.* 1994;91: 4514–4518. doi:10.1073/pnas.91.10.4514
 113. Margolin JF, Friedman JR, Meyer WKH, Vissing H, Thiesen - HJ, Rauscher FJ. Kruppel-associated boxes are potent transcriptional repression domains. *Proc Natl Acad Sci.* 1994;91: 4509–4513. doi:10.1073/pnas.91.10.4509
 114. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. *Curr Genomics.* 2013;14: 268–278. doi:10.2174/13892029113149990002
 115. Thakore PI, Black JB, Hilton IB, Gersbach CA. Editing the epigenome: technologies for programmable transcription and epigenetic modulation. *Nat Methods.* 2016;13: 127–137. doi:10.1038/nmeth.3733
 116. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell.* 2013;154: 442–451. doi:10.1016/j.cell.2013.06.044
 117. Eckhardt F, Beck S, Gut IG, Berlin K. Future potential of the Human Epigenome Project. *Expert Review of Molecular Diagnostics.* 2004. doi:10.1586/14737159.4.5.609
 118. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, et al. The NIH roadmap epigenomics mapping consortium. *Nat Biotechnol.* 2010;28: 1045–1048. doi:10.1038/nbt1010-1045
 119. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Wspolczesna Onkol.* 2015;19: A68–A77. doi:10.5114/wo.2014.47136
 120. Bujold D, Morais DA de L, Gauthier C, Côté C, Caron M, Kwan T, et al. The International Human Epigenome Consortium Data Portal. *Cell Syst.* 2016;3: 496–499. doi:10.1016/j.cels.2016.10.019
 121. Fernández JM, de la Torre V, Richardson D, Royo R, Puiggròs M, Moncunill V, et al. The BLUEPRINT Data Analysis Portal. *Cell Syst.* 2016;3: 491–495. doi:10.1016/j.cels.2016.10.021

122. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA elements (ENCODE): Data portal update. *Nucleic Acids Res.* 2017;46: D794–D801. doi:10.1093/nar/gkx1081
123. Therapeutically Applicable Research to Generate Effective Treatments (TARGET). 2018. Available: <https://ocg.cancer.gov/programs/target/>
124. Jurkowski TP, Ravichandran M, Stepper P. Synthetic epigenetics—towards intelligent control of epigenetic states and cell identity. *Clin Epigenetics.* 2015;7: 18. doi:10.1186/s13148-015-0044-x
125. Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature.* 1958. doi:10.1038/182064a0
126. Graf T. Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell.* 2011;9: 504–516. doi:10.1016/j.stem.2011.11.012
127. Takahashi K, Okita K, Nakagawa M, Yamanaka S. Induction of pluripotent stem cells from fibroblast cultures. *Nat Protoc.* 2007; 3081–3089. doi:10.1038/nprot.2007.418
128. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell.* 2007;131: 861–872. doi:10.1016/j.cell.2007.11.019
129. De Groote ML, Verschure PJ, Rots MG. Epigenetic Editing: Targeted rewriting of epigenetic marks to modulate expression of selected target genes. *Nucleic Acids Res.* 2012;40: 10596–10613. doi:10.1093/nar/gks863
130. Wolfe SA, Nekludova L, Pabo CO. DNA Recognition by Cys2His2 Zinc Finger Proteins. *Annu Rev Biophys Biomol Struct.* 2000;29: 183–212. doi:10.1146/annurev.biophys.29.1.183
131. Choo Y, Klug A. Toward a code for the interactions of zinc fingers with DNA : Selection of randomized fingers displayed on phage. *Proc Natl Acad Sci.* 1994;91: 11163–11167. doi:10.1073/pnas.91.23.11163
132. Fu F, Sander JD, Maeder M, Thibodeau-Beganny S, Joung JK, Dobbs D, et al. Zinc Finger Database (ZiFDB): a repository for information on C2H2 zinc fingers and engineered zinc-finger arrays. *Nucleic Acids Res.* 2009;37: 279–283. doi:10.1093/nar/gkn606
133. Sander JD, Yeh J-RJ, Peterson RT, Joung JK. Engineering Zinc Finger Nucleases for Targeted Mutagenesis of Zebrafish. Third Edit. *Methods in Cell Biology.* Elsevier Inc.; 2011. doi:10.1016/B978-0-12-374814-0.00003-3
134. Boch J, Bonas U. Xanthomonas AvrBs3 Family-Type III Effectors : Discovery and Function. *Annu Rev Phytopathol.* 2010;48: 419–436. doi:10.1146/annurev-phyto-080508-081936
135. Scholze H, Boch J. TAL effector-DNA specificity. *Virulence.* 2010;1: 428–432. doi:10.4161/viru.1.5.12863
136. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol.* 2012;30: 460–465. doi:10.1038/nbt.2170
137. Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc.* 2012;7: 171–

192. doi:10.1038/nprot.2011.431
138. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 2014;157: 1262–1278. doi:10.1016/j.cell.2014.05.010
139. Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. *Nat Methods*. 2013;10: 957–963. doi:10.1038/nmeth.2649.Cas9
140. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans -encoded small RNA and host factor RNase III. *Nature*. 2011;471: 602–607. doi:10.1038/nature09886
141. Chylinski K, Rhun A Le, Charpentier E. The tracrRNA and Cas9 families of type II CRISPR- Cas immunity systems. *RNA Biol*. 2013;10: 726–737. doi:10.4161/rna.24321
142. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (80-)*. 2012;337: 816–821. doi:10.1126/science.1225829
143. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Sci Technol Adv Mater*. 2013;339: 819–824. doi:10.1126/science.1231143
144. Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*. 2018;556: 57–63. doi:10.1038/nature26155
145. Jamieson AC, Miller JC, Pabo CO. Drug discovery with engineered zinc-finger proteins. *Nat Rev Drug Discov*. 2003;2: 361–368. doi:10.1038/nrd1087
146. Bochtler M. Structural basis of the TAL effector–DNA interaction. *Biol Chem*. 2012;393: 1055–1066. doi:10.1515/hsz-2012-0164
147. Razin A, Riggs AD. DNA Methylation and Gene Function. *Science (80-)*. 1980;210: 604–610. doi:10.1126/science.6254144
148. Liu Y, Toh H, Sasaki H, Zhang X, Cheng X. An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. *Genes Dev*. 2012;26: 2374–2379. doi:10.1101/gad.202200.112
149. Kubik G, Schmidt MJ, Penner JE, Summerer D. Programmable and highly resolved in vitro detection of 5-methylcytosine by TALEs. *Angew Chemie - Int Ed*. 2014;53: 6002–6006. doi:10.1002/anie.201400436
150. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013;152: 1173–1183. doi:10.1016/j.cell.2013.02.022
151. Kungulovski G, Jeltsch A. Epigenome Editing: State of the Art, Concepts, and Perspectives. *Trends Genet*. 2016;32: 101–113. doi:10.1016/j.tig.2015.12.001
152. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res*. 1996;6: 986–994. doi:10.1101/gr.6.10.986
153. Laird CD, Pleasant ND, Clark AD, Sneed JL, Hassan KMA, Manley NC, et al. Hairpin-bisulfite PCR: Assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc Natl Acad Sci*. 2004;101: 204–209. doi:10.1073/pnas.2536758100

154. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One*. 2010;5: 1–9. doi:10.1371/journal.pone.0008888
155. Huebert DJ, Kamal M, O'Donovan A, Bernstein BE. Genome-wide analysis of histone modifications by ChIP-on-chip. *Methods*. 2006;40: 365–369. doi:10.1016/j.ymeth.2006.07.032
156. Xu GL, Bestor TH. Cytosine methylation targeted to pre-determined sequences. *Nat Genet*. 1997; 376–378. doi:10.1038/ng1297-376
157. Li F, Papworth M, Minczuk M, Rohde C, Zhang Y, Ragozin S, et al. Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes. *Nucleic Acids Res*. 2007;35: 100–112. doi:10.1093/nar/gkl1035
158. Smith AE, Ford KG. Specific targeting of cytosine methylation to DNA sequences in vivo. *Nucleic Acids Res*. 2007;35: 740–754. doi:10.1093/nar/gkl1053
159. Smith AE, Hurd PJ, Bannister AJ, Kouzarides T, Ford KG. Heritable gene repression through the action of a directed DNA methyltransferase at a chromosomal locus. *J Biol Chem*. 2008;283: 9878–9885. doi:10.1074/jbc.M710393200
160. Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, et al. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol*. 2013;31: 1137–1142. doi:10.1038/nbt.2726
161. Chen H, Kazemier HG, De Groote ML, Ruiters MHJ, Xu GL, Rots MG. Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter. *Nucleic Acids Res*. 2014;42: 1563–1574. doi:10.1093/nar/gkt1019
162. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol*. 2014;32: 347–350. doi:10.1038/nbt.2842
163. Hilton IB, D'Ippolito AM, Vockley CM, Thakore, Pratiksha I, Crawford GE, Reddy TE, et al. Epigenome editing by a CRISPR/Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotechnol*. 2015;33: 510–517. doi:10.1038/nbt.3199
164. Vojta A, Dobrinic P, Tadic V, Bockor L, Korac P, Julg B, et al. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res*. 2016;44: 5615–5628. doi:10.1093/nar/gkw159
165. Xu X, Tao Y, Gao X, Zhang L, Li X, Zou W, et al. A CRISPR-based approach for targeted DNA demethylation. *Cell Discov*. 2016;2: 1–12. doi:10.1038/celldisc.2016.9
166. Cano-Rodriguez D, Gjaltema RAF, Jilderda LJ, Jellema P, Dokter-Fokkens J, Ruiters MHJ, et al. Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. *Nat Commun*. 2016. doi:10.1038/ncomms12284
167. Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, et al. Functional annotation of native enhancers with a Cas9 -histone demethylase fusion. *Nat Methods*. 2015;12: 401–403. doi:10.1038/nmeth.3325

168. Amabile A, Migliara A, Capasso P, Biffi M, Cittaro D, Naldini L, et al. Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell*. 2016;167: 219–232. doi:10.1016/j.cell.2016.09.006
169. Yeo NC, Chavez A, Lance-Byrne A, Chan Y, Menn D, Milanova D, et al. An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat Methods*. 2018;15: 611–616. doi:10.1038/s41592-018-0048-5
170. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. 2014;159: 635–646. doi:10.1016/j.cell.2014.09.039
171. Kungulovski G, Nunna S, Thomas M, Zanger UM, Reinhardt R, Jeltsch A. Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. *Epigenetics and Chromatin*. 2015;8: 1–10. doi:10.1186/s13072-015-0002-z
172. Saunderson EA, Stepper P, Gomm JJ, Hoa L, Morgan A, Allen MD, et al. Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. *Nat Commun*. 2017;8: 1–14. doi:10.1038/s41467-017-01078-2
173. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. *Nature*. 1990;346: 818–822. doi:10.1038/346818a0
174. Peabody DS. The RNA binding site of bacteriophage MS2 coat protein. *EMBO J*. 1993;12: 595–600. doi:10.1002/j.1460-2075.1993.tb05691.x
175. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2015;517: 583–588. doi:10.1038/nature14136
176. Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*. 2015;106: 339–350. doi:10.1016/j.cell.2014.11.052
177. Lim F, Peabody DS. RNA recognition site of PP7 coat protein. *Nucleic Acids Res*. 2002;30: 4138–4144. doi:10.1093/nar/gkf552
178. Chattopadhyay S, Hung SC, Stuart AC, Palmer AG, Garcia-Mena J, Das A, et al. Interaction between the phage HK022 Nun protein and the nut RNA of phage λ . *Proc Natl Acad Sci*. 1995;92: 12131–12135. doi:10.1073/pnas.92.26.12131
179. Cheng AW, Jillette N, Lee P, Plaskon D, Fujiwara Y, Wang W, et al. Casilio: A versatile CRISPR-Cas9-Pumilio hybrid for gene regulation and genomic labeling. *Cell Res*. 2016;26: 254–257. doi:10.1038/cr.2016.3
180. Ma H, Tu LC, Naseri A, Huisman M, Zhang S, Grunwald D, et al. Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat Biotechnol*. 2016;34. doi:10.1038/nbt.3526
181. Shimizu-Sato S, Huq E, Tepperman JM, Quail PH. A light-switchable gene promoter system. *Nat Biotechnol*. 2002;20: 1041–1044. doi:10.1038/nbt734
182. Zhang F, Wang LP, Brauner M, Liewald JF, Kay K, Watzke N, et al. Multimodal fast optical interrogation of neural circuitry. *Nature*. 2007;446: 633–639. doi:10.1038/nature05744
183. Yazawa M, Sadaghiani AM, Hsueh B, Dolmetsch RE. Induction of protein-protein interactions in live cells using light. *Nat Biotechnol*. 2009;27: 941–945.

doi:10.1038/nbt.1569

184. Kennedy MJ, Hughes RM, Peteya LA, Schwartz JW, Ehlers MD, Tucker CL. Rapid blue-light-mediated induction of protein interactions in living cells. *Nat Methods*. 2010;7: 973–975. doi:10.1038/nmeth.1524
185. Ye H, Baba MD EI, Peng RW, Fussenegger M. A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice. *Science* (80-). 2011;332: 1565–1668. doi:10.1126/science.1203535
186. Zhang F, Vierock J, Yizhar O, Fenno LE, Tsunoda S, Kianianmomeni A, et al. The microbial opsin family of optogenetic tools. *Cell*. 2011;147: 1446–1457. doi:10.1016/j.cell.2011.12.004
187. Strickland D, Lin Y, Wagner E, Hope CM, Zayner J, Antoniou C, et al. TULIPs: Tunable, light-controlled interacting protein tags for cell biology. *Nat Methods*. 2012;9: 379–384. doi:10.1038/nmeth.1904
188. Zhang K, Cui B. Optogenetic control of intracellular signaling pathways. *Trends Biotech*. 2015;33: 92–100. doi:10.1016/j.tibtech.2014.11.007
189. Lungu OI, Hallett RA, Choi EJ, Aiken MJ, Hahn KM, Kuhlman B. Designing photoswitchable peptides using the AsLOV2 domain. *Chem Biol*. 2012;19: 507–517. doi:10.1016/j.chembiol.2012.02.006
190. Müller K, Engesser R, Metzger S, Schulz S, Kämpf MM, Busacker M, et al. A red/far-red light-responsive bi-stable toggle switch to control gene expression in mammalian cells. *Nucleic Acids Res*. 2013;41. doi:10.1093/nar/gkt002
191. Müller K, Siegel D, Rodriguez Jahnke F, Gerrer K, Wend S, Decker EL, et al. A red light-controlled synthetic gene expression switch for plant systems. *Mol Biosyst*. 2014;10: 1679–1688. doi:10.1039/c3mb70579j
192. Polstein LR, Gersbach CA. Light-inducible spatiotemporal control of gene activation by customizable zinc finger transcription factors. *J Am Chem Soc*. 2012;134: 16480–16483. doi:10.1021/ja3065667
193. Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, et al. Optical control of mammalian endogenous transcription and epigenetic states. *Nature*. 2013;500: 472–476. doi:10.1038/nature12466
194. Lo CL, Choudhury SR, Irudayaraj J, Zhou FC. Epigenetic Editing of *Ascl1* Gene in Neural Stem Cells by Optogenetics. *Sci Rep*. 2017;7: 1–12. doi:10.1038/srep42047
195. Polstein LR, Gersbach CA. A light-inducible CRISPR/Cas9 system for control of endogenous gene activation. *Nat Chem Biol*. 2015;11: 198–200. doi:10.1038/nchembio.1753.A
196. Gao Y, Xiong X, Wong S, Charles EJ, Lim WA, Qi LS. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat Methods*. 2016;13: 1043–1049. doi:10.1038/nmeth.4042
197. Bao Z, Jain S, Jaroenpantaruk V, Zhao H. Orthogonal Genetic Regulation in Human Cells Using Chemically Induced CRISPR/Cas9 Activators. *ACS Synth Biol*. 2017;6: 686–693. doi:10.1021/acssynbio.6b00313
198. Tak YE, Kleinstiver BP, Nuñez JK, Hsu JY, Horng JE, Gong J, et al. Inducible and multiplex gene regulation using CRISPR-Cpf1-based transcription factors. *Nat Methods*. 2017;14: 1163–1166. doi:10.1038/nmeth.4483

199. Chiarella AM, Butler K V., Gryder BE, Lu D, Wang TA, Yu X, et al. Dose-dependent activation of gene expression is achieved using CRISPR and small molecules that recruit endogenous chromatin machinery. *Nat Biotechnol.* 2019. doi:10.1038/s41587-019-0296-7
200. Braun SMG, Kirkland JG, Chory EJ, Husmann D, Calarco JP, Crabtree GR. Rapid and reversible epigenome editing by endogenous chromatin regulators. *Nat Commun.* 2017;8. doi:10.1038/s41467-017-00644-y
201. Banaszynski LA, Liu CW, Wandless TJ. Characterization of the FKBP-rapamycin-FRB ternary complex. *J Am Chem Soc.* 2005;127: 4715–4721. doi:10.1021/ja043277y
202. Choi J, Chen J, Schreiber SL, Clardy J. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science (80-).* 1996;273: 239–242. doi:10.1126/science.273.5272.239
203. Crabtree GR, Schreiber SL. Three-part inventions: Intracellular signaling and induced proximity. *Trends Biochem Sci.* 1996;21: 418–422. doi:10.1016/S0968-0004(96)20027-1
204. Haurwitz RE, Jinek M, Wiedenheft B, Kaihong Z, Doudna JA. Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. *Science (80-).* 2010;329: 1355–1358. doi:10.1126/science.1192272
205. Qi L, Haurwitz RE, Shao W, Doudna JA, Arkin AP. RNA processing enables predictable programming of gene expression. *Nat Biotechnol.* 2012;30: 1002–1006. doi:10.1038/nbt.2355
206. Borchardt EK, Vadoros LA, Huang M, Lackey PE, Marzluff WF, Asokan A. Controlling mRNA stability and translation with the CRISPR endoribonuclease Csy4. *RNA.* 2015;21: 1921–1930. doi:10.1261/rna.051227.115
207. Sonenberg N. mRNA 5' Cap-binding protein eIF4E and control of cell growth. In: Hershey JWB, Mathews MB, Sonenberg N, editors. *Translation Control.* New York: Cold Spring Harbor; 1996. pp. 245–269.
208. Sokhi UK, Das SK, Dasgupta S, Emdad L, Shiang R, DeSalle R, et al. Human Polynucleotide Phosphorylase (hPNPaseold-35): Should I Eat You or Not — That Is the Question? *Adv Cancer Res.* 2013;119: 161–190. doi:https://doi.org/10.1016/B978-0-12-407190-2.00005-8
209. Haurwitz RE, Sternberg SH, Doudna JA. Csy4 relies on an unusual catalytic dyad to position and cleave CRISPR RNA. *EMBO J.* 2012;31: 2824–2832. doi:10.1038/emboj.2012.107
210. Lee HY, Haurwitz RE, Apffel A, Zhou K, Smart B, Wenger CD. RNA – protein analysis using a conditional CRISPR nuclease. *Proc Natl Acad Sci.* 2013;110: 5416–5421. doi:10.1073/pnas.1302807110
211. Oberacker P, Stepper P, Jurkowski TP. CsyTag: An inducible, multipurpose RNA-tag for tuneable control of gene expression. submitted. 2021.
212. Berman HM, Westbrook JD, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. *Nucleic Acids Res.* 2000;28: 235–242. doi:10.1093/nar/28.1.235
213. Sun X, Cheng G, Hao M, Zheng J, Zhou X, Zhang J, et al. CXCL12 / CXCR4 / CXCR7 chemokine axis and cancer progression. *Cancer Metastasis Rev.*

- 2010;29: 709–722. doi:10.1007/s10555-010-9256-x
214. Susek KH, Karvouni M, Alici E, Lundqvist A. The Role of CXC Chemokine Receptors 1-4 on Immune Cells in the Tumor Microenvironment. *Front Immunol.* 2018;9: 2159. doi:10.3389/fimmu.2018.02159
 215. Cabrero-De Las Heras S, Martínez-Balibrea E. CXC family of chemokines as prognostic or predictive biomarkers and possible drug targets in colorectal cancer. *World J Gastroenterol.* 2018;24: 4738–4749. doi:10.3748/wjg.v24.i42.4738
 216. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science (80-).* 1996;272: 872–877. doi:10.1126/science.272.5263.872
 217. Münz M, Kieu C, Mack B, Schmitt B, Zeidler R, Gires O. The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation. *Oncogene.* 2004;23: 5748–5758. doi:10.1038/sj.onc.1207610
 218. Terris B, Cavard C, Perret C. EpCAM, a new marker for cancer stem cells in hepatocellular carcinoma. *J Hepatol.* 2010;52: 280–281. doi:10.1016/j.jhep.2009.10.026
 219. Van Der Gun BTF, Huisman C, Stolzenburg S, Kazemier HG, Ruiters MHJ, Blancafort P, et al. Bidirectional modulation of endogenous EpCAM expression to unravel its function in ovarian cancer. *Br J Cancer.* 2013;108: 881–886. doi:10.1038/bjc.2013.45
 220. Nunna S, Reinhardt R, Ragozin S, Jeltsch A. Targeted Methylation of the Epithelial Cell Adhesion Molecule (EpCAM) Promoter to Silence Its Expression in Ovarian Cancer Cells. *PLoS One.* 2014;9: e87703. doi:10.1371/journal.pone.0087703
 221. Aisen P. Transferrin receptor 1. *Int J Biochem Cell Biol.* 2004;36: 2137–2143. doi:10.1016/j.biocel.2004.02.007
 222. Haddad E, Moura IC, Arcos-Fajardo M, Macher MA, Baudouin V, Alberti C, et al. Enhanced expression of the CD71 mesangial IgA1 receptor in Berger disease and Henoch-Schönlein nephritis: Association between CD71 expression and IgA deposits. *J Am Soc Nephrol.* 2003;14: 327–337. doi:10.1097/01.ASN.0000046961.04917.83
 223. Shen Y, Li X, Dong D, Zhang B, Xue Y, Shang P. Transferrin receptor 1 in cancer: a new sight for cancer therapy. *Am J Cancer Res.* 2018;8: 916–931. Available: <http://www.ncbi.nlm.nih.gov/pubmed/30034931> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6048407>
 224. Liu Q, Wang M, Hu Y, Xing H, Chen X, Zhang Y, et al. Significance of CD71 expression by flow cytometry in diagnosis of acute leukemia. *Leuk Lymphoma.* 2014;55: 892–898. doi:10.3109/10428194.2013.819100
 225. Chouchane M, Costa MR. Instructing neuronal identity during CNS development and astroglial-lineage reprogramming: Roles of NEUROG2 and ASCL1. *Brain Res.* 2019;1705: 66–74. doi:10.1016/j.brainres.2018.02.045
 226. Chavez A, Tuttle M, Pruitt BW, Ewen-Campen B, Chari R, Ter-Ovanesyan D, et al. Comparison of Cas9 activators in multiple species. *Nat Methods.* 2016;13: 563–567. doi:10.1038/nmeth.3871

227. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science* (80-). 2007;318. doi:10.1126/science.1151526
228. Wang T, Wang G, Hao D, Liu X, Wang D, Ning N, et al. Aberrant regulation of the LIN28A/LIN28B and let-7 loop in human malignant tumors and its effects on the hallmarks of cancer. *Mol Cancer*. 2015;14: 1–13. doi:10.1186/s12943-015-0402-5
229. Hawkins TL, O'connor-Morin T, Roy A, Santillan C. DNA purification and isolation using a solid-phase. *Nucleic Acids Res*. 1994;22: 4543–4544. doi:10.1093/nar/22.21.4543
230. Deangelis MM, Wang DG, Hawkins TL. Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Res*. 1995;23: 4742–4743. doi:10.1093/nar/23.22.4742
231. Caruso F, Susha AS, Giersig M, Möhwald H. Magnetic Core–Shell Particles: Preparation of Magnetite Multilayers on Polymer Latex Microspheres. *Adv Mater*. 1999;11: 950–953. doi:10.1002/(sici)1521-4095(199908)11:11<950::aid-adma950>3.3.co;2-k
232. Choi J, Kim JC, Lee YB, Kim IS, Park YK, Hur NH. Fabrication of silica-coated magnetic nanoparticles with highly photoluminescent lanthanide probes. *Chem Commun*. 2007; 1644–1646. doi:10.1039/b617608a
233. Singh D, McMillan JM, Liu XM, Vishwasrao HM, Kabanov A V., Sokolsky-Papkov M, et al. Formulation design facilitates magnetic nanoparticle delivery to diseased cells and tissues. *Nanomedicine*. 2014;9: 469–485. doi:10.2217/nnm.14.4
234. Majidi S, Zeinali Sehrig F, Farkhani SM, Soleymani Goloujeh M, Akbarzadeh A. Current methods for synthesis of magnetic nanoparticles. *Artif cells, nanomedicine, Biotechnol*. 2014; 1–13. doi:10.3109/21691401.2014.982802
235. Stöber W, Fink A, Bohn E. Controlled growth of monodisperse silica spheres in the micron size range. *J Colloid Interface Sci*. 1968;26: 62–69. doi:10.1016/0021-9797(68)90272-5
236. Aggarwal A, Saxena R, Wang B, Caneba GT. Studies of the polymerization of methacrylic acid via free-radical retrograde precipitation polymerization process. *J Appl Polym Sci*. 1996;62: 2039–2051. doi:10.1002/(SICI)1097-4628(19961219)62:12<2039::AID-APP7>3.0.CO;2-I
237. Oberacker P, Stepper P, Bond DM, Hoehn S, Focken J, Meyer V, et al. Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation. *PLoS Biol*. 2019;17. doi:10.1371/journal.pbio.3000107
238. Oberacker P, Stepper P, Bond D, Hipp K, Hore TA, Jurkowski TP. Simple Synthesis of Functionalized Paramagnetic Beads for Nucleic Acid Purification and Manipulation. *Bio-Protocol*. 2019;9: 1–10. doi:10.21769/bioprotoc.3394
239. Kim JW, Kim LU, Kim CK. Size control of silica nanoparticles and their surface treatment for fabrication of dental nanocomposites. *Biomacromolecules*. 2007;8: 215–222. doi:10.1021/bm060560b
240. Vogelstein B, Gillespie D. Preparative and analytical purification of DNA from agarose. *Proc Natl Acad Sci*. 1979;76: 615–619. doi:10.1073/pnas.76.2.615

241. Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM, van der Noordaa J. Rapid and simple method for purification of nucleic acids. *J Clin Microbiol.* 1990;28: 495–503. doi:10.1556/AMicr.58.2011.1.7
242. Boom R, Sol CJ, Salimans M, Jansen CL, Wertheim-van Dillen PM, van der Noordaa J. Rapid and simple method for purification of nucleic acids. *J Clin Microbiol.* 1990;28: 495–503. doi:10.1556/AMicr.58.2011.1.7
243. Marko MA, Chipperfield R, Birnboim HC. A procedure for the large-scale isolation of highly purified plasmid DNA using alkaline extraction and binding to glass powder. *Anal Biochem.* 1982;121: 382–387. doi:10.1016/0003-2697(82)90497-3
244. Borodina TA, Lehrach H, Soldatov A V. DNA purification on homemade silica spin-columns. *Anal Biochem.* 2003;321: 135–137. doi:10.1016/S0003-2697(03)00403-2
245. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium extraction by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem.* 1987;162: 156–159. doi:10.1016/0003-2697(87)90021-2
246. Chomczynski P, Sacchi N. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: Twenty-something years on. *Nat Protoc.* 2006;1: 581–585. doi:10.1038/nprot.2006.83
247. Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One.* 2008. doi:10.1371/journal.pone.0003647
248. Engler C, Gruetzner R, Kandzia R, Marillonnet S. Golden gate shuffling: A one-pot DNA shuffling method based on type II restriction enzymes. *PLoS One.* 2009. doi:10.1371/journal.pone.0005553
249. Kearns NA, Genga RMJ, Enameh MS, Garber M, Wolfe SA, Maehr R. Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Co Biol.* 2014;141: 219–223. doi:10.1242/dev.103341
250. Ravichandran M. Biochemical Characterisation of TET DNA Hydroxylases. 2017. doi:http://dx.doi.org/10.18419/opus-9194
251. Ma H, Naseri A, Reyes-Gutierrez P, Wolfe SA, Zhang S, Pederson T. Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc Natl Acad Sci.* 2015. doi:10.1073/pnas.1420024112
252. Heard E, Clerc P, Avner P. X-chromosome inactivation in mammals. *Annu Rev Genet.* 1997;31: 571–610. doi:10.1146/annurev.genet.31.1.571
253. Lucchesi JC, Kelly WG, Panning B. Chromatin Remodeling in Dosage Compensation. *Annu Rev Genet.* 2005;39: 615–651. doi:10.1146/annurev.genet.39.073003.094210
254. Peters J. The role of genomic imprinting in biology and disease: An expanding view. *Nat Rev Genet.* 2014;15: 517–530. doi:10.1038/nrg3766
255. Brown EJ, Albers MW, Bum Shin T, Ichikawa K, Keith CT, Lane WS, et al. A mammalian protein targeted by G1-arresting rapamycin-receptor complex. *Nature.* 1994. doi:10.1038/369756a0
256. Midon M, Gimadutdinow O, Meiss G, Friedhoff P, Pingoud A. Chemical Rescue

- of Active Site Mutants of *S. pneumoniae* Surface Endonuclease EndA and Other Nucleases of the HNH Family by Imidazole. *ChemBioChem*. 2012; 713–721. doi:10.1002/cbic.201100775
257. Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev*. 2003;17: 896–911. doi:10.1101/gad.252103
 258. Buyse IM, Shao G, Huang S. The retinoblastoma protein binds to RIZ, a zinc-finger protein that shares an epitope with the adenovirus E1A protein. *Proc Natl Acad Sci*. 1995;92: 4467–4471. doi:10.1073/pnas.92.10.4467
 259. Manzur KL, Farooq A, Zeng L, Plotnikova O, Koch AW, Sachchidanand, et al. A dimeric viral SET domain methyltransferase specific to Lys27 of histone H3. *Nat Struct Biol*. 2003;10: 187–196. doi:10.1038/nsb898
 260. Wood K, Tellier M, Murphy S. DOT1L and H3K79 methylation in transcription and genomic stability. *Biomolecules*. 2018;8: 1–16. doi:10.3390/biom8010011
 261. Huang C, Xiang Y, Wang Y, Li X, Xu L, Zhu Z, et al. Dual-specificity histone demethylase KIAA1718 (KDM7A) regulates neural differentiation through FGF4. *Cell Res*. 2010;20: 154–165. doi:10.1038/cr.2010.5
 262. Baba A, Ohtake F, Okuno Y, Yokota K, Okada M, Imai Y, et al. PKA-dependent regulation of the histone lysine demethylase complex PHF2-ARID5B. *Nat Cell Biol*. 2011;13: 668–675. doi:10.1038/ncb2228
 263. Liu H, Wang C, Lee S, Deng Y, Wither M, Oh S, et al. Clipping of arginine-methylated histone tails by JMJD5 and JMJD7. *Proc Natl Acad Sci*. 2017;114: 7717–7726. doi:10.1073/pnas.1706831114
 264. Shen J, Xiang X, Chen L, Wang H, Wu L, Sun Y, et al. JMJD5 cleaves monomethylated histone H3 N-tail under DNA damaging stress. *EMBO Rep*. 2017;18: 2131–2143. doi:10.15252/embr.201743892
 265. Liu H, Wang C, Lee S, Ning F, Wang Y, Zhang Q, et al. Specific Recognition of Arginine Methylated Histone Tails by JMJD5 and JMJD7. *Sci Rep*. 2018;8: 1–11. doi:10.1038/s41598-018-21432-8
 266. Ge W, Wolf A, Feng T, Ho CH, Sekirnik R, Zayer A, et al. Oxygenase-catalyzed ribosome hydroxylation occurs in prokaryotes and humans. *Nat Chem Biol*. 2012;8: 960–962. doi:10.1038/nchembio.1093
 267. Shechner DM, Hacısüleyman E, Younger ST, Rinn JL. CRISPR Display: A modular method for locus-specific targeting of long noncoding RNAs and synthetic RNA devices in vivo. *Nat Methods*. 2015;12: 664–670. doi:10.1038/nmeth.3433.CRISPR
 268. Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal Cas9 Proteins for RNA-Guided Gene Regulation and Editing. *Nat Methods*. 2013;10: 1116–1121. doi:10.1038/nmeth.2681
 269. Briner AE, Donohoue PD, Goma AA, Selle K, Slorach EM, Nye CH, et al. Guide RNA functional modules direct Cas9 activity and orthogonality. *Mol Cell*. 2014;56: 333–339. doi:10.1016/j.molcel.2014.09.019
 270. Braff JL, Yaung SJ, Esvelt KM, Church GM. Characterization of Cas9–guide RNA orthologs. *Cold Spring Harb Protoc*. 2016; 422–425.

doi:10.1101/pdb.top086793

271. Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci*. 2013;110: 15644–15649. doi:10.1073/pnas.1313587110
272. Fu Y, Rocha PP, Luo VM, Raviram R, Deng Y, Mazzone EO, et al. CRISPR-dCas9 and sgRNA scaffolds enable dual-colour live imaging of satellite sequences and repeat-enriched individual loci. *Nat Commun*. 2016;7: 1–8. doi:10.1038/ncomms11707
273. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*. 2013;155: 1479–1491. doi:10.1016/j.cell.2013.12.001
274. Zhang S, Feng S, Jiang W, Huang X, Chen J. Construction and optimization of a base editor based on the MS2 system. *Anim Model Exp Med*. 2019;2: 185–190. doi:10.1002/ame2.12080
275. Rivera VM, Berk L, Clackson T. Dimerizer-Mediated Regulation of Gene Expression. *Cold Spring Harb Protoc*. 2012. doi:10.1101/pdb.top070128
276. Godley BF, Shamsi FA, Liang FQ, Jarrett SG, Davies S, Boulton M. Blue light induces mitochondrial DNA damage and free radical production in epithelial cells. *J Biol Chem*. 2005;280: 21061–21066. doi:10.1074/jbc.M502194200
277. Hörner M, Chatelle C, Mühlhäuser WWD, Stocker DR, Coats M, Weber W, et al. Optogenetic control of focal adhesion kinase signaling. *Cell Signal*. 2018;42: 176–183. doi:10.1016/j.cellsig.2017.10.012
278. Marín-Prida J, Pavón-Fuentes N, Llópiz-Arzuaga A, Fernández-Massó JR, Delgado-Roche L, Mendoza-Marí Y, et al. Phycocyanobilin promotes PC12 cell survival and modulates immune and inflammatory genes and oxidative stress markers in acute cerebral hypoperfusion in rats. *Toxicol Appl Pharmacol*. 2013;272: 49–60. doi:10.1016/j.taap.2013.05.021
279. Ferry QRV, Lyutova R, Fulga TA. Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs. *Nat Commun*. 2017;8: 1–10. doi:10.1038/ncomms14633
280. Chung HK, Jacobs CL, Huo Y, Yang J, Krumm S a, Plemper RK, et al. Tunable and reversible drug control of protein production via a self-excising degron. *Nat Chem Biol*. 2015;11: 1–10. doi:10.1038/nchembio.1869
281. Oberacker P. Development of recombinant influenza A reporter viruses for drug discovery. *Univ Stuttgart*. 2016;MSc-thesis: 21–24.
282. Farcas AM, Blackledge NP, Sudbery I, Long HK, McGouran JF, Rose NR, et al. KDM2B links the polycomb repressive complex 1 (PRC1) to recognition of CpG islands. *Elife*. 2012;1: 1–26. doi:10.7554/eLife.00205
283. Tzatsos A, Paskaleva P, Ferrari F, Deshpande V, Stoykova S, Contino G, et al. KDM2B promotes pancreatic cancer via Polycomb-dependent and -independent transcriptional programs. *J Clin Invest*. 2013;123: 727–739. doi:10.1172/JCI64535
284. Kang JY, Kim JY, Kim KB, Park JW, Cho H, Hahm JY, et al. KDM2B is a histone H3K79 demethylase and induces transcriptional repression via sirtuin-1–

- mediated chromatin silencing. *FASEB J.* 2018;32: 5737–5750. doi:10.1096/fj.201800242R
285. Kipreos ET, Pagano M. The F-box protein family. *Genome Biol.* 2000;1: 1–7. doi:10.1186/gb-2000-1-5-reviews3002
286. Kim KC, Geng L, Huang S. Inactivation of a Histone Methyltransferase by Mutations in Human Cancers. *Cancer Res.* 2003;63: 7619–7623.
287. Tu S, Teng YC, Yuan C, Wu YT, Chan MY, Cheng AN, et al. The ARID domain of the H3K4 demethylase RBP2 binds to a DNA CCGCCC motif. *Nat Struct Mol Biol.* 2008;15: 419–421. doi:10.1038/nsmb.1400
288. Kao HY, Downes M, Ordentlich P, Evans RM. Isolation of a novel histone deacetylase reveals that class I and class II deacetylases promote SMRT-mediated repression. *Genes Dev.* 2000;14: 55–66. doi:10.1101/gad.14.1.55
289. Dequiedt F, Kasler H, Fischle W, Kiermer V, Weinstein M, Herndier BG, et al. HDAC7, a thymus-specific class II histone deacetylase, regulates Nur77 transcription and TCR-mediated apoptosis. *Immunity.* 2003;18: 687–698. doi:10.1016/S1074-7613(03)00109-2
290. Metzger E, Wissmann M, Yin N, Müller JM, Schneider R, Peters AHFM, et al. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature.* 2005;437: 436–439. doi:10.1038/nature04020
291. Forneris F, Binda C, Vanoni MA, Battaglioli E, Mattevi A. Human histone demethylase LSD1 reads the histone code. *J Biol Chem.* 2005;280: 41360–41365. doi:10.1074/jbc.M509549200
292. Steger DJ, Lefterova MI, Ying L, Stonestrom AJ, Schupp M, Zhuo D, et al. DOT1L/KMT4 Recruitment and H3K79 Methylation Are Ubiquitously Coupled with Gene Transcription in Mammalian Cells. *Mol Cell Biol.* 2008;28: 2825–2839. doi:10.1128/mcb.02076-07
293. Jones B, Su H, Bhat A, Lei H, Bajko J, Hevi S, et al. The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet.* 2008;4: e1000190. doi:10.1371/journal.pgen.1000190
294. Bernt KM, Zhu N, Sinha AU, Vempati S, Faber J, Krivtsov A V., et al. MLL-Rearranged Leukemia Is Dependent on Aberrant H3K79 Methylation by DOT1L. *Cancer Cell.* 2011;20: 66–78. doi:10.1016/j.ccr.2011.06.010
295. Chen CW, Koche RP, Sinha AU, Deshpande AJ, Zhu N, Eng R, et al. DOT1L inhibits SIRT1-mediated epigenetic silencing to maintain leukemic gene expression in MLL-rearranged leukemia. *Nat Med.* 2015;21: 335–343. doi:10.1038/nm.3832
296. O’Geen H, Ren C, Nicolet CM, Perez AA, Halmaj J, Le VM, et al. dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Res.* 2017;45: 9901–9916. doi:10.1093/nar/gkx578
297. Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J, et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods.* 2017;14: 297–301. doi:10.1038/nmeth.4177
298. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and

- prospects for epigenetic therapy. *Nature*. 2004;429: 457–463. doi:10.4137/GEG.S12270
299. Heerboth S, Lapinska K, Snyder N, Leary M, Rollinson S, Sarkar S. Use of epigenetic drugs in disease: An overview. *Genet Epigenetics*. 2014;1: 9–19. doi:10.4137/GeG.s12270
300. Liang Y, Richardson S, Yan J, Benites VT, Cheng-Yue C, Tran T, et al. Endoribonuclease-Based Two-Component Repressor Systems for Tight Gene Expression Control in Plants. *ACS Synth Biol*. 2017;6: 806–816. doi:10.1021/acssynbio.6b00295
301. Gossen M, Bujard H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci*. 1992;89: 5547–5551. doi:10.1073/pnas.89.12.5547
302. Matsuda T, Cepko CL. Controlled expression of transgenes introduced by in vivo electroporation. *Proc Natl Acad Sci*. 2007;104: 1027–1032. doi:10.1073/pnas.0610155104
303. Schmidl SR, Sheth RU, Wu A, Tabor JJ. Refactoring and optimization of light-switchable *Escherichia coli* two-component systems. *ACS Synth Biol*. 2014;3: 820–831. doi:10.1021/sb500273n
304. Schleif R. Regulation of the L -arabinose operon of *Escherichia coli* Over forty years of research on the L -arabinose operon of *Escherichia coli* have provided insights into the. *Regulation*. 2000;16: 559–65. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0168952500021533>
305. Blumenthal T. Operons in eukaryotes. *Brief Funct Genomic Proteomic*. 2004;3: 199–211. doi:10.1093/bfgp/3.3.199
306. Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*. 2004;428: 281–286. doi:10.1038/nature02362
307. Wilson RC, Doudna JA. Molecular Mechanisms of RNA Interference. *Annu Rev Biophys*. 2013;42: 217–239. doi:10.1146/annurev-biophys-083012-130404
308. Beisel CL, Bayer TS, Hoff KG, Smolke CD. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Mol Syst Biol*. 2008;4. doi:10.1038/msb.2008.62
309. Tian S, Huang Q, Fang Y, Wu J. FurinDB: A Database of 20-residue furin cleavage site motifs, substrates and their associated drugs. *Int J Mol Sci*. 2011;12: 1060–1065. doi:10.3390/ijms12021060
310. Drake K, Hore TA. SARS-CoV-2 RNA purification from nasal/throat swabs collected in Viral Transfer Media. *BOMB.bio*. 2020; 1–4. Available: https://bomb.bio/wp-content/uploads/2020/04/SARS-CoV-2-RNA-purification-from-nasal-swabs_BOMBv2.pdf
311. Aitken J, Ambrose K, Barrell S, Beale R, Biswas D, Byrne R, et al. Scalable and Resilient SARS-CoV2 testing in an Academic Centre. *Nat Biotechnol*. 2020. doi:10.1101/2020.04.19.20071373
312. Hebert PDN, Prosser SWJ, Ivanova N V, Zakharov E V, Ratnasingham S. Massive Multiplexing Can Deliver a \$ 1 Test for COVID-19. *bioRxiv*. 2020; 1–29. doi:10.1101/2020.05.05.079400

313. Caro-Quintero A, Yockteng R, Navarrete A, Bernal J, Daza G, Ruiz-Avila C. Protocolo de extracción de ARN viral optimizado para SARS-Cov2 en AGROSAVIA. Researchgate. 2020. doi:10.13140/RG.2.2.34818.50881
314. Ianevski A, Yao R, Fenstad MH, Biza S, Zusinaite E, Lysvand H, et al. Antiviral options against SARS-CoV-2 infection. *Viruses*. 2020;12: 1–19. doi:10.3390/v12060642
315. Walker KT, Donora M, Thomas A, Phillips AJ, Ramgoolam K, Pilch KS, et al. CONTAIN: An open-source shipping container laboratory optimised for automated COVID-19 diagnostics. *bioRxiv*. 2020; 1–29. doi:10.1101/2020.05.20.106625
316. Aygar G, Kaya M, Özkan N, Kocabiyik S, Volkan M. Preparation of silica coated cobalt ferrite magnetic nanoparticles for the purification of histidine-tagged proteins. *J Phys Chem Solids*. 2015;87: 64–71. doi:10.1016/j.jpcs.2015.08.005
317. Furlan PY, Furlan AY, Kisslinger K, Melcer ME, Shinn DW, Warren JB. Water as the Solvent in the Stober Process for Forming Ultrafine Silica Shells on Magnetite Nanoparticles. *ACS Sustain Chem Eng*. 2019;7: 15578–15584. doi:10.1021/acssuschemeng.9b03554
318. Cano-Serrano E, Campos-Martin JM, Fierro JLG. Sulfonic acid-functionalized silica through quantitative oxidation of thiol groups. *Chem Commun*. 2003; 246–247. doi:10.1039/b210766j
319. Fang W, Chen X, Zheng N. Superparamagnetic core-shell polymer particles for efficient purification of his-tagged proteins. *J Mater Chem*. 2010;20: 8624–8630. doi:10.1039/c0jm02081h
320. Buten C, Lamping S, Körsgen M, Arlinghaus HF, Jamieson C, Ravoo BJ. Surface Functionalization with Carboxylic Acids by Photochemical Microcontact Printing and Tetrazole Chemistry. *Langmuir*. 2018. doi:10.1021/acs.langmuir.7b03678
321. Ouadahi K, Allard E, Oberleitner B, Larpent C. Synthesis of azide-functionalized nanoparticles by microemulsion polymerization and surface modification by click chemistry in aqueous medium. *J Polym Sci Part A Polym Chem*. 2012;50: 314–328. doi:10.1002/pola.25035
322. Gao J, Chen J, Li X, Wang M, Zhang X, Tan F, et al. Azide-functionalized hollow silica nanospheres for removal of antibiotics. *J Colloid Interface Sci*. 2015. doi:10.1016/j.jcis.2014.12.054
323. Presolski SI, Hong VP, Finn MG. Copper-Catalyzed Azide-Alkyne Click Chemistry for Bioconjugation. *Curr Protoc Chem Biol*. 2012;3: 153–162. doi:10.1002/9780470559277.ch110148.Copper-Catalyzed
324. Ji T, Lirtsman VG, Avny Y, Davidov D. Preparation, characterization, and application of Au-shell/polystyrene beads and Au-shell/magnetic beads. *Adv Mater*. 2001;13: 1253–1256. doi:10.1002/1521-4095(200108)13:16<1253::AID-ADMA1253>3.0.CO;2-T
325. Zheng B, Zhang M, Xiao D, Jin Y, Choi MMF. Fast microwave synthesis of Fe₃O₄ and Fe₃O₄/Ag magnetic nanoparticles using Fe²⁺ as precursor. *Inorg Mater*. 2010;46: 1106–1111. doi:10.1134/S0020168510100146
326. Wang M, Yan Y, Wang R, Wang L, Zhou H, Li Y, et al. Simultaneous Detection of Bovine Rotavirus, Bovine Parvovirus, and Bovine Viral Diarrhea Virus Using a Gold Nanoparticle-Assisted PCR Assay With a Dual-Priming Oligonucleotide

- System. *Front Microbiol.* 2019;10: 1–9. doi:10.3389/fmicb.2019.02884
327. Biswas S, Khimulya G, Alley EC, Esvelt KM, Church GM. Low-N protein engineering with data-efficient deep learning. *bioRxiv.* 2020. doi:10.1101/2020.01.23.917682
328. Booth DS, King N. Genome editing enables reverse genetics of multicellular development in the choanoflagellate *Salpingoeca rosetta*. *Elife.* 2020;9: e56193. doi:10.7554/eLife.56193
329. Niehaus M, Straube H, Künzler P, Rugen N, Hegermann J, Giavalisco P, et al. Rapid Affinity Purification of Tagged Plant Mitochondria (Mito-AP) for Metabolome and Proteome Analyses. *Plant Physiol.* 2020;182: 1194–1210. doi:10.1104/pp.19.00736
330. Petrov DG, Makarova ED, Germash NN, Antifeev IE. Methods for Isolation and Purification of DNA from Cell Lysates (Review). *NAUCHNOE Priborostr.* 2019;29: 28–50. doi:10.18358/np-29-4-i2850
331. Raymond CK, Raymond FC, Hill K. UltraPrep is a scalable, cost-effective, bead-based method for purifying cell-free DNA. *PLoS One.* 2020;15: 1–11. doi:10.1371/journal.pone.0231854
332. Schubert MG, Goodman DB, Wannier TM, Kaur D, Farzadfard F, Lu TK, et al. High throughput functional variant screens via in-vivo production of single-stranded DNA. *bioRxiv.* 2020. doi:10.1101/2020.03.05.975441
333. Sosa-Acosta JR, Iriarte-Mesa C, Ortega GA, Díaz-García AM. DNA–Iron Oxide Nanoparticles Conjugates: Functional Magnetic Nanoplatfoms in Biomedical Applications. *Top Curr Chem.* 2020;378: 1–29. doi:10.1007/s41061-019-0277-9
334. Tang C, He Z, Liu H, Xu Y, Huang H, Yang G, et al. Application of magnetic nanoparticles in nucleic acid detection. *J Nanobiotechnology.* 2020;18: 1–19. doi:10.1186/s12951-020-00613-6
335. Knot IE, Zouganelis GD, Weedall GD, Wich SA, Rae R. DNA Barcoding of Nematodes Using the MinION. *Front Ecol Evol.* 2020;8: 1–11. doi:10.3389/fevo.2020.00100
336. Bollmann-giolai A, Giolai M, Heavens D, Macaulay I, Malone J, Clark MD. A low-cost pipeline for soil microbiome profiling. *bioRxiv.* 2020; 1–22. doi:10.1101/2020.05.07.082305
337. Knapp BD, Zhu L, Huang KC. SiCTeC: an inexpensive, easily assembled Peltier device for rapid temperature shifting during single-cell imaging. *bioRxiv.* 2020; 1–50. doi:10.1101/2020.05.29.123166
338. Ramos-Mandujano G, Salunke R, Mfarrej S, Rachmadi A, Hala S, Xu J, et al. A Robust, Safe and Scalable Magnetic Nanoparticle Workflow for RNA Extraction of Pathogens from Clinical and Environmental Samples. *medRxiv.* 2020; 2020.06.28.20141945. doi:10.1101/2020.06.28.20141945
339. Klein S, Mueller TG, Khalid D, Sonntag-Buck V, Heuser A-M, Glass B, et al. SARS-CoV-2 RNA extraction using magnetic beads for rapid large-scale testing by RT-qPCR and RT-LAMP. *Viruses.* 2020;12: 1–14. doi:10.3390/v12080863
340. Ware SA, Desai N, Lopez M, Leach D, Zhang Y, Giordano L, et al. An automated, high throughput methodology optimized for quantitative cell-free mitochondrial and nuclear DNA isolation from plasma. *bioRxiv.* 2020; 1–29.

doi:10.1101/2020.07.16.206987

341. Dukatz M, Holzer K, Choudalakis M, Emperle M, Lungu C, Bashtrykov P, et al. H3K36me_{2/3} Binding and DNA Binding of the DNA Methyltransferase DNMT3A PWWP Domain Both Contribute to its Chromatin Interaction. *J Mol Biol.* 2019;431: 5063–5074. doi:<https://doi.org/10.1016/j.jmb.2019.09.006>
342. Jeunen GJ, Knapp M, Spencer HG, Taylor HR, Lamare MD, Stat M, et al. Species-level biodiversity assessment using marine environmental DNA metabarcoding requires protocol optimization and standardization. *Ecol Evol.* 2019;9: 1323–1335. doi:10.1002/ece3.4843
343. Ortega-Recalde O, Day RC, Gemmell NJ, Hore TA. Zebrafish preserve global germline DNA methylation while sex-linked rDNA is amplified and demethylated during feminisation. *Nat Commun.* 2019;10. doi:10.1038/s41467-019-10894-7
344. Todd E V., Ortega-Recalde O, Liu H, Lamm MS, Rutherford KM, Cross H, et al. Stress, novel sex genes, and epigenetic reprogramming orchestrate socially controlled sex change. *Sci Adv.* 2019;5: 1–15. doi:10.1126/sciadv.aaw7006
345. Vanyorek L, Ilosvai ÁM, Szőri-Dorogházi E, Váradi C, Kristály F, Prekob Á, et al. Synthesis of iron oxide nanoparticles for DNA purification. *J Dispers Sci Technol.* 2019; 1–8. doi:10.1080/01932691.2019.1708380
346. Abbate J, Lacayo JC, Prichard M, Pari G, McVoy MA. Bifunctional protein conferring enhanced green fluorescence and puromycin resistance. *Biotechniques.* 2001.
347. Lam AJ, St-Pierre F, Gong Y, Marshall JD, Cranfill PJ, Baird MA, et al. Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat Methods.* 2012. doi:10.1038/nmeth.2171
348. Norrander J, Kempe T, Messing J. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene.* 1983. doi:10.1016/0378-1119(83)90040-9
349. Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet.* 2013;29: 569–574. doi:10.1016/j.tig.2013.05.010
350. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001. doi:10.1093/nar/29.9.e45
351. Schork SM, Bee G, Thumm M, Wolf DH. Catabolite inactivation of fructose-1,6-bisphosphatase in yeast is mediated by the proteasome. *FEBS Lett.* 1994. doi:10.1016/0014-5793(94)00668-7

Supplementary Data

Table 0-1: gRNA spacer sequences.

Primer	Sequence [5'...3']
TFRC sg1	ctttaatcctcttatcaacgggg
TFRC sg2	gaacacctgcggttcagtcgg
TFRC sg3	tcaggtgtaccagagccgaagg
TFRC sg4	cccgagcccggggaaccgaagg
TFRC sg5	cagcccggagccggcgactagg
TFRC sg6	gggaagggacgagaggcgaaggg
TFRC sg7	cccacaggctcgcgccccgctc
TFRC sg9	tgctctgacagatgccccgggg
CXCR4 sg1	tgttatcactggcgcgttgg
CXCR4 sg2	ccaccgacggttaaacatcaca
CXCR4 sg3	ccgtcgccgcacagagttaac
EPCAM sg1	gttcccgaccctccccccagg
EPCAM sg2	gcagcgtcctccggttaaagg
EPCAM sg3	ggagacgaagcacctggggcggg
EPCAM sg4	gcgccccaaactgcagcgccggg
EPCAM sg5	ggccagaggtgagcagtcgggg
EPCAM sg6	gcgagcgagcacctcgacgcgg
EPCAM sg7	gccccaggcctcgcgctccgg
LIN28A sg1	cccgcgggggttggtcattgt
LIN28A sg2	ttcgagctcgggacttagcggg
LIN28A sg3	agtctctggaaacggtgtgg
LIN28A sg4	cccgagacggcctccgattccg
LIN28A sg5	ctgtccatgaccgccgcgg
LIN28A sg6	cctccgcttaccacgtgtctat
ASCL1 sg1	ctgtccaagcggtcgagcgacgg
ASCL1 sg2	cgggagaaaggaacggaggggg
ASCL1 sg3	gcagccgctcgtgcagcagcgg
ASCL1 sg4	cgaagccaaccgcgaaggagg
ASCL1 sg5	cctgttctttgccagcggcga
TELS	aaactaacctaaccctaacc
Chr9	aaactcattccattccattccac

Table 0-2: Cloning primers 1 – gRNAs plasmids

Primer	Sequence [5'...3']
ccdB_out_f	gcgattaagttggtaacg
ccdB_out_r	ctggggaatataaatgtcagg
U6_crRNA_f	GACCtgatcggattGGTCTCAactctttttctagaccagctttc
U6_crRNA_r	TCTgtcaCAAAACAGCATAGCTCTCAAACAGGTCTTCTCGAAGAC
U6_gRNA_in_f	tatcaactgaaaaagttggcaccgagtcggtgctgacAGAGACctgatcggattGGTC
U6_gRNA_in_r	acggactagccttattgaactgctatttctagctctcaaacAGGTCTTCTCGAAGAC

Table 0-3: U6_gRNA constructs (examples). Promoter sequences are depicted in grey, the stuffer region for the insertion of the spacer sequence in yellow, the stuffer for aptamer cloning in cyan and the gRNA scaffold is coloured in green. After cloning of MS2 and/or CsyTag, both aptamers are flanked by spacer sequences. The Csy4 restriction site is marked with an *.

Construct	Sequence [5'...3']
gRNA	gagggcctattcccatgattcctcatatttgcataacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattag tacaaaatacgtgacgtagaagaataaattctggtagttgcagtttaaaattatgtttaaaatggactatcatatgcttacgtaactgaaagt atctcgattctggctttatatactgtgaaaggacgaaacaccgggtcttcgagaagacctggttagagctagaaatagcaagttcaataag gctagtccgttatcaactgaaaaagttggcaccgagtcggtgctgacagagacctgatcggattggtctcaactctttttt
gRNA-MS2	gagggcctattcccatgattcctcatatttgcataacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattag tacaaaatacgtgacgtagaagaataaattctggtagttgcagtttaaaattatgtttaaaatggactatcatatgcttacgtaactgaaagt atctcgattctggctttatatactgtgaaaggacgaaacaccgggtcttcgagaagacctggttagagctagaaatagcaagttcaataag gctagtccgttatcaactgaaaaagttggcaccgagtcggtgctgaccacagagcaggctcgtgacatgaggatcaccatgtgtccaagcgca cggactctttttt
gRNA-CsyTag-MS2	gagggcctattcccatgattcctcatatttgcataacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattag tacaaaatacgtgacgtagaagaataaattctggtagttgcagtttaaaattatgtttaaaatggactatcatatgcttacgtaactgaaagt atctcgattctggctttatatactgtgaaaggacgaaacaccgggtcttcgagaagacctggttagagctagaaatagcaagttcaataag gctagtccgttatcaactgaaaaagttggcaccgagtcggtgctgaccacagagcaggctcgtggttactgcccgtataggcag*ctaagaagtg ccaagcgcacggacttacgacagagcaggctcgtgacatgaggatcaccatgtgtccaagcgacggactctttttt

Table 0-5: template sequences

Construct	Sequence [5'...3']
DmrA_FLAG_CDS	atgtctagaggagtgacaggtggaaacatctccccaggagacggggcgacctccccaagcgccagacctgctggtgactacaccggg atgctgaagatggaagaaattgattcctcccgggacagaaacaagcccttaagttatgctaggcaagcaggaggatccgaggctgggaa gaaggggtgcccagatgagtggtgacagagccaaactgactatctccagattatgcctatggtgccactgggcacccaggcatcatcca ccacatgccactctcgtctcgatggtgagctctaaaactggaactagttatgattacaaggatgacgacgataag
DmrC_HA_CDS	atggctctagaatcctctggcatgagatggtgatgaagcctggaagggcatctcgtttgtactttgggaaaggaacgtgaaggcatggtga ggtgctggagcccttgcacgctatgatggaacggggccccagactctgaaggaacatccttaacaggcctatggtcagatggtgagggcc caagagtggtgcaggaagtacatgaaatcaggaatgtcaaggacctcctcaagcctgggacctctattatcatgtgtccgacgaatctcaaag actagttatccgtacgacgtaccagactacgca
gBLOCK-rain-vTES	tgacccgaggatccgggggactggtccggtacaggcgaagcgggaacaggaggatcagccagcgggtccggacgggctccccaggaattc ccgggtcactctgcaggattacaaggacgacgatgacaagctcaggaatgacaatgaccgctgacgtggaagaaagtcctggtggcgg ctacggggttttgcacgaaaatctttgagaaggagaattggtcgaagagtcctctgtatagtcgccacaacgacgactggggcagacactc gaagattaccttttccgaaagaatgctccgcatggcgtgggatttggcaatttcaaccacagtaagacccaatgctcgacacgagtt gacagcaggtctaaaagaatgcggatcttacaatcaaaccatcgcgataggcaggaatcaccatttcacggagacgattactggcttag ccgacccccggctgacacagaat

Table 0-6: Cloning primers 3 – EPIC'RISPR vector backbone

Primer	Sequence [5'...3']
gBLOCK-rain-f	tgacccgaggatccgg
gBLOCK-rain-r	attctgtgcagccggg
open_gBLOCK_f	gattacaaggacgacgatgacaagctcg
open_gBLOCK_r	ctgcagagtcgacccggaattc

Table 0-7: pHAGE_EFS plasmid vector backbone for EPIC'RISPR ED cloning. Promoter sequences are depicted in grey, the GGS linker in pink with the PstI restriction site as a *. The FLAG tag is coloured in green and the NLS in red. The MCP aptamer binder can be arbitrarily exchanged for any other protein.

Construct	Sequence [5'...3']
MCP-gBLOCK	<pre> gggcagagcgcacatcgccacagtccccgagaagtggggggaggggtcggcaaltgaaccggtcctagagaagtgccgcggggtaaa ctgggaaagtgatgtcgtgactggctccgcttttccgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgtcttttcgcaac gggttgcgccagaacacaggtgctgtgacgcggtccgggtaaggatgaacgcgtgcaggctggccaccatggcttctaacttactcagttcgtt ctcgtcgacaatggcggaactggcgactgctgcgccaagcaactcgtctaacgggatcgtgaatggatcagctctaactcgcgttcacag gcttacaaglaacctgtagcgtcagagctcgcgagaatcgcaaatcacccatcaaaagtcgagggtcctaaaggcgcctggcgttcgtact aaatatggaactaaccattccaatttcgccacgaattccgactgcgagctattgtaaggcaatgaaggctcctaaaagatggaaccggattc cctcagcaatcgcagaaactccggcatctacggggatctggtagcggcgctgatccgggggcaactggttccggtacagcgggaagcggaa caggaggatcagccagcgggtccggacgggctcccaggaattccgggtcgactctca*ggattacaaggacgacgatgacaagctcagac caaagaaaaagcggaaagttaa </pre>

Table 0-8: EPIC'RISPR ED amplification primers. Each primer consists of a universal overhang which matches the linearized vector backbone and a sequence matching the respective ED.

ED	Pf	Pr
	3' overhang [5'...3']	5' overhang [5'...3']
all	ggaattcccgggtcgact	gatccagcggctgtaggggc
	3' match [5'...3']	5' match [5'...3']
KRAB	atggacgcgaaatcacttacg	taccagccaaggtcttcccc
mTET3del1	atcgcccttagagttccctac	gatccagcggctgtaggggc
VPR	agggacgggctgacgattggacattttg	aaacagagatgtgcaagatggacag
VP64	agggacgggctgacgattggacattttg	tctagagtaatcagcatgccaggctg
SUV39H1	tgtgtcgtatcctcaagcagttc	gaagaggtattgcccaggactc
SUV39H2	tgcccgttactgcttcagc	gttgaggtaacctctgcaagtcac
EHMT2	aatcgggccatccgcacagag	cgctggcccgcctggac
EHMT1	tccaagggctctgacatcaacatccg	gacggctgcccgaacacc
SETDB1	ttttactatatttgacatcactatgggaag	cattgaatgagaggacgtcttctt
SETDB2	cagttgctcggaattacccaaag	ctgccaatgtggggttaataaatgtagaaaaaaaaatatta
KMT2D	tcaaccagacgtgccaccag	cgctgccgtcggttccttaac
KMT2C	agcacagtcactggagaactgaacg	gtgaactgccggaagtggatgaac
KMT2B	ctgaacagcaccagcatgtctaagg	cctggaattgtcgaaatggatgaac
SETD1A	cgcgcttacgagccac	ctgccgggctccctaaac
SETD1B	cgctcggagttgaggagatg	gagaactgccgggggacctcaac
ASH1L	ccagatgtcccactatataagaaaattcg	gtcaggaatcatcggaggcaagagt
NSD1	aaagagctaagacagctgcaggaag	gaactcagtggtctctgggt

SMYD2	atgagggccgagggc	gatcaaacaggaaattgaaagccactga
SMYD3	atggagccgctgaaggtgg	gacgccaacatcagagcatcc
NSD3	gcacaaagagaagaaagaagccc	gcagataactgcagtggtttctagga
NSD2	ggggaccggggcagc	gaagggagaggcagtcagaggac
DOT1L	atgggggagaagctggagctg	caagcgggggcgccccagaag
KMT5A	atggctagaggcaggaagatgtcc	cccaccctggctgaagcattaa
KMT5B	aagtgttgggagaaatccaagaacatg	caactgatgcagataccactcaggaa
KMT5C	gggcccgcagagtgacag	cagccgcagggcctg
EZH2	aaacaccggttgggctg	ggcatcgaagagaaatgaaatccct
EZH1	atggaaataccaaatccccctacctc	cgagagggagaccgactctt
SETD7	atggatagcgacgacgagatgg	caggccaccagcaaaagtga
PRDM2	atgaatcagaacactactgagcctg	cggagcaagcggagctcc
vSET	atgttcaatgaccgctgacgctg	gaccccgctgcacagaaatgaatt
EED	atgtccgagagggaagtgtcgac	ggcgctgggctgacttcca
KDM1A	gccccactgaggaagaaaatg	cagcagtcccaagcatgtga
KDM2A	ggaaaaagaacttttgacttgaagag	gtccctcagcatggattggag
KDM2B	acaaaatgctttgaaattgagtcggc	caaagaaccgggctgtgggt
KDM2B F-box	gatggggcagcccagctcatg	gaaaaactcctgcataaactgagt
KDM3A	agcattctggccttgacactcctc	gtgaatccagttttgcaaacct
KDM3B	acttctactcctgcttctgtg	gttccgctgactcaggaattc
KDM3C	ataccacattctggatctgagaagc	cattaacacaggaactgagactttg
KDM4A	atggctctgagctgaaactctg	gctgagttcttaaggagagtgaaactg
KDM4B	atggggtctgaggaccag	gaccacacggggcccacg
KDM4C	gtggccgaggtgaaaatgc	ggatatatacaccattgatcacaggaag
KDM4D	gcccagaatccaaatgtaacataatg	gtggaaactgggcaagac
KDM4E	atgaagctctgactccagtc	gttatgagctctgaaacacaggg
KDM5A	atggcggcgctggggc	gctgaagctgtaacttctgact
KDM5B	ttctgcctccaccgagtg	gcttggaggcaaatcaacaag
KDM5C	gatttctaccgcccaccggag	cgggctgagctcttgacacc
KDM6A	cagatcatacctcaatgtctgtgtcc	ctccattaccatccgcctcatct
KDM6B	gacgtcgtgcgcccag	gcccggcggggcgcg
KDM7A	tcctgatgaaaaaaggaggaactggc	cggcaaacagttaaatctcagggaaatt
KDM7B	atggcctcgggtgccg	ctgggtggaagacatctccaacagaac
KDM7C	tccaccttaagaagaagcggacc	ctccaaagcctccgaccg
KDM8	gcgaccatggttgattcca	cggtcagctctgtgtgtg
RIOX1	ggtggggagccggcctgggac	ctcactaagatgcctctagccctaaat
JMJD6	atgaaccacaagagcaagaagcg	ctcgggtgacctcaggagctc
KAT2A	cgccggcgcacatcaggtccatg	gaatccccgcaccccctacacg
KAT3B	atttcaaaccagaagaactacgacaggc	gcacacgcagagccaggac
KAT5	accggatgaagaacattgagtcattg	ctgcactcactcccaaggac
KAT6A	ccccctgatccacaagtccg	gactccagtcatagtgtccaactct
KAT6B	cggtagcctctgtgattgaattg	cgagaagctgagaaaggagct
KAT7	gaggaagcaacatgataaaacaattgc	gatcccagctgcttaaaatggaccct
KAT8	accaagggaagatgtggacaagatc	cgctgcctcaagtgggacc
KAT9	atacaggtcagatgtgagaaccagag	caaggcccgtacatggtgaagatgctgaaa
HDAC1	atggcgagagcagggc	gtcaaggaggaggtcaagtggcc
HDAC2-CD	aaaaaaaaagctgctactactacgagctg	gcgttgttgaataatgagcagctgta
HDAC3-CD	aaaaaaaaagctgctactactacgagctg	gaacctgcacctagctccag
HDAC3	aagaccgtggcctatttctacgacc	caatgacaaggaagcagatgtggagatt



HDAC4-CD	aagaccgtggcctatttctacgacc	catggaagaggagcccctg
HDAC5-CD	accaagccgaggttcacgac	gagcaggagcctgcctg
HDAC6-CD1	gcaatggaagaagacctaatcgaggga	ctgggaggttcttgatgagatcaact
HDAC6-CD2	gggctggtctatgacccaaatgatgaatc	gatactggcgagcttacgg
HDAC7	tccgggctcagcttcccag	ctggcctcctgccagactcc
HDAC8	atggaggagccggaggaac	caaagggaatctgaagcatgtggtc
SIRT1	atgattggcacagatcctgaacaattc	ctgtgaaacaggaagtaacagacatgaac
SIRT2	agcctgggcagccagaag	gccagcatagatgccagctg
SIRT6	atgtcggtaattacgcggcg	ctggagcccaaggaggaatct
SIRT7	cgggagctggccagcgcctg	gccagctgggctggagatcccc

Table 0-9: Cloning primers 4 – FKBP (DmrA), FRB (DmrC), PUF and PuroR

Primer	Sequence [5'...3']
pHAGE_DmrA_PstI_f	cgggtcgactctgcaggagtgagggtgaaaccatc
pHAGE_DmrA_PstI_r	ccttgtaatcctgcagttccagtttagaagctccacatcg
pHAGE_DmrC_PstI_f	cgggtcgactctgcagatcctctggcatgagatgtgg
pHAGE_DmrC_PstI_r	ccttgtaatcctgcagcttgagattcgtcgg
pHAGE_EFS_DmrA_NcoI_f	tggcgccaccatgggtggagtgagggtgaaaccatc
pHAGE_EFS_DmrA_SpeI_r	atccacctcactagtagctccgacctccagtttagaagctccacatcg
pHAGE_EFS_DmrC_NcoI_f	tggcgccaccatgggtatcctctggcatgagatgtgg
pHAGE_EFS_DmrC_SpeI_r	atccacctcactagtagctccgaccttgagattcgtcgg
PUF_pcDNA3.1_Gib_f	gaaggtgggcccggaggccgcagccgc
PUF_pcDNA3.1_Gib_r	gccttggatccggcaattggccctaagtcaacaccttctcatg
PuroR_GFPout_f	atgaccgagtacaagcccacggtg
PuroR_GFPout_r	ggtggcgaccgtagcgctagc

Table 0-10: Cloning primers 5 – Csy4 and CsyTag

Primer	Sequence [5'...3']
pHAGE_EFS_Csy4_NcoI_f	ggcgccaccatgggtatggaccactacctcgacattc
pHAGE_EFS_Csy4_SpeI_r	cagagccactagtacctccgccaccgaaccagggaacgaaacctcc
Csy4_H29A_HindIII_f	cttcggcaagcttgcgcaggccctgg
Csy4_H29A_HindIII_r	ccagggcctgcgcaagcttgccgaag
Csy4_H29G_HindIII_f	cttcggcaagcttggccaggccctgg
Csy4_H29G_HindIII_r	ccagggcctggccaagcttgccgaag
NLS_out_f	taatctagataatcaacctctggattac
NLS_out_r	gaaccagggaaacgaaacc
CsyTag_NotI_f	ggccgttactgccgtataggcagctaagaaa
CsyTag_NotI_r	ggccttcttagctgcctatacggcagtgaaac
CsyTag_KpnI_f	gttactgccgtataggcagctaagaaagtac
CsyTag_KpnI_r	ttcttagctgcctatacggcagtgaaactac
CsyTag_EcoRI_f	aattgttactgccgtataggcagctaagaaa
CsyTag_EcoRI_r	aattttcttagctgcctatacggcagtgaaac
multiCsyTag_I1_f	tgtctcatcatttggcaaagaattgttactgccgtataggcagctaagaaa
multiCsyTag_I1_r	tggggaagcccctccgttcttagctgcctatacggcagtgaaac
multiCsyTag_I2_f	cggacgggctcccagttcactgccgtataggcagctaagaaa
multiCsyTag_I2_r	ttccgttccgctgtttcttagctgcctatacggcagtgaaac
multiCsyTag_I3_f	acagggcgaagcgggaagttcactgccgtataggcagctaagaaa
multiCsyTag_end	gcagtgaaactaccgtcgcactgcagtttcttagctgcctatacggcagtgaaac

Table 0-11: Plasmids for expression control. For pHAGE_EFS constructs, the promoter sequences are depicted in grey, Csy4 in cyan, the GGS linker in pink with the PstI restriction site as a . The FLAG tag is coloured in green and the NLS in red. The **H at position 29** can be further mutated into A or G and the linker, FLAG tag and NLS can be removed. For pCAG-constructs, the UTRs are depicted in grey, CsyTag sites in purple with  indicating Csy4 restriction sites. EGFP is coloured in green and the degradation domain d2 in dark green.

Construct	Sequence [5'...3']
pHAGE_EFS_ Csy4 WT	Gggcagagcgcacatcgcccacagctccccgagaagttggggggaggggtcggcaattgaaccgggtgcctagagaagggtggcgcgggtaaacgggaaagtgat gtcgtgactgctcgccttttcccgaggggtgggggagaaccgtatataagtgacagtagtcgocgtgaacgctcttttcgcaacgggttgcgcgagaacacaggtg gtgacgcgggtccggttaagtgaaacgctgagcggctggccaccatgggatggaccactacctcgacattcgttgcgaccggaccggaaattccccggcgca ctcatgagcgtgctcttcggaagctc aa caggccctggggcacagggcggggacaggatcgcgctgagcttccccacctcgacgaaagccgctcccgctggg cgagcgcctgcgcaattcatcctcggcggacacctctgctcctcgcgccggccctggctggaagggttgcgggacacatctgcaattcggagaaccggcagctgt cctacccccacacgctaccgtcaggtcagctcgggttcaggcgaaaagcaatccggaacgctcggcgggcctcatgcgcggcagcatctgagtaggaggagg ctcgaaacgcattccccgatacggctcggagagccttgacctgccctctcgtacgctacgcagccagagcaccggacagcacttcgtctctcatccgccaaggcc gttgacggtagcggcagaggaaggaggattaccctgtacgggttgagcaaggagggttctgtccctggctcgggtggcgaggtactagtggagggtgatccggggc actggtccggtagcgggaagcgggaacaggaggtacgacggctccggacgggttccggacgggttcccgaggaattccgggctgactctgca g gattacaaggacgacga tgacaagctcag caaaagaaaagcggaaag ctaa
pHAGE_EFS_ Csy4 WT ΔNLS	Gggcagagcgcacatcgcccacagctccccgagaagttggggggaggggtcggcaattgaaccgggtgcctagagaagggtggcgcgggtaaacgggaaagtgat gtcgtgactgctcgccttttcccgaggggtgggggagaaccgtatataagtgacagtagtcgocgtgaacgctcttttcgcaacgggttgcgcgagaacacaggtg gtgacgcgggtccggttaagtgaaacgctgagcggctggccaccatgggatggaccactacctcgacattcgttgcgaccggaccggaaattccccggcgca ctcatgagcgtgctcttcggaagctc aa caggccctggggcacagggcggggacaggatcgcgctgagcttccccacctcgacgaaagccgctcccgctggg cgagcgcctgcgcaattcatcctcggcggacacctctgctcctcgcgccggccctggctggaagggttgcgggacacatctgcaattcggagaaccggcagctgt cctacccccacacgctaccgtcaggtcagctcgggttcaggcgaaaagcaatccggaacgctcggcgggcctcatgcgcggcagcatctgagtaggaggagg ctcgaaacgcattccccgatacggctcggagagccttgacctgccctctcgtacgctacgcagccagagcaccggacagcacttcgtctctcatccgccaaggcc gttgacggtagcggcagaggaaggaggattaccctgtacgggttgagcaaggagggttctgtccctggctc
pCAG_EGFPd2 with 5 possible CsyTag sites	ctcctggcaacgtgctggtattgtgctgtctcatcatttggcaagaattgtcactcgcgtataggcag g ctaaagaaacggacgggcttccccagttcactgcgctatagg cag g ctaaagaaacagggcggaagcggaa g ttcactcgcgtataggcag g ctaaagaaacgctcagctcagcggtagctcactcgcgtataggcag g ctaaagaa g tacc ggggcccggtaccaccggtcgcacc at ggtagcaaggcgaggagctgttaccgggggtggtcccatcctggtcagctggaaggcagctaaacggccac aagttcagcgtgctcggcagggcgaggcgatgccactacggcaagctgaccctgaagttcatctgcaaccggcaagctgcccgtgccctgcccacccctcgt accaccctgacctacggcgtgacgtgcttaccgctaccccaccacatgaagcagcagcacttctcaagtccgcatgccgaaggctacgtccaggagcagcacc atctctcaaggacgacggcaactacaagaccgcggaggtgaaagttcagggcgacaccctggtgaaccgcatcagctgaaggcatcagctcaaggaggga cggcaacatcctgggcacaagctggagtacaactacaacagccaacagctatataatgcccgaagcagaagaacggatcaagtgaaactcaagatccgc cacaacatcagggacggcagcgtgagctcggaccactaccagcagaacacccccatcggcagcggccccgtgctgctcccgaacaactactcctgagcacc cagtcgccctgagcaagacccaacgagaagcggatcacatggtcctgctggagtctgaccgccggatcactctcggcatggacgagctgacaagaa gcttagccatggctccccggaggaggaggagcaggatgatggcagctgccatgcttctgcccaggagagcgggatggaccctgacccctgagcctgctctg ctaggatcaatgtagatgctcggccggcc g ttcactcgcgtataggcag g ctaaagaaacggcagcactcctcaggtgagcggctcctcagaagggtggtgctggt ggccaatgccctggtcacaataaccactgagatcttttccctcgcgcaaaatattggggacatcatgaagcccttgagcatctgactctgctaataaggaaattat ttcattgcaatagtggtggaattttgtctctca

Table 0-12: qPCR amplification primers.

Primer	Sequence [5'...3']
Q-EPCAM-f	ctccatgtgctggtgtg
Q-EPCAM-r	tgttttagttcaatgatgatccagta
Q-TFRC-new-f	agggataccttcgtccctg
Q-TFRC-new-r	ccggatgcttcacatttgc
Q-CXCR4-f	agcatgacggacaagtacagg
Q-CXCR4-r	gatgaagtcgggaatagtcagc
Q-REEP5-f	tftggctaccagcctacatc
Q-REEP5-r	caacaggaagccacactcagc
Q-LIN28A-f	ctgtaagtggttcaacgtgcg
Q-LIN28A-r	cttcaagctcgggaaccctt
Q-ASCL1-fn	cggctctatcctactcgtcg
Q-ASCL1-rn	gatcacctgcttcaaagtc
Q_gRNA_f	gttgagagctagaatagcaagt
Q_gRNA_cut_r	gacctgctctggggtcag
Q_gRNA_full_r	tggcacacatgggtgatcc
Q_BleoR_BG_f	tgatgaacagggtcacgtcg
Q_BleoR_BG_r	caagttgaccagtccggtc