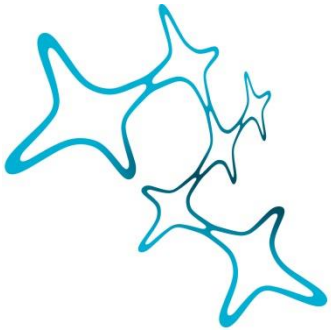
GENOME EDITING AND TRANSCRIPTIONAL ACTIVATION USING CRISPR-CAS9 TECHNOLOGY

EXPLORING AND MANIPULATING EPIGENETIC MECHANISMS IN NEURONS

Victoria Splith



Graduate School of
Systemic Neurosciences

LMU Munich

Dissertation der
Graduate School of Systemic Neurosciences der
Ludwig-Maximilians-Universität München

March, 2019

Supervisor:
PD Dr. Stylianos Michalakis
Department of Pharmacy- Center for Drug Research
Ludwig-Maximilians-Universität München

First Reviewer: PD Dr. Stylianos Michalakis
Second Reviewer: Prof. Dr. Martin Biel
External Reviewer: Dr. Volker Busskamp

Date of Submission: 28.03.2019
Date of Defense: 06.06.2019

Contents

List of Abbreviations

2OG	2-oxoglutarate
5caC	5-carboxycytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
µg	microgram
µl	microliter
µM	micromolar
aa	amino acid
ACHM	achromatopsia
ACTB	actin beta
Alas	aminolevulinic acid synthase
ANOVA	analysis of variance
Anti-Anti	Antibiotic-Antimycotic
APS	ammonium persulfate
ATP	adenosine triphosphate
BER	base excision repair
BES	2-[bis(2-hydroxyethyl)amino]ethanesulfonic acid
BHT	2,6-ditert-butyl-4-methylphenol
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BS	bisulfite
BSA	bovine serum albumin
C574	cytosine/aa574 of *Sp*Cas9
$CaCl_2$	calcium chloride
Cas	CRISPR associated protein

CB	ChemiBLOCKER™
cDNA	complementary DNA
cGMP	cyclic guanosine monophosphate
CMV	cytomegalovirus promoter
CNG	cyclic nucleotide gated
CpG	cytosine-guanine dinucleotide
CRISPR	clustered regularly interspaced short palindromic repeats
crRNA	crispr RNA
Ct	threshold cycle
Cys	cysteine
d	day
DMEM	Dulbecco´s modified eagle medium
dN	deoxynucleotides
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
dNTP	deoxynucleoside triphosphate
DOX	doxycycline
DPBS	Dulbecco´s phosphate-buffered saline
ds	double stranded
DSB	double strand break
DSBH	double stranded β-helix
DTT	dithiothreitol
E573	glutamate/aa573 of *Sp*Cas9
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EF1a	translation elongation factor 1 alpha promoter
EFNA1	ephrin A1

e.g.	exempli gratia, for example
EGTA	ethylenebis(oxyethylenenitrilo)tetraacetic acid
ESC	embryonic stem cell
et al.	et alii, and others
FACS	fluorescence activated cell sorting
FAM122C	family with sequence similarity 122C
FBS	fetal bovine serum
FL	full-length
FSC	forward scatter
fw	forward
g	gram
GABA	γ-aminobutyric acid
Gapdh	glycerinaldehyd-3-phophat-dehydrogenase
GC	guanylyl cyclase
gDNA	genomic deoxyribonucleic acid
GO	Gene Ontology
GOI	gene of interest
G_T	G protein transducin
h	hour(s)
H_2O	water
HCl	hydrochloric acid
HDR	homology directed repair
HEK293	human embryonic kidney 293
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HR	homologous recombination
hRHO	human rhodopsin
HRP	horseradish peroxidase

HSF1	human heat shock factor 1
hU6	human U6 promoter
Hz	hertz
IB	immunoblot
IDT	Integrated DNA Technologies
i.e.	id est, that is
Indel	insertion/deletion
iNGN	inducible Neurogenin iPSC
Int	intein
iPSC	induced pluripotent stem cell
ITR	inverted terminal repeats
kb	kilobase
KCl	potassium chloride
kDa	kilodalton
kHz	kilohertz
KO	knock-out
KOH	potassium hydroxide
KRAB	Krueppel associated box
l	liter
L-opsin	long wavelength-sensitive opsin
M	molar
MeCN	acetonitrile
MEF	mouse embryonic fibroblast
MeOH	methanol
mg	milligram
$MgCl_2$	magnesium chloride
min	minute(s)

ml	milliliter
mM	millimolar
M-MuLV	Moloney Murine Leukemia Virus
M-opsin	middle wavelength-sensitive opsin
mRNA	messenger RNA
ms	milliseconds
MS	mass spectrometry
mSWS	murine short wavelength (S-) opsin
mV	millivolt
N	nitrogen
Na_2HPO_4	disodium hydrogen phosphate
NaCl	sodium chloride
NaOH	sodium hydroxide
NCKX	Na^+/Ca^{2+}-K^+ exchanger
NEB	New England Biolabs
ng	nanogram
NGS	next generation sequencing
NHEJ	non-homologous end joining
nm	nanometer
nt	nucleotides
NUC	nuclease
O	oxygen
pAAV	AAV plasmid
p_{adj}	adjusted p-value
PAGE	polyacrylamide gel electrophoresis
PAM	protospacer adjacent motif
pb	PiggyBac

PBS	phosphate-buffered saline
PCDHA6	protocadherin alpha 6
PCR	polymerase chain reaction
PDE	phosphodiesterase
PFA	paraformaldehyde
pmol	picomole
Pol II	RNA polymerase II
polyA	polyadenylation signal
POU5F1	POU class 5 homeobox 1
PuroR	puromycin resistance gene
qRT-PCR	Real-Time quantitative Reverse Transcription PCR
rAAV	recombinant adeno-associated virus
RAET1L	retinoic acid early transcript 1L
REC	recognition
REST	RE1 silencing transcription factor
rev	reverse
RHO	rhodopsin
Rma	*Rhodothermus marinus*
RNA	ribonucleic acid
RNP	ribonucleoprotein
RP	retinitis pigmentosa
RRBS	reduced representation bisulfite sequencing
RT	room temperature
RT-PCR	Reverse Transcription PCR
rtTA	reverse tetracycline-controlled transactivator
s	seconds
S	sulfur

S714	serine/aa714 of *Sp*Cas9
SAM	synergistic activation mediator
SDS	sodium dodecyl sulfate
seq	sequencing
sgRNA	single guide RNA
S-opsin	short wavelength-sensitive opsin
*Sp*Cas9	*Streptococcus pyogenes* Cas9
ss	single stranded
SSC	side scatter
SV40	simian virus 40
TALEN	transcription activator-like effector nuclease
TBS	Tris-buffered saline
TDG	thymine DNA glycosylase
TEMED	tetramethylethylenediamine
TET	ten-eleven translocation
TK	thymidine kinase
Tm	melting temperature
tracrRNA	trans-activating crRNA
TRE	tetracycline response element
TSS	transcription start site
TX	Triton X-100
U	units
UHPLC	ultra-high performance liquid chromatography
V713	valine/aa713 of *Sp*Cas9
VP16	*Herpes simplex* virus protein 16
VPR	VP64-p65-Rta
v/v	volume per volume

WT	wild type
w/v	weight per volume
ZFN	zinc finger nuclease
ZNF577	zinc finger protein 577
$ZnSO_4$	zinc sulfate

List of Tables

List of Figures

Abstract

The adaption of the bacterial clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 system for genome engineering has revolutionized research in life sciences. Due to its two-component structure, the system provides a so far unreached simplicity in retargeting as well as multiplexing. In addition to genome editing applications using CRISPR-Cas9, a nuclease-deficient mutant Cas9 variant (dCas9) can be customized to guide various fused effector domains to any desired locus in the genome.

In this study, the CRISPR-Cas9 system was used in two ways. On the one hand, classical genome editing was performed to knock out catalytically active TET3 protein in the human inducible Neurogenin iPS (iNGN) cell line. On the other hand, a dCas9 fusion to transcriptional activation domains, termed dCas9-VPR, was tested for the ectopic activation of photoreceptor specific genes.

This study showed that the depletion of functional TET3 resulted in a loss of 5-hydroxymethylcytosine accumulation commonly observed with neuronal differentiation. Globally, TET3KO and wild type iNGN-derived neurons show very different DNA methylation as well as gene expression patterns. As analyzed by RNA-sequencing, genes that were upregulated in TET3KO neurons compared to WT neurons were involved in neuronal function and signaling. In contrast, downregulated genes in TET3KO neurons were significantly enriched for biological processes such as development and proliferation. However, these dramatic changes in the methylome and transcriptome caused by TET3 depletion were in principle still compatible with the differentiation of TET3KO iNGN cells into neurons.

The second part of this study applied the transcriptional activator dCas9-VPR to ectopically induce the expression of specific genes *in vitro*. In particular, cone photoreceptor-like 661W cells could be triggered to express the rod photoreceptor-specific gene *Cnga1* by guiding dCas9-VPR to its promoter region. This transactivation resulted in profound expression of functional Cnga1 protein as determined by immunostaining and electrophysiological measurement of CNG channel-mediated ion conductance. Conversely, expression of the cone-specific *Opn1mw* gene could be successfully activated in mouse embryonic fibroblasts. To enable delivery of the transactivation constructs by recombinant adeno-associated viruses (rAAV), split dCas9-VPR rAAV vectors were developed and shown to mediate a successful transactivation of both genes. These results lay the foundations for further development of the transactivation technique into a potential gene therapy.

1 Introduction

1.1 CRISPR-Cas9: A Revolution in Genome Editing

1.1.1 History of Genome Editing

Genome or gene editing refers to the targeted alteration of the nucleotide sequence in the genomic deoxyribonucleic acid (DNA) of a cell or even a whole organism. It comprises different technologies for the replacement, elimination or modification of DNA sequences. Genome editing is a powerful approach to elucidate gene function in biological processes, such as development and disease. Nowadays, it emerges into novel approaches for the correction of disease-causing mutations (Charpentier and Marraffini, 2014; Hsu et al., 2014).

All genome editing techniques aim at introducing a site-specific DNA double strand break (DSB) in the target genome. To make changes to the surrounding DNA sequence, the researcher simply relies on the endogenous DSB repair machinery of the cell. Naturally, DSBs can occur due to both exogenous and endogenous influences, such as ionizing radiation, reactive oxygen species or replication errors, and can have detrimental effects on the cell (Pannunzio et al., 2018; Sallmyr and Tomkinson, 2018). Thus, to maintain DNA integrity, the cell is equipped with several possible repair pathways that are initiated by a DNA DSB and decision for one or the other mainly depends on cell cycle phase. The two major pathways are non-homologous DNA end-joining (NHEJ) and homology-directed repair (HDR). NHEJ is the prevailing repair pathway and accounts for almost all DSB repair outside of the S and G_2 phases of the cell cycle. Due to the involvement of a variable set of multiple enzymes the outcome of repair can be quite diverse and often results in errors in the original DNA sequence (Pannunzio et al., 2018). Frequently, small insertions or deletions of nucleotides, so-called “indels”, are introduced at the break site by NHEJ (Chang et al., 2017). In a protein coding sequence, such Indels can act as frameshift mutations that cause a premature stop codon in the transcribed messenger ribonucleic acid (mRNA), thereby leading to truncated or non-functional protein. Hence, targeting of the nucleases to early coding exons is preferred for generating a knock-out (KO) of the corresponding gene. In contrast to the canonical NHEJ, which uses little to no homology for repair, HDR is highly dependent on homology to a sister chromatid or a homologous sequence elsewhere in the genome and thereby represents a high-fidelity repair pathway (Pannunzio et al., 2018; Wright et al., 2018). Homologous recombination (HR) is an essential cellular process which ensures genomic stability and survival (Wright et al., 2018). In genome editing a

synthetic partially homologous repair template can be supplied for precise manipulation of the DNA sequence near the created DSB ranging from the addition of a single point mutation to the insertion of a complete reporter gene. Moreover, there are alternative end-joining pathways, which are poorly understood and only make a minor contribution to DSB repair in healthy cells. Based on the differing amount of sequence homology used to align the DNA molecules, three distinct pathways are distinguished: single-strand annealing (Ivanov et al., 1996), microhomology-mediated end-joining (Ma et al., 2003) and end-joining (Sallmyr and Tomkinson, 2018; Wang and Xu, 2017).

Genome engineering has a long history of diverse editing tools beginning with the usage of homing endonucleases (Jacquier and Dujon, 1985; Szostak et al., 1983). Homing endonucleases are distinguished from other site-specific endonucleases like restriction enzymes by rather large recognition sites (14-40 base pairs (bp)) that often occur only once in mammalian genomes (Belfort and Bonocora, 2014). To expand the target locus possibilities in the genome, screening for naturally occurring enzymes with different target sites (Jacoby et al., 2012; Li et al., 2012; Takeuchi et al., 2011) as well as the generation of mutant homing endonucleases with altered binding and cleavage specificity (Arnould et al., 2006; Ashworth et al., 2006, 2010; Ulge et al., 2011) was explored. In the 1990s, with the engineering of zinc finger nucleases (ZFNs) the field made a big step into the direction of tailor-made sequence specificity of endonucleases (Kim et al., 1996). This was achieved by linking engineered zinc finger DNA-binding domains to cleavage domains of *Fok*I endonuclease (Figure 1). In this approach, two *Fok*I domains need to dimerize for DNA cleavage, therefore always a pair of ZFNs that bind to opposite strands at the target locus is used (Maeder et al., 2008; Porteus and Baltimore, 2003; Urnov et al., 2005). In 2010, transcription activator-like effector nucleases (TALENs) were developed (Christian et al., 2010). Deciphering of the DNA recognition mechanism of the plant pathogen *Xanthomonas´* TAL effectors gave rise to the idea of using them as novel DNA-binding domains in biotechnology (Boch et al., 2009; Moscou and Bogdanove, 2009). Fusions of custom-made TALE domains to *Fok*I cleavage domains have been shown to facilitate DNA cleavage at any desired location in the genome of various organisms and have been used for the generation of many KO animals (Christian et al., 2010; Davies et al., 2013; Hockemeyer et al., 2011; Li et al., 2011; Miller et al., 2011; Tesson et al., 2011; Zhang et al., 2013).

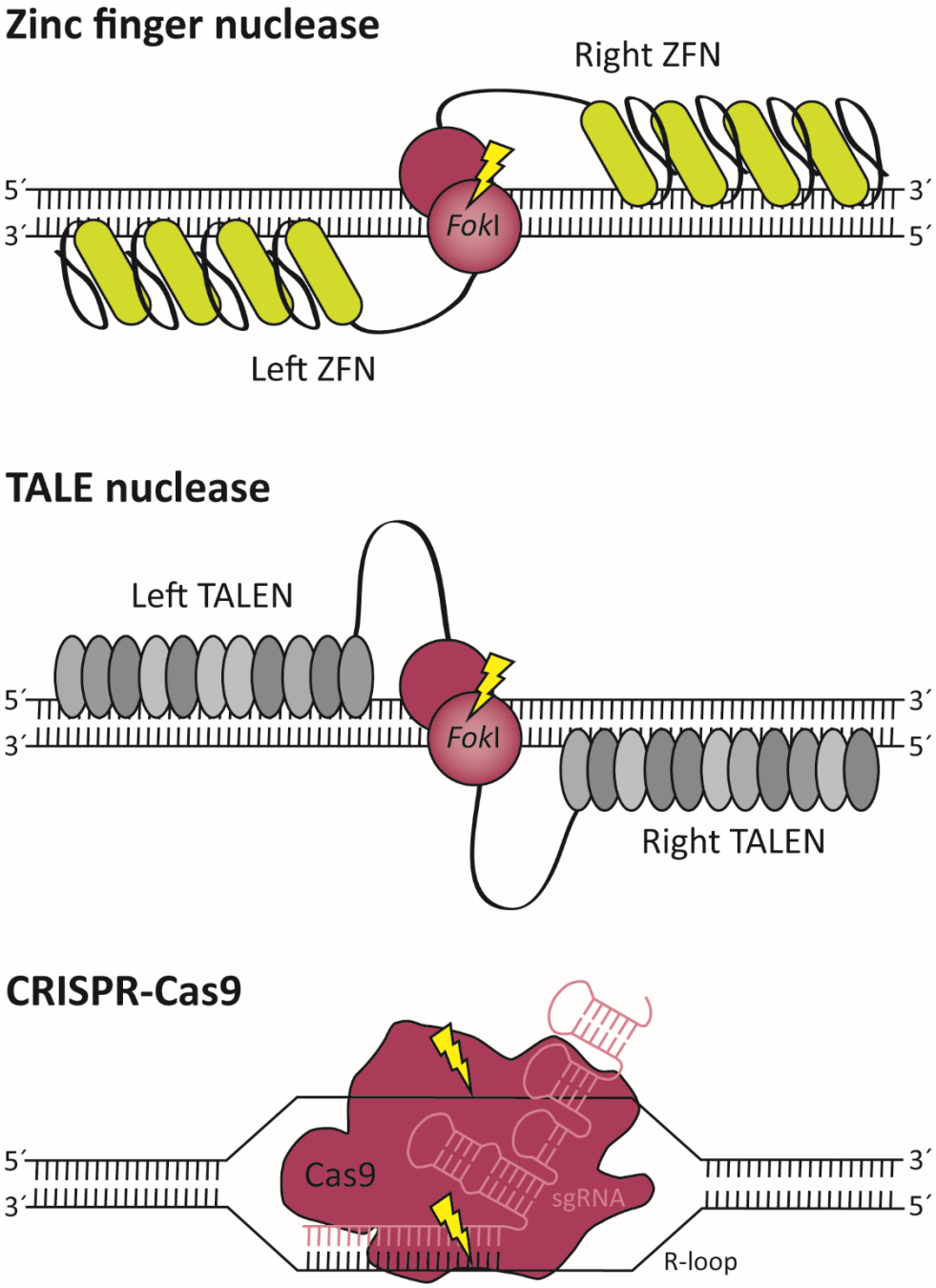


Figure 1 Genome engineering tools
Zinc finger nucleases (ZFNs) consist of a zinc finger DNA binding domain fused to *Fok*I endonuclease domain. In transcription activator-like effector nucleases (TALEN) a *Fok*I domain is fused to a TALE DNA recognition array. Both, ZFN and TALEN, act as dimers to induce a site-specific DNA double strand break (yellow flash). In the CRISPR-Cas9 system, targeting of the Cas9 endonuclease to a specific DNA locus is facilitated by a single guide RNA (sgRNA). Partial base pairing of the sgRNA and the complementary strand of genomic DNA results in R loop formation and Cas9-mediated cleavage of the complementary and non-complementary DNA strand.

One of the biggest scientific breakthroughs of the 21st century was the application of the bacterial clustered regularly interspaced short palindromic repeats (CRISPR)-

CRISPR associated protein (Cas) system as a genome editing tool (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013a). In contrast to all previously described tools, sequence specific targeting of the Cas endonuclease is mediated by an RNA molecule and does not have to be encoded in the effector itself. This two-component system defines a new class of RNA programmable genome engineering tools with highly increased flexibility and vast new possibilities (Cong et al., 2013).

1.1.2 CRISPR-Cas9 Origin & Development

Many bacterial and archaeal species have evolved CRISPR-Cas systems as a defense mechanism against invading phages and transferred plasmid (Jiang and Doudna, 2017; Mojica et al., 2005). A CRISPR-Cas system always comprises a CRISPR array, which consists of repetitive sequences that are interspaced by short variable sequences and an operon of several Cas genes. It represents an adaptive immune system, in which variable new spacers are acquired by processing and incorporation of little fragments of foreign genetic elements (protospacers) into the array (Barrangou et al., 2007). There are two classes with several types of CRISPR-Cas systems (Koonin et al., 2017). Since the class 2 type II CRISPR-Cas9 system is most widely used for genome editing, the functional principles of the CRISPR-Cas9 system are explained in the following.

The CRISPR array gets transcribed into a precursor CRISPR RNA (pre-crRNA). A separately transcribed trans-activating crRNA (tracrRNA) shows partial complementarity to the repeat sequences in the pre-crRNA. By base pairing to the pre-crRNA, it initiates the enzymatic cleavage of the pre-crRNA by endogenous RNase III into several single crRNA:tracrRNA duplexes (Deltcheva et al., 2011). The targeting crRNA component of this RNA duplex can base pair with protospacer sequences in invading DNA enabling recruitment of the Cas9 endonuclease to the phage DNA upon reinfection (Gasiunas et al., 2012; Jinek et al., 2012). Cas9 then cleaves the invading DNA and thereby destroys the phage genome (Garneau et al., 2010). The position of DNA cleavage is determined by the ~20 bp region of base pairing between crRNA and invader protospacer as well as the recognition of the so-called protospacer adjacent motif (PAM) sequence immediately downstream of the target region on the non-complementary DNA strand (Jinek et al., 2012).

In 2012, the groups of Jennifer A. Doudna and Emmanuelle Charpentier collectively adapted the type II CRISPR-Cas9 system of *Streptococcus pyogenes* (*Sp*) for the sequence specific cleavage of double stranded (ds) DNA and already highlighted the potential to exploit this system for RNA-programmable genome editing (Jinek et al., 2012). In this study, they also engineered a synthetic crRNA:tracrRNA chimera, termed single guide RNA (sgRNA), which can efficiently direct ds-DNA cleavage by

*Sp*Cas9. For this, they fused the 3´end of a mature crRNA via a synthetic loop to the 5´end of the tracrRNA mimicking the dual RNA structure needed for efficient Cas9 recruitment to the target locus. The resulting minimal two component system consisting of endonuclease and targeting sgRNA obviously proves more practical for genome engineering approaches and is the established standard format nowadays. Furthermore, it was shown that *Sp*Cas9 target DNA cleavage occurred three base pairs upstream of the PAM sequence (NGG for *Sp*Cas9) (Jinek et al., 2012).

With the first application of CRISPR-Cas9 technology in human and mouse cells in 2013 (Cong et al., 2013; Mali et al., 2013a), a new era of genome editing began making this extremely easy-to-use technique available for almost any laboratory. Two independent studies by the groups of Feng Zhang and George M. Church described the generation of a human codon optimized version of the *Sp*Cas9 gene with amino (N-) and/or carboxy (C-) terminal nuclear localization signals to ensure nuclear compartmentalization of *Sp*Cas9 in mammalian cells. They showed successful induction of indels as well as HDR-mediated integration of a DNA repair template at reporter and endogenous genes (Cong et al., 2013; Mali et al., 2013a). Importantly, both studies demonstrated that by simply supplying different sgRNAs simultaneously with the Cas9 endonuclease multiple genes can be targeted at once or defined regions of one gene can be cut out. This possibility of multiplexing and the ease of retargeting are enormous advantages of the CRISPR-Cas9 technology over previously used genome editing tools, which enable powerful new applications (Cong et al., 2013; Mali et al., 2013a). Shortly after, even sgRNA libraries were being used in CRISPR-Cas9-based genome-wide KO screens for the identification of new genes involved in cell viability as well as toxin or drug resistance (Koike-Yusa et al., 2014; Shalem et al., 2014, 2015).

With the publication of the crystal structure of *Sp*Cas9 in complex with sgRNA and ds-DNA target in 2014, novel insights about DNA binding and DSB generation were revealed (Nishimasu et al., 2014). *Sp*Cas9 has a bilobed architecture with an alpha-helical DNA recognition (REC) lobe and a nuclease (NUC) lobe. In between those two lobes there is a positively charged groove where the negatively charged DNA resides. The NUC lobe contains the conserved HNH domain, which cleaves the crRNA complementary strand of the DNA and split RuvC nuclease domains for cleavage of the non-complementary strand (Jinek et al., 2012; Makarova et al., 2011; Mali et al., 2013b; Sapranauskas et al., 2011).

Moreover, several groups developed algorithms for the design of specific and efficient sgRNAs and released open-source web tools (Haeussler et al., 2016; Hsu et al., 2013; Labun et al., 2016; Montague et al., 2014). By now, several companies

provide custom-made sgRNAs, Cas9 encoding vectors, mRNA and recombinant Cas9 protein. With all these tools at hand, CRISPR-Cas9 has within a short time become the most widely used gene editing method (Cong et al., 2013). Notably, the CRISPR-Cas9 genome editing technology is currently also being explored for the treatment of genetic disorders (Jiang and Doudna, 2017), such as Duchenne muscular dystrophy (Long et al., 2016; Nelson et al., 2016; Tabebordbar et al., 2016), sickle cell disease (DeWitt et al., 2016; Park et al., 2016), cystic fibrosis (Firth et al., 2015) as well as cardiovascular (Strong and Musunuru, 2017) and neurological diseases (Heidenreich and Zhang, 2016). Successful proof-of-concept studies demonstrated CRISPR-Cas9-mediated *in vivo* genome editing in various tissues like murine liver (Ding et al., 2014; Yin et al., 2014), muscle (Long et al., 2016) and brain (Swiech et al., 2015). Moreover, in the past five years CRISPR-Cas9 technology was extensively used for the generation of genetically modified animals across a broad range of species including worm (Friedland et al., 2013), fly (Bassett et al., 2013), zebrafish (Chang et al., 2013; Hwang et al., 2013; Jao et al., 2013), rat (Li et al., 2013), dog (Zou et al., 2015), pig (Hai et al., 2014), mouse (Wang et al., 2013) and monkey (Niu et al., 2014; Zuo et al., 2017). Recently, even RNA targeting class 2 type VI CRISPR-Cas systems, such as Cas13a (Abudayyeh et al., 2016, 2017; East-Seletsky et al., 2016) and Cas13d (CasRx) (Konermann et al., 2018) have been adapted for highly efficient and specific transcript silencing as well as targeted manipulation of alternative splicing events. Additional applications have arisen by developing nuclease-deficient Cas9 into an RNA guided shuttle for various effector domains.

1.1.3 Dead Cas9: Everyone´s Favorite Fusion Partner

Already in the pioneering paper of Jinek at al. in 2012, a nuclease deficient *Sp*Cas9 variant containing the inactivating points mutations D10A and H840A in its catalytic domains was introduced (Jinek et al., 2012; Sapranauskas et al., 2011). This mutant Cas9 is also referred to as catalytically "dead" Cas9 (dCas9). Introduction of only one of the mutations renders Cas9 into a DNA nickase (Cong et al., 2013; Jinek et al., 2012). The laboratory of George M. Church even created a Cas9 variant with up to four mutated amino acids putatively involved in magnesium coordination (D10A+D839A+H840A+ N863A) and no detectable nuclease activity as analyzed by deep sequencing of the targeted loci (Mali et al., 2013b).

Fusions of dCas9 to various effector domains for epigenetic modulation, transcriptional regulation, base editing (Gaudelli et al., 2017; Komor et al., 2016) as well as genome imaging (Chen et al., 2013; Tanenbaum et al., 2014) have been published in the past six years. The aim of epigenome editing lies in the locus-specific deposition or removal of epigenetic marks with the goal of altering

chromatin state and thereby gene expression. Fusions of dCas9 to the following epigenetic modifiers or catalytic domains thereof have been published: DNA methyltransferase DNMT3A (Huang et al., 2017; Liu et al., 2016; McDonald et al., 2016; Vojta et al., 2016), 5-methylcytosine dioxygenase TET1 (Choudhury et al., 2016; Liu et al., 2016; Xu et al., 2016), histone acetyltransferase p300 (Hilton et al., 2015), histone methyltransferase G9a (Agne et al., 2014) and histone demethylase LSD1 (Kearns et al., 2015).

Additionally, several dCas9 fusions have been developed to specifically up- or downregulate transcript expression levels (Figure 2). With CRISPRi, a modest degree of gene repression could be achieved in mammalian cells by simply blocking the transcription machinery with dCas9 alone (Qi et al., 2013). Targeting dCas9 to the promoter region sterically prevented transcription initiation while binding of the dCas9-sgRNA complex to the non-template DNA strand of the coding region hindered transcription elongation (Larson et al., 2013; Qi et al., 2013). This approach was developed further by fusing dCas9 to different repressive domains for the recruitment of chromatin modifying complexes that cause transcriptional silencing (Figure 2). The Krueppel associated box (KRAB) repression domain (Margolin et al., 1994) turned out to be the most efficient fusion partner enabling a gene repression of up to 15-fold (Gilbert et al., 2013). Recently, this fusion was further improved with the construction of the dCas9-KRAB-MECP2 repressor (Yeo et al., 2018). It was also suggested that CRISPRi could be used to map and interfere with the regulatory roles of distal and proximal enhancers (Gilbert et al., 2013; Thakore et al., 2015). Conversely, Gilbert et al. also showed the efficient transcriptional activation by dCas9 fusion constructs. For their so-called CRISPRa approach they successfully used a dCas9 fusion to VP64 (tetrameric repeat of *Herpes simplex* virus protein 16 (VP16) minimal activation domain (Beerli et al., 1998; Sadowski et al., 1988)) or the p65 activation domain of NF-κB (Schmitz and Baeuerle, 1991) to increase the expression of a reporter gene (Gilbert et al., 2013) (Figure 2). Several studies validated and characterized dCas9-VP64 as transcriptional activator (Maeder et al., 2013; Mali et al., 2013b; Perez-Pinera et al., 2013) or even expanded it to dCas9-VP160 (CRISPR-on) (Cheng et al., 2013), dCas9-VP192 (Balboa et al., 2015) or VP64-dCas9-BFP-VP64 (Chakraborty et al., 2014). From here on, artificial stimulation of gene transcription by recruiting diverse activation domains in *trans*, here via dCas9, to the target gene or its promoter will also be referred to as transactivation. A first significant improvement in gene activation techniques could be reached with development of dCas9-SunTag (Tanenbaum et al., 2014). In this approach, a chain of ten GCN4 peptide epitopes was fused to the C-terminus of dCas9. When single chain antibodies against the

GCN4 epitope each carrying a VP64 domain are co-expressed with dCas9-SunTag and the sgRNA, a recruitment of multiple VP64 domains by SunTag to the target locus and thus an increased gene activation could be achieved (Tanenbaum et al., 2014).

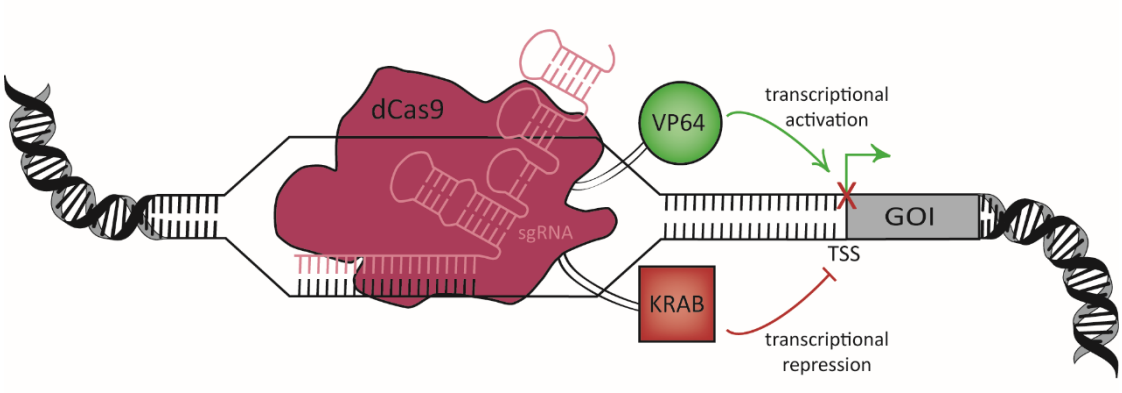


Figure 2 Transcriptional regulation by dead Cas9 fusion to effector domains
Targeting of dCas9-VP64 fusion constructs to the promoter region of a gene of interest (GOI) via a specific single guide RNA (sgRNA) results in transcriptional activation. In contrast, fusion of dCas9 to a KRAB domain can be used for transcriptional repression of the GOI.

Furthermore, to increase the transactivation efficiency, the recruitment of multiple transcriptional activators via the sgRNA has been explored (Konermann et al., 2015; Mali et al., 2013b; Zalatan et al., 2015; Zhang et al., 2015). To achieve this, bacteriophage MS2 RNA hairpin aptamers for the recruitment of MS2 coat protein (Bertrand et al., 1998; Witherell et al., 1991) were incorporated into the sgRNA sequence (Mali et al., 2013b). In the novel dCas9 synergistic activation mediator (SAM) system (Konermann et al., 2015), dCas9-VP64 is used in combination with a modified sgRNA (sgRNA 2.0) containing two MS2 hairpins. Per hairpin a dimer of MS2 coat protein can bind. Each of these MS2 proteins is in turn fused to one p65 and one human heat shock factor 1 (HSF1) activation domain (MS2-p65-HSF1). Right next to dCas9-VP64, this sgRNA design results in the binding of four additional activation domain dimers and consequently very potent induction of gene expression (Chavez et al., 2016; Konermann et al., 2015). Recently, another activator containing these same activation domains fused to dCas9, dCas9-VP192-p65-HSF1 (dCas9-VPH), was used for the reprogramming of primary human skin fibroblasts into iPSCs (Weltner et al., 2018). Zalatan et al. even demonstrated the simultaneous activation of one gene and repression of another gene by converting the sgRNA into a RNA scaffold that encodes both information, which DNA sequence to bind to and what regulatory domain to recruit (Zalatan et al., 2015). Also in the year 2015, George M. Church and colleagues presented their improved transcriptional activator design of a dCas9 fusion to the hybrid tri-partite transcriptional activator VP64-p65-

Rta (VPR) (Chavez et al., 2015). This combination of activator domains turned out to be the most efficient one in a systematic screen of 20 candidate domain combinations. Rta is a transcription factor expressed by Epstein-Barr virus during the immediate-early phase of the lytic cycle where it activates several genes required for lytic progression (Chang et al., 2005). Notably, dCas9-VPR was even strong enough to mediate robust and rapid differentiation of human induced pluripotent stem cells (iPSCs) into a neuronal phenotype by highly activating either neurogenin 2 or neurogenic differentiation factor 1 expression (Chavez et al., 2015).

Such a range of efficient epigenetic and transcriptional modifiers enabling precise control over gene expression decisions might have an enormous therapeutic potential. This is particularly relevant for diseases characterized by a dysregulation of the epigenome including Prader-Willi syndrome, fragile X syndrome or cancer (Holtzman and Gersbach, 2018). In cancer for example, it could be possible to selectively shut down oncogenes or, vice versa, reactivate tumor suppressor genes that are silenced in some sorts of cancer (Jiang and Doudna, 2017). Furthermore, epigenetics is believed to play a crucial role in learning, memory formation, addiction and psychopathology (Heidenreich and Zhang, 2016; Sweatt, 2013) opening possible therapeutic applications of the dCas9-based modifiers in neuropsychiatric disorders or neurodegenerative diseases.

1.1.4 Cas9 in Gene Therapy

Up to now, recombinant adeno-associated viruses (rAAV) represent the preferred gene delivery vehicle in gene therapy approaches. This is due to their broad cell tropism and superior transduction efficiency in dividing and postmitotic cells combined with an excellent safety profile (Basner-Tschakarjan and Mingozzi, 2014; Mingozzi and High, 2011). Notably, rAAVs exhibit a low mutagenic risk since they mainly reside in an episomal state in the transduced cells and if at all, integrate into safe regions of the host genome (Kaeppel et al., 2013; Ramlogan-Steel et al., 2018). Furthermore, rAAVs are easy to produce at large scale and depending on the capsid variant used for packaging, rAAVs can be employed to selectively infect distinct cell types in a tissue (Michalakis et al., 2018; Palfi et al., 2015; Schön et al., 2015). However, the biggest obstacle in rAAV based gene therapy is their limited packaging capacity of about 4.7 kb (Zinn and Vandenberghe, 2014). As the *Sp*Cas9 coding sequence alone is ~4.2 kb, together with sgRNA and regulatory sequences it is barely possible to fit all components needed for genome editing into a single rAAV vector. Several strategies have been explored to overcome this limitation. One approach was to identify novel smaller Cas9 orthologues from other bacterial strains, such as ~3.3 kb *St*1Cas9 from *Streptococcus thermophilus* LMD-9 CRISPR1

(Cong et al., 2013; Esvelt et al., 2013), ~3.2 kb *Nm*Cas9 from *Neisseria meningitidis* (Esvelt et al., 2013), ~3.2 kb *Sa*Cas9 from *Staphylococcus aureus* (Ran et al., 2015) and ~3 kb *Cj*Cas9 from *Campylobacter jejuni* (Kim et al., 2017). However, most orthogonal Cas9 proteins have more complex PAM requirements limiting the density of possible target sites per genome compared to *Sp*Cas9 whose NGG restricts the target space to on average every 8 bp in the human genome (Cong et al., 2013). Moreover, CRISPR from *Prevotella* and *Francisella* (Cpf1) from the novel class 2 type V CRISPR-Cas12a system emerged as an alternative more compact single RNA-guided endonuclease (Zetsche et al., 2015a, 2017). A rationally designed truncated version of *Sp*Cas9 unfortunately only exhibited ~50 % of full-length *Sp*Cas9 efficiency (Nishimasu et al., 2014).

Another strategy to overcome the low AAV cargo size is based on separation of *Sp*Cas9 into two fragments for split Cas9 delivery by a dual rAAV vector system. Splitting Cas9 allows for more complex and longer genetic elements to be packaged into an rAAV vector including tissue-specific promoters, multiple sgRNA cassettes and dCas9 effector domain fusions (Fine et al., 2015). Different approaches were chosen to enable a reconstitution of the Cas9 fragments into functional full-length Cas9 in the cell. Zetsche et al. reported the fusion of split Cas9 to rapamycin-binding dimerization domains. In this approach, they used subcellular separation of the Cas9 halves combined with rapamycin-induced dimerization and nuclear transport to achieve an inducible activation of Cas9 endonuclease or dCas9-VP64 transcriptional control (Zetsche et al., 2015b). Nihongaki et al. fused split Cas9 fragments to photoinducible dimerization domains to engineer photoactivatable Cas9 for optogenetic control of genome editing (Nihongaki et al., 2015). Another strategy resulting in scarless reconstitution of full-length Cas9 is the fusion of *Sp*Cas9 fragments to split intein sequences (Chew et al., 2016; Truong et al., 2015). Split intein-mediated protein *trans*-splicing was already successfully applied in rAAV-based therapeutic supplementation of other large genes (Li et al., 2008). Inteins can catalyze their own excision from an amino acid sequence rejoining the remaining flanking polypeptides (extein sequence) with a peptide bond. Truong et al. used the N- and C-intein parts of DNA polymerase III DnaE from *Nostoc punctiforme* and exchanged the original extein sequence with the Cas9 fragments. For the first time, they could show functional reconstitution of *Sp*Cas9 from a dual rAAV vector system additionally containing sgRNA sequences and a HDR template (Truong et al., 2015). Fine et al. split the *Sp*Cas9 sequence after tyrosine 656 (Y656) and fused the fragments to the Gyrase A intein from *Mycobacterium xenopi*. Although the Y656 split variant did induce DNA DSB in HEK293T cells upon co-transfection, this study failed to reproduce functionality of split Cas9 in other cell types (Fine et al., 2015).

However, two other studies showed successful splitting of *Sp*Cas9 between glutamate 573 (E573) and cysteine 574 (C574) (Truong et al., 2015; Zetsche et al., 2015b). Furthermore, Ma et al. presented a split site after lysine 1153 in the PAM-interacting domain of *Sp*Cas9 as their most promising candidate (Ma et al., 2016). In addition, splitting of *Sp*Cas9 between valine 713 (V713) and serine 714 (S714) in the disordered linker, which connects the REC and NUC lobe, was proven very efficient (Chew et al., 2016; Nihongaki et al., 2015). For *in vivo* gene editing, Chew et al. delivered fusions of the N-terminal *Sp*Cas9 lobe with the N-intein of *Rhodothermus marinus* (*Rma*) and the C-terminal lobe with *Rma* C-intein by injection of dual rAAV particles. Even the large transcriptional activator Cas9-VPR (5.8 kb) was already successfully split and fit into two rAAV vectors capable of inducing *in vitro* and *in vivo* gene activation (Chew et al., 2016) (Figure 3).

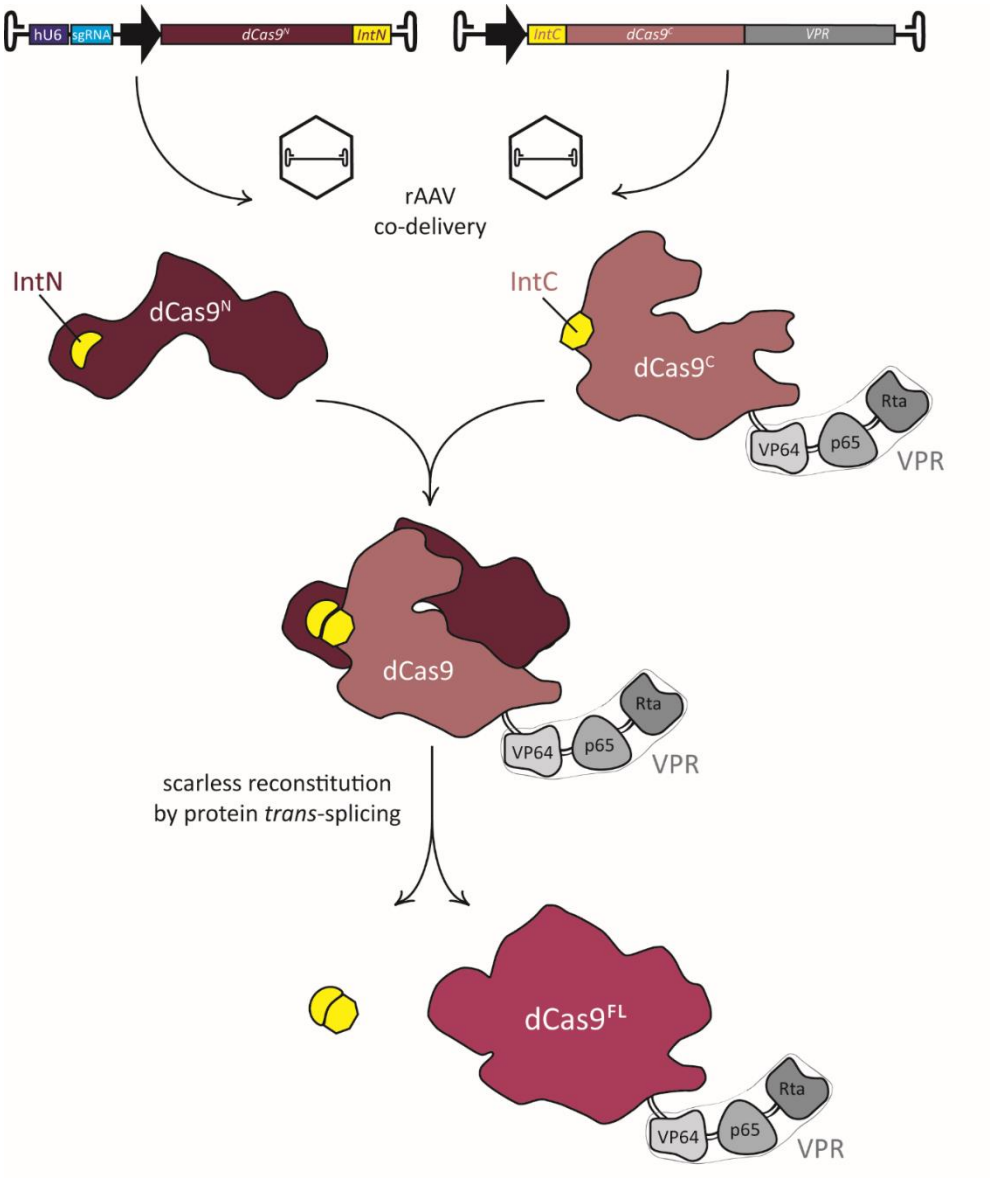


Figure 3 Reconstitution of full-length dCas9-VPR by split intein-mediated protein trans-splicing

Co-expression of both, N-terminal fragment of dCas9 ($dCas9^N$) fused to the N-intein domain (IntN) and C-terminal dCas9 fragment ($Cas9^C$) with the VPR transactivation domains fused to the C-intein (IntC) is achieved using a dual vector rAAV system. Scarless reconstitution of full-length (FL) dCas9-VPR can be catalyzed by split intein-mediated protein *trans*-splicing.

1.2 Retina: From Photon to Electrical Signal

1.2.1 Structure & Functioning of the Retina

The retina is the photosensitive innermost layer of the eye where visual stimuli are transformed into electrical signals. The neuronal retina consists of three main layers comprising different neuronal types, which are interconnected by synapses. The outer nuclear layer is formed by cell bodies of the photoreceptors, the highly specialized cell type responsible for the detection of light (Müller, 1854). There are two types of photoreceptors that fulfill distinct tasks in the retina. Cone photoreceptors enable daylight (photopic) vision, which is marked by high visual acuity and color discrimination, while rod photoreceptors mediate dim light (scotopic) vision (Schultze, 1866). In primates, trichromatic color vision is mediated by three distinct types of cone photoreceptors with unique spectral sensitivities (Hecht, 1930; Jacobs, 2018). The adjacent inner nuclear layer contains the nuclei of bipolar, horizontal, amacrine and Müller glial cells. Bipolar cells establish a direct connection between the synaptic terminals of photoreceptor cells and retinal ganglion cells, which form the innermost layer of the retina. While horizontal cells have the single broad function of lateral inhibition at the first retinal synapse, amacrine cells transmit and shape light responses at the second retinal synapse (Masland, 2001). The axons of retinal ganglion cells form the optic nerve which leaves the eye and transmits the visual information to the brain.

The phototransduction cascade is initiated upon the incidence of light onto the visual pigments located in photoreceptor outer segments. Rhodopsin (RHO) is the photopigment of rod photoreceptors whereas the three human types of cones contain three spectrally different cone opsins. Short (S), medium (M) and long (L) wavelength-sensitive cones express S-, M- and L-opsin, respectively (Yokoyama, 2000). Mice only have two cone types with short wavelength sensitive pigments showing an absorption maximum in the ultraviolet (S-opsin) and middle wavelength sensitive (M-opsin) pigments (Jacobs et al., 1991).

While rods and cones express distinct sets of homologous key proteins, the principle signaling process initiated by light (transduction cascade) is similar in both cell types and shall be here described for rods (Figure 4). In the dark, constant activity of transmembrane guanylyl cyclases results in high levels of cyclic guanosine

monophosphate (cGMP) (Yang et al., 1995) keeping cyclic nucleotide-gated (CNG) ion channels in an open state (Fesenko et al., 1985). These CNG channels facilitate the constant influx of sodium (Na^+) and calcium (Ca^{2+}) ions, also referred to as dark current. This steady influx of cations depolarizes the photoreceptor promoting synaptic glutamate release. Rhodopsin is a G-protein coupled receptor, so when photons hit the rhodopsin molecule a G-protein-mediated signaling cascade is initiated (Palczewski et al., 2000). In detail, the photon causes the opsin bound chromophoric group 11-*cis*-retinal to isomerize to all-*trans*-retinal initiating the remodeling of rhodopsin into the active intermediate metarhodopsin II. This activation results in a nucleotide exchange at the heterotrimeric G-protein transducin and subsequent dissociation of the activated transducin α-subunit (Choe et al., 2011). The α-subunit in turn activates the cGMP phosphodiesterase (PDE6), which catalyzes the hydrolysis of cGMP to 5´-GMP. The decreasing cGMP levels result in the closure of CNG channels. Since the influx of Na^+ and Ca^{2+} is reduced, the photoreceptor hyperpolarizes leading to a reduced transmitter release at its synaptic terminal. In addition, the reduced intracellular Ca^{2+} concentration is essential to initiate several mechanisms of photoreceptor regeneration including phosphorylation of rhodopsin and PDE6 as well as the activation of guanylyl cyclases for restoration of cGMP levels (Michalakis et al., 2018).

The possibility of substituting rod-specific components of the phototransduction cascade with the cone-specific homologs or vice versa has been investigated before. By generating transgenic mice that express S-opsin in rods it could be shown that S-opsin could substitute the structural and functional role of rhodopsin (Shi et al., 2007). S-opsin correctly located to the outer segment where it was able to activate rod transducin and be inactivated by rhodopsin kinase and rod arrestin. Importantly, expression of S-opsin in $Rho^{-/-}$ rods partially reversed the effects of Rho deletion by promoting outer segment formation and cell viability. Moreover, the S-opsin expression enabled the $Rho^{-/-}$ rods to respond to light, although with shifted spectral sensitivity corresponding to S-pigment (Shi et al., 2007).
In addition, rod and cone transducin α-subunits were shown to be functionally interchangeable (Deng et al., 2009) and the expression of the cone PDE6 α´-subunit was sufficient to restore rod function in a PDE6β deficient mouse model (Deng et al., 2013).

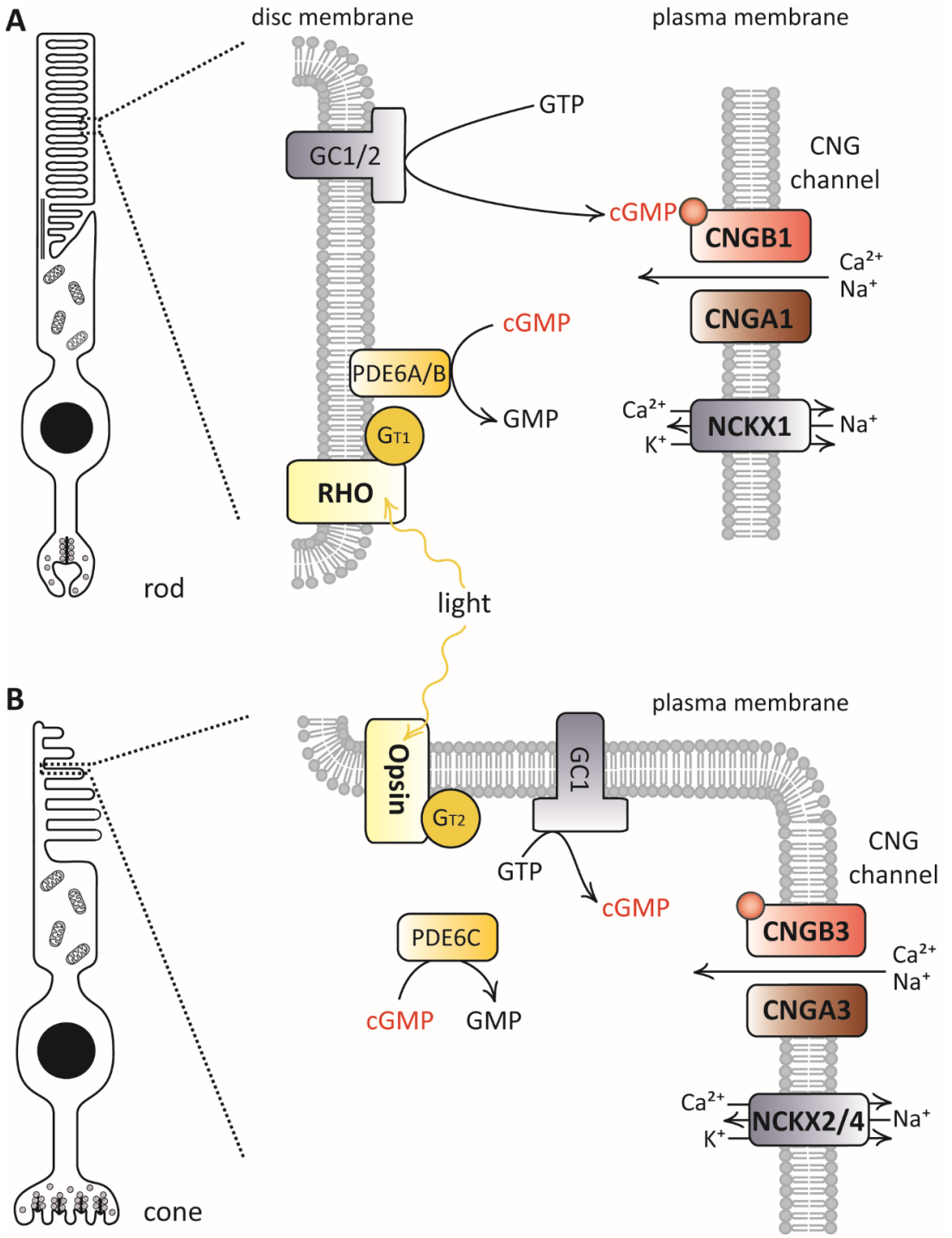


Figure 4 Phototransduction cascade in rod and cone photoreceptors
The principal components of the phototransduction cascade in the outer segments of ***(A)*** rod and ***(B)*** cone photoreceptors are depicted. In the dark, membrane-bound guanylyl cyclases (GC) produce cGMP from GTP. Binding of cGMP to the cyclic nucleotide gated (CNG) channel enables the influx of Ca^{2+} and Na^{+} ions in the dark state. Upon the incidence of light onto the visual pigments rhodopsin (RHO) in rods or cone opsins, a G protein-mediated cascade is initiated. The G protein transducin (G_T) activates a phosphodiesterase (PDE6), which hydrolyzes cGMP to GMP. Decreasing cGMP levels result in the closure of CNG channels. Since Ca^{2+} is continuously exported from the cell by the Na^{+}/Ca^{2+}-K^{+} exchanger (NCKX), the

decreased cation influx via the CNG channel causes a hyperpolarization of the photoreceptor. Modified from (Michalakis et al., 2018).

1.2.2 CNG Channels

The mammalian CNG channel family comprises CNGA1-4, CNGB1 and CNGB3. CNGA2, CNGA4 and CNGB1b (protein encoded by splice variant b of *CNGB1*) subunits assemble into the CNG channel of olfactory receptor neurons, which are essential for odor reception (Munger et al., 2001). In contrast, the expression of *CNGA1* and *CNGB1a* as well as *CNGA3* and *CNGB3* is restricted to rod and cone photoreceptors, respectively (Biel and Michalakis, 2007). Three A subunits (CNGA1 in rods, CNGA3 in cones) and one B subunit (CNGB1 in rods, CNGB3 in cones) assemble into photoreceptor specific hetero-tetrameric channels that localize to the disc membranes (Shuart et al., 2011; Weitz et al., 2002; Zheng et al., 2002; Zhong et al., 2002). However, *in vitro* different combinations of CNG subunits can form functional CNG channels and heterologously expressed CNGA1, CNGA2 and CNGA3 proteins can even form functional homo-tetrameric channels (Kaupp and Seifert, 2002; Michalakis et al., 2018). All channel subunits share the same membrane topology of six transmembrane α-helices, a reentrant pore loop between transmembrane segment 5 and 6 as well as cytosolic N and C termini. In the assembled CNG channel, the pore loops and sixth transmembrane domains of the four subunits line the ion conducting pore (Biel and Michalakis, 2007; Li et al., 2017). The so-called C linker couples the channel gate in transmembrane domain 6 to the C-terminal cyclic nucleotide binding domain and was shown to be the determining factor for channel gating (Zong et al., 1998). CNG channels are permeable for several mono- and divalent cations but Ca^{2+} and Mg^{2+} can also cause a voltage-dependent blockage of monovalent cation flux (Dzeja et al., 1999; Frings et al., 1995).

1.2.3 Inherited Retinal Degenerative Diseases & Gene Therapy

In general, mutations in genes involved in the visual transduction cascade or required for proper photoreceptor architecture are often associated with inherited retinal degenerative diseases (Verbakel et al., 2018). Mutations in genes specifically expressed in rods, such as *CNGA1* (Dryja et al., 1995), *CNGB1* (Bareil et al., 2001) or *RHO* (Dryja et al., 1990; al-Maghtheh et al., 1993), have been found in patients suffering from retinitis pigmentosa (RP). In RP, the initial loss of rod photoreceptor function causes tunnel vision and night blindness. Moreover, this stage is often followed by the progressive degeneration of primarily not affected cone photoreceptors, which eventually leads to complete blindness (Dryja et al., 1995). Importantly, mouse models like $Cngb1^{-/-}$ (Hüttl et al., 2005; Zhang et al., 2009) or $Rho^{-/-}$ (Humphries et al., 1997; Lem et al., 1999) were generated that closely

resemble the RP phenotype and enable the *in vivo* validation of potential therapeutic approaches (Michalakis et al., 2014).

Mutations in cone-specific proteins like *CNGA3* (Kohl et al., 1998) and *CNGB3* (Kohl et al., 2000; Sundin et al., 2000) were shown to cause achromatopsia (ACHM). More than 100 mutations in *CNGA3* and more than 40 mutations in *CNGB1* have been identified in ACHM patients so far (Remmer et al., 2015). This autosomal recessive eye disease is characterized by a congenital absence of cone photoreceptor function causing symptoms such as low visual acuity, total color blindness, photophobia and nystagmus (Michalakis et al., 2018; Sundin et al., 2000). ACHM mouse models showing a progressive cone degeneration were generated by Cnga3 KO (Biel et al., 1999; Michalakis et al., 2005) and Cngb3 KO (Ding et al., 2009; Xu et al., 2011).

No treatment has been developed yet to cure RP or ACHM and only some amelioration of the symptoms is achieved by vision aids. Nevertheless, several promising gene therapy approaches are ongoing (Jiang et al., 2018). For the autosomal recessively inherited CNG channelopathies, therapeutic approaches mainly focus on rAAV-mediated gene supplementation providing treated cells with a healthy copy of the mutated gene (Michalakis et al., 2018). For *CNGB1*-linked RP, proof-of-concept studies demonstrate beneficial effects of rAAV-based gene therapy in mice (Michalakis et al., 2014) and dogs (Petersen-Jones et al., 2018). For *CNGA3*- and *CNGB3*-linked ACHM, in addition to several successful pre-clinical studies (Banin et al., 2015; Carvalho et al., 2011; Komáromy et al., 2010; Michalakis et al., 2010), even four phase I/II clinical trials for rAAV-based gene therapies (Fischer et al., 2016) are currently in progress (Michalakis et al., 2018). RHO-linked RP is mostly inherited in an autosomal dominant manner and often a toxic gain-of-function phenotype or dominant-negative effects of misfolded mutant RHO on the wild type (WT) protein are observed (Mendes and Cheetham, 2008). Although studies exist that show positive effects for rAAV-mediated RHO expression in $Rho^{-/-}$ mice (Palfi et al., 2015), just providing a healthy copy of *RHO* alone might not be sufficient for a long term rescue of the affected photoreceptors. In addition to gene supplementation, expression of mutated *RHO* needs to be suppressed (Rossmiller et al., 2012). Moreover, well-balanced Rho expression levels are essential, since an overexpression of WT Rho was shown to cause RP-like retinal degeneration in healthy mice (Olsson et al., 1992; Tan et al., 2001). In summary, *RHO* gene therapy proves to be much more challenging possibly explaining the lack of current clinical studies.

Notably, in 2017, the first rAAV-based gene therapy product for the treatment of biallelic RPE65-linked retinal dystrophy (LUXTURNA™) was approved by the U.S.

Food and Drug Administration highlighting the fast and promising development in the field of retinal gene therapy (Michalakis et al., 2018).

1.2.4 iNGNs as an *In Vitro* Model of Neuronal Differentiation

The human inducible Neurogenin iPS (iNGN) cell line (Busskamp et al., 2014) contains a doxycycline (DOX) inducible bicistronic neurogenin expression cassette stably integrated into the genome. Application of DOX leads to the overexpression of neurogenin 1 and neurogenin 2 triggering the rapid differentiation of iNGN cells into neurons within only four days (Figure 5). The resulting homogeneous neuronal population mainly consists of cells with bipolar-shaped morphology and expresses neuronal marker proteins like microtubule associated protein 2 and synapsin 1 (Busskamp et al., 2014).

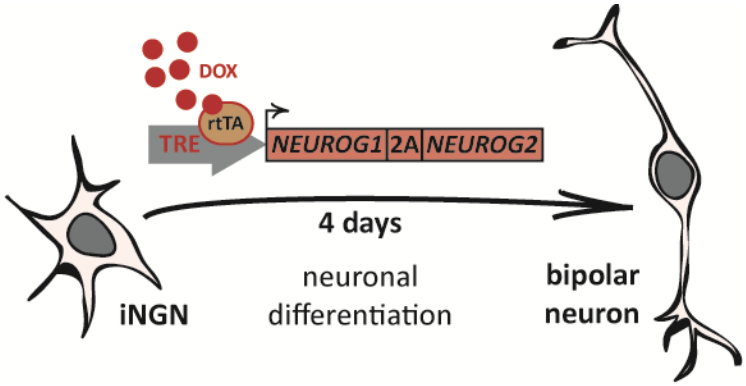


Figure 5 Neuronal differentiation of iNGN cells
iNGN cells are human iPS cells that contain an inducible neurogenin expression cassette integrated in their genome. In the presence of doxycycline (DOX) the reverse tetracycline-controlled transactivator (rtTA) can bind to the tetracycline response element (TRE) resulting in the expression of neurogenin 1 and neurogenin 2. This overexpression in turn triggers iNGN differentiation into bipolar neurons within four days.

When co-cultured with astrocytes, iNGN-derived neurons become electrically active within 14 days being capable of firing trains of action potentials. Even spontaneous postsynaptic currents were detected in mature neuronal cultures indicating functional synaptic connections (Busskamp et al., 2014; Lam et al., 2017). Comparison of transcriptional expression profiles of d4 differentiated iNGNs to human brain showed increased similarity to *in vivo* neuronal development validating the iNGN cell line as a suitable *in vitro* model for neuronal differentiation (Busskamp et al., 2014).

1.3 Epigenetics

The term epigenetics comprises inheritable changes in a chromosome, which define cellular gene expression without, however, altering the actual nucleotide sequence (Berger et al., 2009; Weinhold, 2006). Different epigenetic landscapes lead to different gene expression patterns and thus explain the development of a vast range of highly specialized cell types and tissues that in principle all contain the same DNA (Waddington, 1957).

Chromatin structure is largely determined by histones, the basic proteins enabling the massive compaction of the genomic DNA within the nucleus (McGhee and Felsenfeld, 1980). As a first level of genomic compaction, DNA is organized in nucleosomes (Kornberg, 1974). In a nucleosome, 145-147 bp of DNA are wrapped around an octameric core complex comprising a pair of each core histone H2A, H2B, H3 and H4 (Luger et al., 1997). In between nucleosome cores, there are short segments of free DNA, called linker DNA, which associate with the linker histone H1 (McGinty and Tan, 2015). Histone H1 is involved in the higher-order compaction of the nucleosomal array into the 30 nm chromatin fiber, which is then subject to the formation of large-scale tertiary structures that build an entire chromosome (Tremethick, 2007). Chromatin can reside in two forms: either euchromatin or heterochromatin. Euchromatin represents a loosely packed more open chromatin state that is considered transcriptionally active. In contrast, heterochromatin is tightly packed chromatin, which is not accessible to the transcription machinery (Heitz, 1928; Kebede et al., 2015). The chromatin structure varies with the dynamic deposition and removal of various epigenetic marks that are often associated with distinct gene expression patterns.

1.3.1 Histone Modifications

There are two main mechanisms by which epigenetic information is stored: covalent histone modifications and DNA methylation. Already in the 1960s, post-translational modifications of histone proteins have been discovered (Allfrey et al., 1964). These post-translational modifications of histones mainly occur at amino acid residues in the N-terminal histone tails protruding from the core octamer and highly affect inter-nucleosomal interactions and thus the overall chromatin state (Bannister and Kouzarides, 2011). By now, many modifications on several amino acids of histones have been identified, including histone acetylation, phosphorylation, methylation, ubiquitinylation, deimination, ADP ribosylation, β-N-acetylglucosamine modification, sumoylation, crotonylation, propionylation, butyrylation, malonylation, and succinylation (Bannister and Kouzarides, 2011;

Chen et al., 2007; Kebede et al., 2015; Tan et al., 2011; Xie et al., 2012). As numerous as these modifications are, as diverse are their combinatorial effects on chromatin packaging and gene expression leading to the development of the histone code theory (Strahl and Allis, 2000). Unlike acetylation and phosphorylation, which mainly exhibit their effects by adding a negative charge to the histone and thereby disrupting charge interactions between histone and DNA, other modifications act via more complex mechanisms. Some histone modifications alone or in combination with other epigenetic marks recruit effector proteins including chromatin remodeling enzymes that catalyze the adenosine triphosphate (ATP)-dependent repositioning of nucleosomes, thereby altering chromatin density (Bannister and Kouzarides, 2011; Musselman et al., 2012).

1.3.2 DNA Methylation

Next to epigenetic information being encoded in histone proteins, DNA methylation serves as an epigenetic mark in the DNA itself (Johnson and Coghill, 1925). In mammals, two forms of naturally occurring DNA methylation marks are known, either 5-methylcytosine (5mC) (Wyatt, 1951) or the recently discovered N6-methyladenine (Wu et al., 2016). Cytosine methylation mainly occurs in the context of cytosine-guanine dinucleotides, referred to as CpGs. However, especially in embryonic stem cells (ESC), iPSCs and neurons, substantial amounts of 5mC were also shown in a non-CpG context (Jang et al., 2017; Lister et al., 2009, 2013). The transfer of a methyl group to the 5th carbon atom of cytosine is catalyzed by DNA methyltransferases (DNMTs) using S-adenosyl-L-methionine as a methyl donor (Chen et al., 1991; Doerfler, 1983). There are four members of the DNMT family: DNMT1, DNMT3A, DNMT3B and DNMT3L. DNMT3L is the only member which does not exhibit enzymatic activity (Aapola et al., 2000; Bourc'his et al., 2001; Hata et al., 2002). DNMT1 acts as a maintenance methyltransferase on hemi-methylated CpGs resulting from DNA replication where it methylates the unmodified cytosines in the newly synthesized daughter DNA strand and thereby preserves existing methylation patterns (Bestor et al., 1988; Robertson et al., 1999). In contrast, DNMT3A and DNMT3B can also catalyze *de novo* DNA methylation of unmodified cytosines (Okano et al., 1998). Generally, it is believed that the accumulation of 5mC in the promoter or gene body is a repressive epigenetic mark associated with low or no expression of the respective gene (Holliday and Pugh, 1975; Mcghee and Ginder, 1979; Taylor and Jones, 1979). How this gene repression is exactly induced by 5mC is still unknown. It was postulated that 5mC in the promoter interferes with the binding of transcription factors necessary for gene transcription (Becker et al., 1987; Cedar, 1988). Alternatively, it is believed that methylated CpGs might be specifically recognized by methyl-CpG binding proteins. Binding of these proteins to

5mC could either directly block transcription of the related gene or recruit co-repressor proteins for chromatin condensation (Bird, 1992; Boyes and Bird, 1991; Buschhausen et al., 1987; Jones et al., 1998). However, DNA methylation is essential for proper cell differentiation and embryonic development, irreversible promoter silencing of imprinted genes and genes subjected to X chromosome inactivation as well as repression of retrotransposons (Bestor, 2000; Kareta et al., 2006; Li et al., 1992; Riggs, 1975). Moreover, aberrant DNA methylation patterns are involved in many diseases including cancer, autoimmune diseases, metabolic and neurological disorders (Jin and Liu, 2018; Robertson, 2005). Hence, understanding the 5mC deposition and removal dynamics has always been of high interest.

1.3.3 DNA Demethylation by TET Enzymes

In 2009, it was discovered that the methyl group of 5mC can be oxidized by a family of iron (FeII) and 2-oxoglutarate (2OG) dependent methylcytosine dioxygenases, called ten-eleven translocation (TET) enzymes (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). This enzyme family includes three members: TET1, TET2 and TET3. All TET enzymes contain a C-terminal catalytic domain, which comprises a conserved cysteine (Cys)-rich region followed by a double stranded β-helix (DSBH) fold core region (Ito et al., 2010; Kohli and Zhang, 2013). The DSBH catalytic core contains the crucial residues for the binding of FeII and 2OG while the Cys-rich domain is indispensable for proper stabilization of the DNA above the DSBH core (Hu et al., 2013). TET enzymes catalyze the gradual oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) and further to 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) (Ito et al., 2010, 2011; Tahiliani et al., 2009) (Figure 6). Importantly, 5fC and 5caC are targets for a rapid excision by the thymine DNA glycosylase (TDG) and subsequent base excision repair (BER) results in unmodified cytosine (He et al., 2011; Maiti and Drohat, 2011). Hence, this TET-mediated oxidation process is considered as a mechanism for active DNA demethylation. However, a recent study also proposed that 5fC and possibly also 5caC are converted into unmodified cytosine by a direct C-C bond cleavage (Iwan et al., 2018). Moreover, 5hmC, 5fC and 5caC can be turned into unmodified cytosine by replication dependent dilution, a pathway called “active modification and passive dilution”. In contrast, the term passive DNA demethylation refers to the loss of 5mC during consecutive replication rounds in the absence a functional DNA methylation maintenance machinery (Kohli and Zhang, 2013).

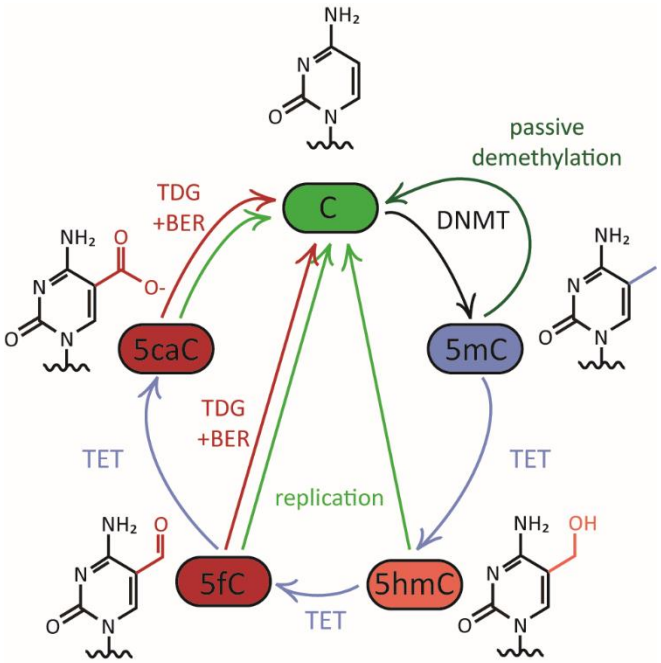


Figure 6 Pathways of DNA demethylation
DNA methyltransferases (DNMTs) catalyze the transfer of a methyl group to the 5th position of cytosine (C) resulting in 5-methylcytosine (5mC). As part of the active DNA demethylation pathway, 5mC can be oxidized by TET enzymes to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxycytosine (5caC). 5fC and 5caC can be excised from the DNA by thymine DNA glycosylase (TDG) and subsequent base excision repair (BER) results in unmodified cytosine. Moreover, 5hmC, 5fC and 5caC get diluted by DNA replication and 5mC can be lost due to passive demethylation.

The removal of 5mC from a gene or its regulatory regions is commonly believed to lead to an activation of gene expression. While initially, 5hmC was considered only an intermediate of DNA demethylation, it is nowadays believed to be a stable epigenetic mark (Hahn et al., 2013; Münzel et al., 2010, 2011). Genomic levels of 5hmC were shown to increase with neurogenesis where its accumulation in gene bodies, promoters and enhancers correlates positively with gene expression (Colquitt et al., 2013; Hahn et al., 2013; Lister et al., 2013; Melamed et al., 2018; Szulwach et al., 2011). Concurrently, in comparison to several other murine tissues genomic 5hmC was shown to be most abundant in brain correlating with TET3 expression levels (Globisch et al., 2010; Khare et al., 2012; Szwagierczak et al., 2010). While TET1 and TET2 are highly expressed in mouse ES cells and primordial germ cells (Ito et al., 2010; Koh et al., 2011), TET3 was shown to be the major TET enzyme in neurons of brain and retina (Colquitt et al., 2013; Hahn et al., 2013; Perera et al., 2015; Szwagierczak et al., 2010).

1.3.4 TET3 in Neurons

In mammals, TET3 exists in several isoforms generated by alternative splicing and differential promoter usage (Jin et al., 2016; Liu et al., 2013; Melamed et al., 2018) (Figure 7).

Figure 7 Human TET3 transcript variants
In humans, TET3 exists in three main isoforms. The two longer transcripts, named *TET3* full-length (*TET3FL*), contain a N-terminal CXXC DNA-binding domain and only differ in the inclusion or exclusion of exon 2 by alternative splicing. The expression of the short *TET3* variant is initiated from an alternative promoter resulting in the *TET3S* isoform lacking the CXXC domain. Light blue regions mark the 5´ and 3´untranslated regions, while the coding sequence is depicted in dark blue.

Both *TET3* full-length (*TET3FL*) splice variants encode a N-terminal CXXC DNA binding domain but differ in the splicing of exon 2. The expression of another shorter *TET3* isoform is initiated from an alternative promoter resulting in *TET3S* without the CXXC domain (Liu et al., 2013). In mice, another oocyte-specific *Tet3* splice variant lacking the CXXC domain was identified and termed *Tet3o* (Jin et al., 2016). *Tet3S* was shown to be the prevailing *Tet3* isoform in the murine retina where it plays a crucial role in retinal development and maturation. TET3S is targeted to the DNA by interacting with RE1-silencing transcription factor (REST) for 5hmC generation and induction of histone 3 lysine 36 trimethylation resulting in the activation of neuronal genes (Perera et al., 2015). Also Liu et al. demonstrated that in different brain regions of the mouse *Tet3S* was higher expressed than *Tet3FL* supporting the dominant role of Tet3S in neurons (Liu et al., 2013).

2 Aim of the Study

In this study, the CRISPR-Cas9 system was used in two ways. On the one hand, classical genome editing was performed to knock out catalytically active TET3 protein in the human iNGN cell line. On the other hand, dCas9-based transcriptional activators, such as dCas9-VPR, were tested for the ectopic activation of photoreceptor specific genes.

The TET3KO iNGN project aims at generating and validating a TET3KO iNGN cell line. This cell line enables detailed investigation of the role of TET3 in neuronal differentiation. The newly generated cell line should be characterized in detail regarding TET expression and neuronal differentiation. Of special interest are the effects of the TET3KO on the methylome and transcriptome. Global and site-specific changes in cytosine modifications as well as the resulting changes in gene expression should be analyzed.

The aim of the transactivation project was the development and *in vitro* assessment of a novel approach of ectopic activation of homologous genes in photoreceptors. Rod and cone photoreceptors share a set of similar but distinct proteins involved in the phototransduction cascade. As a novel therapeutic approach for inherited retinal degenerative diseases, the transcriptional activation of the healthy homologous gene by transcriptional activators could compensate the mutation-induced functional deficit in diseased photoreceptors. The potential of dCas9-VPR-mediated transactivation of suitable candidate genes in different cell lines should be evaluated *in vitro*. This includes testing the activation efficiency of rod-specific cyclic nucleotide gated channel subunit A1 in cone-like 661W cells as well as cone-specific M- and S-opsin in mouse embryonic fibroblasts. Moreover, to pave the way for a future rAAV-mediated gene therapy, the concept of splitting the transcriptional activator dCas9-VPR into two fragments for dual rAAV vector delivery and subsequent reconstitution of functional full-length dCas9-VPR by split intein-mediated protein *trans*-splicing should be tested.

3 Material & Methods

All used chemicals were obtained from VWR, Sigma-Aldrich, Merck, Bio-Rad, Fluka or Roth if not stated otherwise. They had "*pro analysi*" or "for molecular biological use" quality. For all solutions, high pure and deionized water with a resistivity of 18.2 MΩ x cm at 25 °C from the Milli-Q Plus System (Millipore) was used.

3.1 Cloning

Cloning was performed either using classical restriction digest followed by ligation or restriction enzyme independent methods such as Gibson cloning (Gibson et al., 2009) and commercial cloning kits like NEBuilder (New England Biolabs (NEB)). All FastDigest restriction enzymes, Fast-Alkaline Phosphatase and T4 Ligase were obtained from Fermentas (Thermo Fisher Scientific). Insert sequences were either produced by restriction digest, polymerase chain reaction (PCR) amplification from existing plasmids or commercially synthesized (BioCat GmbH). The following Addgene plasmids were used (Table 1).

Table 1 Addgene plasmids
List of used Addgene plasmids including plasmid number, vector backbone, expressed feature and name of the depositor lab.

Plasmid name	**Addgene number**	**Vector backbone**	**Expressed feature**	**Depositor**	**Citation**
Cas9m4-VP64	47319	pcDNA3.3	Cas9m4-VP64	George Church	(Mali et al., 2013b)
SP-dCas9-VPR	63798	pcDNA3.3	dCas9-VPR		(Chavez et al., 2015)
pSMVP-Cas9N	80934	pMAX	$Cas9^{N}$		(Chew et al., 2016)
pSMVP-Cas9C	80939	pMAX	$Cas9^{C}$		
pAAV-CMV-Cas9C-VPR	80933	pZac2.1	Cas9C-VPR		

PCR primers were designed using the NCBI primer-BLAST tool (U.S. National Library of Medicine). Primer synthesis and DNA tube sequencing was performed by Eurofins Genomics. For amplification of DNA sequences needed for cloning, Q5 High-Fidelity DNA Polymerase (NEB) was used according to the following scheme

and PCR was run in a ProFlex PCR system cycler (Applied Biosystems, Thermo Fisher Scientific) under the PCR cycling parameter noted in Table 2.

Q5 PCR pipetting scheme

5x Q5 reaction buffer	10 µl
dNTPs (1 mM)	10 µl
Forward primer (10 µM)	2.5 µl
Reverse primer (10 µM)	2.5 µl
DNA template	1 ng
Q5 Polymerase	0.5 µl
H_2O	*ad* 50 µl

Table 2 Q5 PCR cycling parameter
Annealing temperatures were calculated with the NEB Q5 Tm calculator, as annealing temperatures for Q5 PCR tend to be even higher than the melting temperature (Tm) of the primer pair.

Step	**Number of cycles**	**Temperature**	**Duration**
Initial denaturation	1x	98 °C	30 s
Denaturation	35x	98 °C	10 s
Annealing		50-72 °C	20 s
Elongation		72 °C	30 s/kilobase(kb)
Final elongation	1x	72 °C	2 min
Cool-down	1x	8 °C	10 min

Extraction of DNA from agarose gels was performed using the QIAquick gel extraction kit (Qiagen) according to the manufacturer´s instruction. For purification of DNA sequences from reaction mixtures, the QIAquick PCR purification kit (Qiagen) was applied.

Chemically competent 10-beta *Escherichia coli* (*E.coli*) K12 bacterial cells (NEB) were transformed by heat shock method (Cosloy and Oishi, 1973; Mandel and Higa, 1970). In detail, 3 µl ligation product were added to bacteria that had been thawed on ice and were gently mixed with the pipette tip. After 30 minutes (min) incubation on ice, a heat shock was applied by placing the tube into a 42 °C Thermomixer compact (Eppendorf) for 30-45 s and subsequently placing the cells on ice again for approximately 2 min. Finally, cells were plated directly onto lysogeny broth (LB)

agar plates containing 50 µg/ml ampicillin and were placed either in a 37 °C incubator (Heraeus, Thermo Scientific) overnight or at room temperature (RT) on the benchtop for the period of a weekend.

In Table 3 all plasmids that were cloned during this project are listed.

Table 3 Cloned plasmids
The cloned plasmids are either on a pcDNA3.3, pAAV2.1 or PiggyBac (pb) vector backbone containing β-lactamase coding sequence for ampicillin resistance. The number of base pairs in kilobases (kb) of the sequence encoded between the inverted terminal repeats (ITR) is given (including ITR length).

Plasmid name	**Size between ITR**
pcDNA3.3-CMV-Cas9^{N} (split after E573)	-
pcDNA3.3-CMV-Cas9^{C} (split after E573)	-
pAAV2.1-3x-hU6-sgCnga1-CMV-dCas9^{N}-RmaInt^{N}-SV40polyA	4.77 kb
pAAV2.1-hU6-sgLacZ-CMV-dCas9^{N}-RmaInt^{N}-SV40polyA	4.05 kb
pAAV2.1-3x-hU6-sgOpn1mw-CMV-dCas9^{N}-RmaInt^{N}-SV40polyA	4.77 kb
pAAV2.1-CMV-RmaInt^{N}-dCas9^{C}-VPR-synpA	4.95 kb
pAAV2.1-3x-hU6-sgCnga1-mSWS-dCas9^{N}-RmaInt^{N}-SV40polyA	4.68 kb
pAAV2.1-hU6-sgLacZ-mSWS-dCas9^{N}-RmaInt^{N}-SV40polyA	3.96 kb
pAAV2.1-mSWS-RmaInt^{N}-dCas9^{C}-VPR-synpA	4.85 kb
pAAV2.1-3x-hU6-sgOpn1mw-hRHO-dCas9^{N}-RmaInt^{N}-SV40polyA	4.36 kb
pAAV2.1-hU6-sgLacZ-hRHO-dCas9^{N}-RmaInt^{N}-SV40polyA	3.64 kb
pAAV2.1-hRHO-RmaInt^{N}-dCas9^{C}-VPR-synpA	4.53 kb
pb-TRE-dCas9-VPR-TKpolyA-3x-hU6-sgCnga1-EF1a-TetO-T2A-PuroR-SV40polyA	11.18 kb
pb-TRE-dCas9-VPR-TKpolyA-3x-hU6-sgOpn1mw-EF1a-TetO-T2A-PuroR-SV40polyA	11.18 kb
pb-TRE-dCas9-VPR-TKpolyA-hU6-sgLacZ-EF1a-TetO-T2A-PuroR-SV40polyA	10.43 kb

For plasmid preparations in the mini format a lab internal protocol using alkaline lysis was applied, for midi or maxi plasmid DNA preparations the PureLink™ HiPure Plasmid Midiprep or Maxiprep Kit (Invitrogen, Thermo Fisher Scientific), respectively, was used.

3.2 Single guide RNA design

Genomic DNA and transcript sequences of the human (Dec.2013, GRCh38/hg38) and murine reference genome (Dec.2011, GRCm38/mm10) were downloaded from UCSC Genome Browser using the Table Browser tool, the NCBI gene interface or Ensembl genome browser 93. Additionally, the Eukaryotic Promoter Database was consulted for the identification of promoter regions (Dreos et al., 2015, 2017). Single guide RNA sequences in the genomic target region (Table 4) were chosen using the CRISPOR website with 20 bp-NGG PAM settings for *Sp*Cas9 (Haeussler et al., 2016) and elimination of sgRNAs with a specificity score (Hsu et al., 2013) lower than 50. Off-targets were also predicted by the CRISPOR website and ranked by CFD off-target score from most to least likely (Doench et al., 2016). Top three global off-targets plus top two exonic off-targets per *TET3* sgRNA were chosen for experimental off-target analysis. Sequence information on the primer pairs used for off-target analysis is provided in the appendix (Table 10).

Table 4 sgRNA sequences and positioning
The sequences of used sgRNAs are given with their position relative to the transcription start site (TSS) or within the *TET3* gene, respectively.

Target gene	**sgRNA position**	**sgRNA sequence 5´-3´**
Cnga1 promoter	-104 bp to TSS	TAGGCGACCGGCTTTGAGAA
	-270 bp to TSS	CTGTGGAAGTCTCCAAACGC
	-309 bp to TSS	TCTTCTCTCTCGGCGCTATG
Opn1mw promoter	-60 bp to TSS	GTTTGGGGGCCTTTAAGGTA
	-159 bp to TSS	CCTGAGCCACCCCTGTGGAT
	-260 bp to TSS	TAGCTCTTGCTTGTTTACAA
Opn1sw promoter	-90 bp to TSS	GAACACAATCAATCAGCTTT
	-142 bp to TSS	AGCACCCAGCCTTGTTATCC
	-278 bp to TSS	AGCATCTTTCACAAAGGGCC
E. coli LacZ	Gene body	GTCTGACCGATGATCCGCGC
TET3 up 1	*TET3* Intron 6	CACATGGGTTTACGGAAGTA
TET3 up 2	*TET3* Intron 6	CAACCCAAGTGCCCCTAACT
TET3 down 1	*TET3* Intron 8	CAGGTAGCCTGGTGTAAAAA
TET3 down 2	*TET3* Intron 8	TATCGCTTTACTTCTGGACT

3.3 Cell culture and transfection

All cell culture media and supplements were obtained from Gibco™ (Thermo Fisher Scientific), unless stated otherwise. For all cell culture experiments, materials were

either obtained in cell culture grade or sterilized by autoclaving (Sterilisator, Münchener Medizin Mechanik).

3.3.1 Inducible Neurogenin iPS cells

Inducible Neurogenin iPS cells (iNGN) were a kind gift of the laboratory of Volker Busskamp (DFG-Center for Regenerative Therapies Dresden (CRTD)). The iNGN cells were generated by lentiviral gene delivery integrating a doxycycline-inducible neurogenin expression cassette into the genome of human PGP1 induced pluripotent stem cells (Busskamp et al., 2014).

iNGNs were maintained in feeder-free cultures with StemFlex Medium + 1 % Antibiotic-Antimycotic (Anti-Anti) in a CO_2 incubator (Heraeus, Thermo Scientific) at 37 °C and 5 % CO_2 setting. All used cell culture plates were pre-coated with Geltrex™ LDEV-free, hESC-qualified reduced growth factor basement membrane matrix, which was diluted 1:50 in Dulbecco´s modified eagle medium (DMEM)/F-12 (1:1) + L-Glutamine + 15 mM HEPES + 1 % Anti-Anti. For passaging, cells were washed with Dulbecco´s phosphate-buffered saline (DPBS), detached using TrypLE™ Express, washed again with DPBS and centrifuged for 5 min at 300 x g and RT. Afterwards, cells were seeded in StemFlex + 1 % Anti-Anti + 1 % RevitaCell™ Supplement improving cell survival in the single-cell state. Medium was changed after 18-24 h to StemFlex + 1 % Anti-Anti. Subsequently, medium was changed every other day.

Application of doxycycline (DOX) induces the overexpression of neurogenin 1 and neurogenin 2 and thereby causes a rapid differentiation of the iPSCs into bipolar neurons within 4 days (Busskamp et al., 2014). The first day, on which doxycycline hyclate (Sigma) was added at a concentration of 500 ng/ml in StemFlex + RevitaCell™, was considered day 0 (d0). On d1 of differentiation, medium was changed to a 1:1 mixture of StemFlex + 1 % Anti-Anti and Neurobasal A medium + 1 % Anti-Anti + 1 % GlutaMAX + 2 % B-27 supplemented with DOX. From d2 on, differentiating cells were kept in Neurobasal A medium + 1 % Anti-Anti + 1 % GlutaMAX + 2 % B-27 + DOX and medium was changed every other day. iNGN-derived neurons were harvested either at d4 or d8 of DOX application for further experiments. In contrast, the time point d0 indicates the pluripotent state of iNGN cells that were not exposed to DOX.

Images of iNGN cells at different stages of differentiation were taken with the EVOS® FL cell imaging system (Life Technologies, Thermo Fisher Scientific).

3.3.2 661W cells

661W cells (kindly provided by Dr. Muayyad Al-Ubaidi, University of Houston) were cloned from retinal tumors of a transgenic mouse line that expresses the simian virus 40 (SV40) T antigen under control of the human interphotoreceptor retinol-binding protein promoter (al-Ubaidi et al., 1992). They were shown to demonstrate cellular and biochemical characteristics exhibited by cone photoreceptor cells (Tan et al., 2004). 661W cells were cultured in DMEM medium + GlutaMAX + 1 g/l glucose + pyruvate + 10 % FBS (Biochrom) + 1 % Anti-Anti in a CO_2 incubator (Heraeus, Thermo Scientific) at 37 °C and 10 % CO_2 setting. Transient transfections of 661W cells were performed using Xfect Transfection Reagent (Takara Bio) according to Table 5.

Table 5 Xfect transfection scheme
The amount of DNA as well as volumina of Xfect buffer and Xfect reagent used for transfection in 12-well, 6-well or 10 cm plate format are given. Reagents were pipetted in the indicated order.

	12-well plate format	**6-well plate format**	**10 cm plate format**
DNA	2 µg	5 µg	20 µg
Xfect Buffer	*ad* 50 µl	*ad* 100 µl	*ad* 600 µl
Xfect Reagent	0.6 µl	1.5 µl	6 µl

After a 10 min incubation of the reaction mix at RT, solution was added dropwise onto the culture plate. Culture medium was changed after 4 hours (h) of incubation or the next morning, respectively.

3.3.3 MEF cells

Immortalized mouse embryonic fibroblasts (MEF) were generated as previously described (Jat et al., 1986; Xu, 2005). MEF cells were cultured in DMEM medium + GlutaMAX + 1 g/l glucose, + pyruvate + 10 % FBS (Biochrom) + 1 % penicillin/streptomycin (Biochrom) in a CO_2 incubator (Heraeus, Thermo Scientific) at 37 °C and 5 % CO_2 setting. Transient transfections of MEF cells were performed using Xfect Transfection Reagent (Takara Bio). See Table 5 above for transfection details.

3.3.4 HEK293 cells

Human embryonic kidney 293 (HEK293) cells (DMSZ) were cultured in DMEM medium + GlutaMAX + 1 g/l glucose, + pyruvate + 10 % FBS (Biochrom) + 1 % penicillin/streptomycin (Biochrom) in a CO_2 incubator (Heraeus, Thermo Scientific) at 37 °C and 10 % CO_2 setting. Transient transfections of HEK293 cells were

performed in 10 cm culture plates using the calcium phosphate technique (Kingston et al., 2003). The following reagents were mixed in the indicated order by vortexing:

DNA	10 µg
H_2O	*ad* 450 µl
2.5 M $CaCl_2$	50 µl
2x BBS	500 µl

Reaction mix was shortly incubated at RT and then added dropwise onto the cells. Subsequently, plates were incubated at 37 °C and 5 % CO_2 overnight. On the next day, medium was changed and plates were placed back into a CO_2 incubator at 37 °C and 10 % CO_2 setting until they were harvested.

2 x BBS solution

BES	8.52 g
NaCl	13.08 g
Na_2HPO_4	0.17 g
H_2O	*ad* 800 ml

The pH was adjusted to 7.080 with NaOH and solution was sterile filtrated using a 0.2 µm pore size syringe filter (VWR).

3.4 Generation of stable cell lines using PiggyBac technology

For the generation of stable cell lines, PiggyBac transposon technology was used (System Biosciences). First, PiggyBac transposon vectors containing the FL dCas9-VPR sequence under the control of the TRE, different sgRNAs each expressed from a single hU6 promoter and a bicistronic expression cassette consisting of the reverse tetracycline-controlled transactivator and the puromycin resistance gene (rtTA-T2A-PuroR) (see also Figure 22). To achieve the insertion of the sequence encoded between the ITRs of the transposon vector into random TTAA sites throughout the genome, 661W and MEF cells were co-transfected with a PiggyBac transposase encoding vector and the cloned transposon vector in a 1:2.5 DNA mass ratio. In detail, for the 6-well format, 1.4 µg transposase vector and 3.5 µg transposon vector were mixed for transfection using Xfect transfection reagent (see also Table 5). 48-72 h after transfection, cells were selected for successful integration by addition of puromycin dihydrochloride (Gibco, Thermo Fisher Scientific) for approximately one week at the following concentrations:

661W cells	4.5 µg/ml puromycin
MEF cells	1 µg/ml puromycin

3.5 Generation of TET3KO iNGN

For the generation of the TET3KO iNGN cell line, CRISPR-Cas9 technology was applied. The used protocol was developed and adjusted based on the Alt-R® CRISPR-Cas9 system user guide for the delivery of ribonucleoprotein (RNP) complexes using the Amaxa™ Nucleofector™ System (Integrated DNA Technologies (IDT)).

To erase catalytically active TET3, exon 7 and 8 that partially encode for the first oxygenase domain were deleted using two sgRNAs located upstream of exon 7 and two sgRNAs downstream of exon 8. For information on sgRNA sequences see Table 4.

The ordered RNA oligos (Alt-R® crRNAs and Alt-R® tracrRNA, ATTO™ 550 (IDT)) were resuspended in RNase-free duplex buffer (IDT) to final concentrations of 200 µM. Each crRNA was mixed with tracrRNA in equimolar concentrations in a sterile microcentrifuge tube to a final duplex concentration of 40 µM. For duplex formation, the mixture was heated for 5 min at 95 °C and afterwards cooled to RT. To form the RNP complex, 0.6 µl of each crRNA:tracrRNA duplex and 3.4 µl Alt-R® S.p. Cas9 Nuclease 3NLS (IDT) were diluted with PBS to a total volume of 10 µl and incubated for 20 min at RT.

For nucleofection of the iNGN cells with the assembled RNP complexes, the Amaxa™ Human Stem Cell Nucleofector™ Kit 2 (Lonza) in combination with the Nucleofector™ 2b Device (Lonza) was used. First, iNGN cells were dissociated as described above, washed with PBS and counted using a Neubauer Improved cell counting chamber (BRAND). 8×10^5 cells were pelleted in a falcon tube and resuspended in 80 µl Nucleofection Solution (nucleofection solution 2 and supplement 1 at a 5.5:1 ratio, IDT). The cell suspension was added to the tube with the RNP complex and mixed by pipetting up and down. Moreover, 3 µl of 100 µM Alt-R® Cas9 Electroporation Enhancer (IDT) were added before electroporation. The whole amount of cell/RNP complex mixture was transferred to an electroporation cuvette (Lonza) and air bubbles were removed by tapping the cuvette on the benchtop. Then, the cuvette was placed in the Nucleofector™ device and the program A-033 was selected for electroporation. After electroporation, 500 µl of pre-warmed StemFlex medium containing RevitaCell were added to the cuvette and cells were gently pipetted up and down. The resuspended cells were

transferred to a 24-well plate, which had been coated with 0.5 mg/cm² rhLaminin-521 (Gibco™, Thermo Fisher Scientific) 2 h prior to the electroporation procedure and contained 500 µl of pre-warmed StemFlex medium with RevitaCell™. Cells were cultivated under the conditions described above. Medium was changed after 24 h to StemFlex without RevitaCell™ supplement.

Clonal isolation was performed several days post nucleofection after the cells had reached 60-70 % confluency. For this, single cells of the gene edited iNGN cell pool were replated into separate wells of a 96-well plate by fluorescence activated cell sorting (FACS) at the Immunoanalytics Core Facility of the Helmholtz Zentrum München. In detail, iNGN cells were dissociated according to the standard protocol and dissolved in 1 ml DPBS without calcium and magnesium (Gibco, Thermo Fisher Scientific) supplemented with 2 % BSA (pluriSelect) and 2 mM EDTA (pluriSelect). Cell solution was mixed with propidium iodide as a viability dye and was, directly prior sorting, passed through a cell-strainer using 5 ml Polystyrene round-bottom tubes with a cell-strainer cap with 35µm nylon mesh (Falcon). The FACSAria III (BD Biosciences) distributed single viable (propidium iodide negative) iNGN cells into the inner 60 wells of the prepared 96-well plates under sterile conditions. The 96-well plate had been coated with rhLaminin-521 (Gibco™, Thermo Fisher Scientific) and the inner 60 wells contained 200µl of pre-warmed StemFlex medium with RevitaCell added while the outermost wells were filled with PBS to reduce evaporation from the inner wells. Cells were first gated for size and granularity (FSC vs. SSC) to exclude cell debris, then propidium iodide positive dead cells were excluded from the population and finally doublets were excluded (FSC-A vs. FSC-H) leaving a population of living single cells to be sorted. Medium was changed after 48 h to StemFlex without RevitaCell™ supplement. When the colonies had grown to a passagable size, they were dissociated and transferred to a 24-well plate under the standard conditions. When colonies were expanded, a portion of each potentially edited iNGN cell clone was taken for genomic DNA isolation (3.11) and subsequent genotyping.

For detection of the wanted deletion of exon 7 and 8 from the *TET3* locus, the following primers flanking the deletion site were used in a PCR based genotyping assay. WT allele gives a 4601 bp DNA product while an efficient deletion can be detected by an 852 bp band.

hTET3KO Exon 7+8 fw	CTGCCCTTATCAGAGTTTTTCCCAT
hTET3KO del seq rev	GCCTCACCAGTGGATACCAG

3.6 Immunocytochemistry

For immunocytochemistry, 661W-pb cells were seeded onto sterile 12 mm diameter cover slips in 12-well plates that had been coated with poly-L-lysine hydrobromide (Sigma Aldrich). After 48 h of DOX application, coverslips with cells were removed from the culture plate and placed onto Parafilm® M (Bemis) in a humidity chamber for all following steps. Cover slips were washed three times with DPBS containing calcium and magnesium (Gibco, Thermo Fisher Scientific) and then fixed with 4 % paraformaldehyde pH 7.4 (PFA, Sigma Aldrich) in DPBS for 10 min at RT. After another two washing steps with DPBS, cells were permeabilized for 30 min in 0.3 % Triton X-100 in DPBS at RT. Coverslips were washed three times with DPBS and then incubated with blocking solution (5 % ChemiBLOCKER™ (CB, Merck Millipore) in DPBS) for 30 min at RT. Incubation with primary antibodies (Table 6) in DPBS with 5 % CB was performed overnight at 4 °C. Next, primary antibodies were removed by four washing steps with DPBS. From this point in the protocol, all washing and incubation steps were performed in conditions protecting the samples from light. Secondary antibodies (Table 6) in DPBS with 2 % CB were applied to the cells for 1 h at RT. After three more washing steps with DPBS, cells were incubated for 5 min with a 5 µg/ml Hoechst 33342 solution (Invitrogen, Thermo Fisher Scientific) to stain DNA. Finally, coverslips were washed twice with DPBS and then mounted onto object slides using Permafluor Mountant (Thermo Scientific). Mounted slides were stored in the dark at 4 °C and imaged soon after the staining.

Table 6 Antibodies used in immunocytochemistry

Antibody name	**Source**	**Clone or article number**	**Dilution**
αCnga1 mouse monoclonal	Molday laboratory	PMc1D1	1:30
α*Sp*Cas9 rabbit polyclonal	Diagenode	C15310258	1:1,000
AlexaFluor™ 488 Goat αMouse IgG (H+L)	Invitrogen, Thermo Fisher Scientific	A11001	1:800
Cy™3 AffiniPure Donkey αRabbit IgG (H+L)	Jackson ImmunoResearch	711-165-152	1:400

3.7 Confocal microscopy

For imaging of stained and mounted 661W-pb cells, Leica TCS SP8 spectral confocal laser scanning microscope (Leica Microsystems) equipped with 405 nm diode as well as 488 nm and 552 nm optically pumped semiconductor lasers was used.

Images of 661W-pb cells expressing sgLacZ and sgCnga1 were acquired with a HC PL APO 40x/1.30 Oil CS2 objective (Leica Microsystems) and type F immersion liquid (Leica Microsystems) using LAS X software (Leica Microsystems). All images were taken at identical gain and laser power settings using Z-Stack mode (steps of 0.5 µm size). All images were deconvolved with Huygens Essential version 18.04 (Scientific Volume Imaging, The Netherlands, http://svi.nl) using the CMLE algorithm, with SNR:5 and 40 iterations. Afterwards, single deconvolved scans were merged by maximum projection and further processed using LAS X and Fiji software (Schindelin et al., 2012).

3.8 Patch-clamp measurements

Patch-clamp measurements were performed by René Rötzer (Wahl-Schott group, LMU München). In detail, inside-out patches were excised from 661W-pb cells that were maintained at a DOX concentration of 5 ng/ml. Currents were recorded using an EPC-10 double patch-clamp amplifier (HEKA Elektronik, Harvard Bioscience) and PatchMaster acquisition software (HEKA Elektronik, Harvard Bioscience). Data were digitized at 20 kHz and filtered at 2.9 kHz. All recordings were obtained at RT. The extracellular solution was composed of 140 mM NaCl, 5 mM KCl, 10 mM HEPES and 1 mM EGTA (pH adjusted to 7.4 with NaOH). The intracellular solution contained 140 mM KCl, 5 mM NaCl, 10 mM HEPES and 1 mM EGTA (pH adjusted to 7.4 with KOH). The effect of cGMP was examined by perfusing the patch with extracellular solution supplemented with 300 µM cGMP. To investigate channel blocking, perfusion with a symmetric Ca^{2+}/Mg^{2+} solution composed of 140 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$ and 10 mM HEPES was performed followed by perfusion with a cGMP supplemented Ca^{2+}/Mg^{2+} solution as a control. Currents were evoked from a holding potential of 0 mV by applying a 300 ms pulse of -80 mV followed by a 300 ms pulse of 80 mV every 3 s.

3.9 Protein extraction

3.9.1 Triton X lysates

Whole HEK293 cell protein was isolated using Triton X-100 (TX) lysis buffer. First, HEK293 cells were washed with DPBS and a cell scraper (Sarstedt) was used to detach the cells from the culture plate. Cell suspension was centrifuged for 5 min at 300 x g and 4 °C and supernatant was removed. Subsequently, cell pellet was resuspended in an appropriate volume of TX lysis buffer (150-200 µl for cells cultivated in a 10 cm dish) and transferred to a 2 ml safe-lock tube (Eppendorf). A steel ball was added per tube and tubes were shaken in the mixer mill MM400

(Retsch) at 15 Hz for 30 s. Then tubes containing lysate and steel ball were incubated for 20 min at 4°C with end-over-end rotation (Tube Rotator, VWR). After removing the steel ball with a magnet, tubes were centrifuged for 10 min at 5,000 x g and 4 °C. Protein containing supernatant was transferred into a new reaction tube and stored at - 20 °C until further use.

TX lysis buffer

Triton X-100	0.5 % (v/v)
NaCl	5 M
$CaCl_2$	2.5 M

+ 1 cOmplete ULTRA Protease Inhibitor Cocktail tablet (Roche) per 10 ml buffer

3.9.2 Bradford assay

Bradford assay (Bradford, 1976) was applied to measure protein concentrations in solution. Protein Assay Dye Reagent Concentrate (Bio-Rad) was diluted 1:5 with water. 1 µl of protein lysate was mixed with 49 µl water in a PMMA cuvette (BRAND). Subsequently, 950 µl of diluted Protein Assay Dye Reagent were added and the solution was mixed. After a 2 min incubation at RT, absorption was measured and compared to blank solution containing only resuspension buffer using the BioPhotometer (Eppendorf).

3.10 Western blotting

For Cas9 WB, 30 µg whole cell protein were mixed with 6xLaemmli + DTT buffer an incubated at 72 °C for 10 min. For SDS-PAGE, a 6-12 % polyacrylamide gradient gel for separation topped by a stacking gel was prepared using the Protean 3 gel system (Bio-Rad). Samples were loaded into the pockets of the stacking gel and PageRuler Prestained Protein Ladder (Thermo Fisher Scientific) was used to determine protein size. Electrophoresis was run at 140 V for approximately 70 min.

4x Tris-HCl/SDS pH 6.8 buffer

Tris-HCl	0.5 M
SDS	0.4 %

adjust pH to 6.8

4x Tris-HCl/SDS pH 8.8 buffer	
Tris-HCl	1.5 M
SDS	0.4 %
adjust pH to 8.8	

Stacking gel (for 2 gels)	
30 % acrylamide/bis-acrylamide	1 ml
4x Tris-HCl/SDS pH 6.8 buffer	1.9 ml
H_2O	4.6 ml
APS	37.5 µl
TEMED	7.5 µl

Gradient separation gel (for 2 gels)

	6 %	12 %
30 % acrylamide/bis-acrylamide	2.3 ml	4.6 ml
4x Tris-HCl/SDS pH 8.8 buffer	2.8 ml	2.8 ml
H_2O	6.2 ml	3.9 ml
APS	22.5 µl	22.5 µl
TEMED	7.5 µl	7.5 µl

A 10 ml pipette was used to aspire 4.25 ml of the 6 % gel solution followed by 4.25 ml of the 12 % gel solution. The solutions were gently mixed by aspiring one air bubble before pouring the gradient gel solution into the gel system.

10x electrophoresis buffer	
Tris	30 g
Glycine	144 g
SDS	10 g
H_2O	*ad* 1 l

Transfer of the proteins onto the PVDF membrane (Peqlab) was performed in the Mini Trans-Blot cell (Bio-Rad) filled with blotting buffer at 100 V for 100 min. After blotting, the membrane was transferred into TBS-T with 5 % non-fat dried milk powder (AppliChem) and blocked for 1 h at RT. Subsequently, the membrane was incubated in primary antibody solution consisting of 5 ml TBS-T with 1 % milk powder and 1:1,000 rabbit polyclonal αSpCas9 antibody (C15310258, Diagenode) in a 50 ml falcon tube overnight at 4 °C. The following day, the membrane was washed three times for 5 min in TBS-T applying gentle agitation before it was

incubated with ECL™ donkey αRabbit IgG HRP-linked secondary antibody (NA934V, GE Healthcare) at a 1:2,000 dilution in TBS-T with 1 % milk powder for 1 h at RT. Afterwards, the membrane was washed three times with TBS-T, followed by one washing step with H_2O, before signal detection was performed. For detection, the membrane was incubated with Western Blotting Luminol Reagent (Santa Cruz) prepared according to the manufacturer´s manual. The chemiluminescence signal was detected using the Chemidoc MP Imaging system (Bio-Rad) including Image Lab software (Bio-Rad), which was also utilized for quantification of band intensities.

10x blotting buffer	
Tris	30 g
Glycine	144 g
H_2O	*ad* 1 l

10x TBS	
Tris	12.1 g
NaCl	80.2 g
H_2O	*ad* 1 l

TBS-T	
10x TBS	100 ml
Tween-20	1 ml
H_2O	*ad* 1 l

3.11 Genomic DNA extraction

For genomic DNA (gDNA) extraction, cells were washed with DPBS and then harvested in the culture plate by adding RLT buffer (Qiagen) + 1 % β-mercaptoethanol (Sigma-Aldrich). Plates were placed on an orbital shaker for 2-5 min and then cell suspension was transferred to a 2 ml safe-lock tube (Eppendorf). A steel ball was added per tube and tubes were shaken in the mixer mill MM400 (Retsch) at 30 Hz for 1 min. Afterwards, the steel ball was removed with a magnet and the tubes were centrifuged for 5 min at 21,000 x g and RT. Subsequently, up to 800 µl supernatant at once were transferred to a Zymo-Spin IIC-XL column (Zymo Research) that was placed in 2 ml collection tubes. Columns were centrifuged for 4 min at 1500 x g and RT, followed by a 1 min centrifugation at 10,000 x g and RT. Flow-through contains RNA as well as protein and was frozen at –80 °C for further processing. Next, 400 µl RNase wash buffer (Genomic lysis buffer (Zymo Research) + 0.2 mg/ml RNase A (Qiagen)) were added to the column and incubated for 10-

15 min at RT. The columns were centrifuged for 2 min at 10,000 x g and RT and flow-through was discarded. Subsequently, column membrane was washed by addition of 400-600 µl DNA Pre-Wash buffer (Zymo Research) and centrifugation at 10,000 x g for 1 min. Next, a similar washing step was performed with 600-800 µl gDNA wash buffer (Zymo Research). The third washing step was performed with another 600 µl of gDNA wash buffer, followed by a 2 min centrifugation at 10,000 x g and RT. To eliminate traces of the wash buffer, columns were transferred into a new 2 ml collection tube and centrifuged for 1 min at 10,000 x g. Finally, columns were transferred into new 1.5 ml reaction tubes. For gDNA elution, the column membrane was incubated with 50-100 µl H_2O for 10-15 min at RT and then centrifuged for 2 min at 10,000 x g and RT. For determination of gDNA concentration, 1.5 µl of DNA solution was measured at a NanoDrop 2000c spectrophotometer (Thermo Scientific). gDNA was stored at - 20 °C.

When genomic DNA was isolated for quantification of cytosine modifications, 400 µM 2,6-ditert-butyl-4-methylphenol (BHT) was added to RLT buffer with β-mercaptoethanol and RNase wash buffer. H_2O for elution contained 20 µM BHT.

3.12 Quantification of cytosine modifications by mass spectrometry

Per biological replicate, 1 x 10^6 iNGN cells were plated in a 6 cm dish, incubated without DOX (d0) or with DOX for 4 or 8 days and then harvested for gDNA extraction. gDNA was handed over to Franziska R. Traube (Carell group, LMU München) for further sample preparation and mass spectrometry. With ultra-high-performance liquid chromatography tandem mass spectrometry (UHPLC-MS/MS), three technical replicates were analyzed per biological replicate. For sample preparation, 3 µg of gDNA in 35 µL H_2O were digested per technical replicate. In detail, 7.5 µl of 480 µM $ZnSO_4$ aqueous solution containing 18.4 U nuclease S1 (*Aspergillus oryzae*, Sigma-Aldrich), 5 U Antarctic phosphatase (NEB) and labeled internal standards ([15N2]-5cadC 0.04301 pmol, [15N2,D2]-5hmdC 7.7 pmol, [D3]-5mdC 51.0 pmol, [15N5]-8-oxo-dG 0.109 pmol, [15N2]-5fdC 0.04557 pmol) were added. Then, the mixture was incubated at 37 °C for 3 h. Subsequently, 7.5 µL of a 520 µM [Na]2-EDTA solution containing 0.2 U snake venom phosphodiesterase I (*Crotalus adamanteus*, USB corporation) were added, then the sample was incubated again for 3 h at 37 °C and finally stored at −20 °C. Prior to UHPLC-MS/MS analysis, samples were filtered using an AcroPrep™ Advance 96-well filter plate (0.2 µm Supor membrane, Pall Life Sciences).

For quantitative UHPLC-MS/MS analysis of digested DNA samples an Agilent 1290 UHPLC system equipped with a UV detector and an Agilent 6490 triple quadrupole

mass spectrometer was used. Natural nucleosides were quantified with the stable isotope dilution technique (Traube et al., 2019). An improved method based on (Pfaffeneder et al., 2014) enabled the concurrent analysis of all nucleosides in one single analytical run. The following source-dependent parameters were applied: gas temperature 80 °C, gas flow 15 l/min (N2), nebulizer 30 psi, sheath gas heater 275 °C, sheath gas flow 15 l/min (N2), capillary voltage 2,500 V in the positive ion mode, capillary voltage -2,250 V in the negative ion mode and nozzle voltage 500 V. The fragmentor voltage was 380 V/250 V and for the positive mode Delta EMV was set to 500 V. Chromatography was performed by a Poroshell 120 SB-C8 column (Agilent, 2.7 µm, 2.1 mm × 150 mm) at 35 °C using a gradient of water and MeCN (each containing 0.0085 % (v/v) formic acid) at a flow rate of 0.35 ml/min: 0 → 4 min, 0 → 3.5 % (v/v) MeCN \ 4 → 6.9 min, 3.5 → 5 % MeCN \ 6.9 → 7.2 min, 5 → 80 % MeCN \ 7.2 → 10.5 min, 80 % MeCN \ 10.5 → 11.3 min, 80 → 0 % MeCN \ 11.3 → 14 min, 0 % MeCN. The effluent up to 1.5 min and after 9 min was diverted to waste by a Valco valve. The autosampler was cooled to 4 °C. An injection volume of 39 µL was used.

3.13 Infinium® MethylationEPIC BeadChip

For large-scale quantitative methylation measurements at the single CpG site level, a methylation beadchip analysis was performed for WT and TET3KO iNGN-derived neurons. For this, WT and TET3KO iNGNs were seeded into three wells of a 6-well plate (3n) and incubated with DOX for 8 days. Then, genomic DNA was isolated and 2 n per genotype were submitted to IMGM Laboratories GmbH. There, the Infinium® MethylationEPIC ´850K´ BeadChip Array (Illumina) was used in combination with bisulfite conversion of the DNA samples and the Infinium® HD Methylation Assay (Illumina). Fluorescent signal intensities on the BeadChip were detected with a dual color channel approach and the Illumina iScan System.

Quality control, data preprocessing and differential DNA methylation analysis were performed using the RnBeads package (Assenov et al., 2014) in R (R Development Core Team, 2008). For data preprocessing the wm.dansen method was used. To identify differentially methylated single CpG sites, p-values were computed using the limma method (Smyth, 2004) Hierarchical linear models from the limma package were employed and fitted using an empirical Bayes approach on derived M-values (Du et al., 2010). All methylation sites with a mean methylation difference ≥ 20 % (i.e. differential methylation ≤ -0.2 or ≥ 0.2) and an adjusted (FDR-corrected) p-value ≤ 0.05 were considered to be differentially methylated.

For further analysis, interpretation and integration of data derived from the methylation chip, Ingenuity® Pathway Analysis (Qiagen Bioinformatics) was performed. In detail, the hyper- or hypomethylated sets of genes were uploaded and separate core analyses were performed with MethylationEPIC BeadChip background settings.

3.14 Bisulfite conversion and subsequent PCR amplification

For validation of differentially methylated CpGs identified by the MethylationEPIC BeadChip, pyrosequencing of selected CpGs in the promoter region of *FAM122C* and *PCDHA6* was performed. For this, 500 ng of gDNA from d8 iNGN-derived WT and TET3KO neurons were bisulfite-converted and purified using the EpiTect® Bisulfite Kit (Qiagen) according to the manufacturer´s manual. Bisulfite-converted DNA was eluted twice from the EpiTect spin column using 20 µl water each. Next, the target regions were amplified from the bisulfite-converted DNA using PfuTurbo Cx Hotstart DNA Polymerase (Agilent Genomics) which does not show uracil stalling on bisulfite-treated DNA.

PfuTurbo Cx PCR pipetting scheme

10x PfuTurbo Cx reaction buffer	5 µl
dNTPs (1 mM)	10 µl
Forward primer (10 µM)	2 µl
Reverse primer (10 µM)	2 µl
Bisulfite-converted gDNA	2 µl
PfuTurbo Cx Hotstart Polymerase (2.5 U/µl)	1 µl
H_2O	*ad* 50 µl

Table 7 PfuTurbo Cx PCR cycling parameter

Step	**Number of cycles**	**Temperature**	**Duration**
Initial denaturation	1x	95 °C	2 min
Denaturation	35-50 x	95 °C	30 s
Annealing		Primer Tm - 5 °C	30 s
Elongation		72 °C	1 min/kb
Final elongation	1x	72 °C	10 min
Cool-down	1x	8 °C	10 min

PyroMark Assay Design Software version 2.0.1.15 (Qiagen) was used for pyrosequencing assay design resulting in the primer sets for PCR amplification and pyrosequencing listed in the appendix (Table 11).

3.15 Pyrosequencing

Pyrosequencing was performed using the PyroMark® Q24 System from Qiagen.

First, PyroMark® Gold Q24 Reagents (Qiagen) were loaded into a PyroMark Q24 Cartridge (Qiagen) according to Pre-run information and cartridge was inserted into the PyroMark® Q24 Instrument (Qiagen).

Sequencing sample preparation included binding of Streptavidin Sepharose® High Performance Beads (GE Healthcare) to the biotin-labeled PCR product. For this step, the following reaction was prepared:

H_2O	21 µl
Binding Buffer (Qiagen)	40 µl
PCR product	17 µl
Streptavidin Sepharose HP beads	2 µl

Mixtures were incubated for at least 5 min at RT and constant agitation at 1400 rpm to prevent a deposition of the Sepharose beads. Moreover, a PyroMark® Q24 plate (Qiagen) was prepared by addition of 0.3 µM sequencing primer in 25 µl Annealing Buffer (Qiagen) per well. Then, streptavidin sepharose bead coupled PCR samples were processed using PyroMark™ Q24 Vacuum Prep Workstation (Qiagen). Initially, beads were captured with the vacuum tool, followed by a 5 s flush of 70 % ethanol, a 5 s flush of Denaturation Solution (Qiagen) and a 10 s flush of Washing Buffer (Qiagen). Afterwards, beads were released into the reaction plate containing the sequencing primers. To anneal the sequencing sample to the sequencing primers, the reaction plate was placed on an 80 °C heatblock for 2 min and subsequently cooled to RT again for at least 5 min. Finally, the plate was processed in the PyroMark® Q24 Instrument (Qiagen). Pyrosequencing run and analysis were performed using PyroMark™ Q24 Software (Pyrosequencing AB, Biotage AB).

3.16 RNA extraction

Cells were washed with DPBS and then harvested in the culture plate by adding RLT Plus buffer (Qiagen) + 1 % β-mercaptoethanol (Sigma-Aldrich). Plates were placed on an orbital shaker for 2-5 min and cell suspension was transferred to a 2 ml safe-

lock tube (Eppendorf). For cell dissociation, a steel ball was added per tube and tubes were shaken in the mixer mill MM400 (Retsch) at 30 Hz for 1 min. Afterwards, the steel ball was removed with a magnet and the tubes were centrifuged for 5 min at 21,000 x g and RT. For RNA extraction from the homogenate, RNeasy Plus Kit (Qiagen) was applied according to the manufacturer´s manual. First, gDNA was removed from the sample by applying the mixture to gDNA eliminator spin columns (Qiagen), followed by 30 s centrifugation. Then, RNA containing flow-through was mixed with an equal volume of 70 % ethanol and applied to the RNA binding RNeasy spin column (Qiagen). For extraction of the RNA used for RNA-seq, an additional on-column DNA digest using RNase-free DNase set (Qiagen) was performed according to the manufacturer´s manual. After several washing steps according to the manual, the column was dried by centrifugation and RNA was eluted using 30 µl RNase-free H_2O (Qiagen). RNA extraction was performed at RT but as soon as RNA was eluted, the RNA solution was constantly kept on ice. For determination of RNA concentration, 1.5 µl of RNA solution was measured at a NanoDrop 2000c spectrophotometer (Thermo Scientific). RNA was stored at -80 °C.

3.17 cDNA synthesis

For cDNA synthesis, RevertAid First Strand cDNA synthesis kit (Thermo Scientific) was used according to the manufacturer´s manual.

cDNA synthesis pipetting scheme

Total RNA	up to 1,000 ng
Random hexamer primer	1 µl
Oligo $(dT)_{18}$ primer	1 µl
Nuclease-free H_2O	*ad* 12 µl

RNA was incubated with primers at 65 °C for 5 min and then cooled down to 10 °C. After spinning solution down, following components were added:

5x reaction buffer	4 µl
RiboLock RNase Inhibitor (20 U/µl)	1 µl
10 mM dNTP mix	2 µl
RevertAid M-MuLV RT (200 U/µl)	1 µl
Total volume	*20 µl*

cDNA synthesis cycling parameter

25 °C	5 min
42 °C	60 min
70 °C	5 min
10 °C	∞

cDNA was kept either at - 20 °C for short-term storage (one-two weeks) or at - 80 °C for longer storage. For application in qRT-PCR, cDNA was diluted 1:5 with 80 µl nuclease-free H_2O.

3.18 Reverse Transcription PCR

Reverse Transcription PCR (RT-PCR) was performed using Q5 High-Fidelity DNA polymerase (NEB) according to the protocol noted in 3.1. The following PCR primer pairs were used for amplification of 2 µl cDNA per reaction.

Cnga1 fw	CTCCAGTCTGATTACCTAGAATAC
Cnga1 rev	CCACCCAAACTTGATGTACAGC
Opn1mw fw	GGAGCAGGTACTGGCCTTATG
Opn1mw rev	GGAGGTAGCAGAGCACGATG

Moreover, *Gapdh*-specific control PCR primers from the RevertAid First Strand cDNA synthesis kit (Thermo Scientific) were used in a loading control reaction. RT-PCR products were separated on a 2.5 % agarose gel and subsequently imaged at the Chemidoc MP Imaging system (Bio-Rad).

3.19 Quantitative Reverse Transcription PCR

Real-Time quantitative Reverse Transcription PCR (qRT-PCR) was performed using StepOnePlus Real-Time PCR system (Applied Biosystems, Thermo Fisher Scientific).

PowerUp SYBR Green Master Mix (Applied Biosystems, Thermo Fisher Scientific) was used for quantification of amplified PCR product. Reaction mixtures were pipetted according to the following scheme in two-three technical replicates into MicroAmp Fast 96-well reaction plates (Applied Biosystems, Thermo Fisher Scientific).

qRT-PCR pipetting scheme

Forward primer (10 µM)	0.6 µl (0.3 µM)
Reverse primer (10 µM)	0.6 µl (0.3 µM)
2x PowerUp MM	10 µl
cDNA (diluted 1:5)	5 µl
H_2O	*ad* 20 µl

Table 8 Fast mode Real-time PCR parameter
The asterisks * mark the points of data collection.

Step	**Number of cycles**	**Temperature**	**Duration**
Uracil-DNA-Glycosylase activation	1x	50 °C	2 min
Hot-Start DNA Polymerase activation	1x	95 °C	2 min
Denaturation	40x	95 °C	3 s
Annealing		58 °C	5 s
Elongation		60 °C*	25 s
Melt curve acquisition	1x	95 °C	15 s
	1x	60 °C	1 min
	1x	95 °C Ramp + 0.3 °C/s*	15 s

The used housekeeping and target primer pairs are listed in the appendix (Table 12).

Analysis of acquired data was performed with StepOnePlus Real-Time PCR system software (Applied Biosystems, Thermo Fisher Scientific). Manual adjustments were made to baseline settings and threshold position if necessary. A table with threshold cycle (Ct) and mean Ct value per technical replicates was exported to Excel (Microsoft). Delta delta Ct calculation method (Livak and Schmittgen, 2001; Pfaffl, 2001) was applied in order to evaluate fold changes in mRNA expression between tested conditions using the following equations:

$$\Delta Ct = Ct_{\text{target}} - Ct_{\text{housekeeping}}$$

$$\Delta\Delta Ct = \Delta Ct - \text{Mean}(\Delta Cts \text{ of all replicates of control sample})$$

$$\text{Relative expression to control sample} = \text{primer efficiency}^{-\Delta\Delta Ct}$$

A primer efficiency value of 2 was used in the equation, unless a lower one could be determined in primer efficiency analysis using a cDNA dilution series.

$$\text{Relative expression to control sample} = 2^{-\Delta\Delta Ct}$$

When no SBYR Green signal was detectable in control samples (for example when *Cnga1* specific primers were used in 661W cells expressing dCas9-VPR and sgLacZ), a Ct value of 40 was used to enable a calculation of fold changes.

3.20 RNA-sequencing

RNA-sequencing (RNA-seq) using RNA from d8 iNGN-derived neurons was conducted by Dr. Gilles Gasparoni (Walter group, Universität des Saarlandes, Saarbrücken). For this, poly-A enriched cDNA libraries were prepared and sequenced using the HiSeq 2500 system (Illumina®). Subsequent RNA-seq data analysis was done by Dr. Karl Nordström (Walter group, Universität des Saarlandes, Saarbrücken). Differential analysis of read counts was performed using the DESeq2 R package (Love et al., 2014). Transcripts with a log2FoldChange value < -1 and > 1 as well as an adjusted p-value < 0.05 were considered to be differentially expressed. Gene ontology analysis was performed on two unranked lists of genes (target vs. background) using the Gene Ontology enRIchment anaLysis and visuaLizAtion tool (GOrilla) (Eden et al., 2007, 2009). Biological process was selected as ontology.

3.21 Statistics

To statistically analyze and graph scientific data GraphPad Prism version 7.04 Software was used. The applied statistical tests are indicated in the figure legends.

4 Results

4.1 TET3KO iNGN cell line

4.1.1 TET levels in iNGN cells and iNGN-derived neurons

The iNGN cell line represents a very valuable *in vitro* model to investigate epigenetic and transcriptional changes underlying neuronal differentiation. Within only four days these genetically modified human iPS cells differentiate into bipolar neurons (Figure 8B). As the focus of this work was mainly on deciphering the role of TET enzymes in this process, the dynamics of TET family member expression were analyzed in undifferentiated iNGN cells (DOX d0) and iNGN-derived d4 and d8 neurons (Figure 8A).

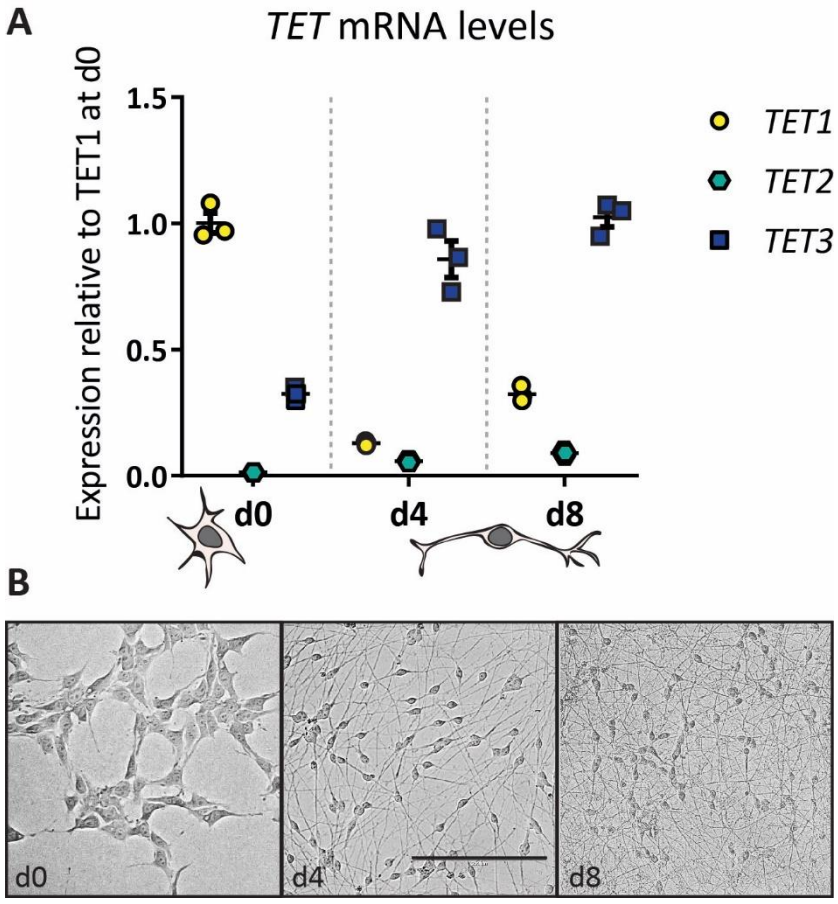


Figure 8 TET expression levels in iNGN cells and derived neurons
(A) Relative *TET1, TET2 and TET3* transcript levels in WT iNGNs without DOX application (d0) or after 4 (d4) and 8 days (8d) of DOX-induced differentiation were analyzed by qRT-PCR. *TET* expression is normalized to *TET1* mRNA levels in d0 iNGNs. ***(B)*** Representative brightfield images of d0 WT iNGNs as well as d4 and d8 iNGN-derived neurons. Scale bar 200 µm.

TET1 was shown to be the dominant TET enzyme in the pluripotent iNGN cells, which is in line with literature (Ito et al., 2010; Koh et al., 2011). In the pluripotent state *TET3* showed much lower expression than *TET1*. However, with neuronal differentiation the roles shift and *TET3* transcript levels increase. In both d4 and d8 iNGN-derived neurons TET3 is the major TET enzyme. Notably, *TET2* is only expressed at low levels before (d0) and after (d4, d8) neuronal differentiation. Based on literature and these experimental findings, it was decided to knock-out (KO) TET3 in the iNGN cell line to investigate the effects on neuronal differentiation and function.

4.1.2 Generation of TET3KO iNGN cell line

For TET3KO generation, exon 7 and 8 of the *TET3* gene that encode for the beginning of the first oxygenase domain were deleted using CRISPR-Cas9 technology (Figure 9A). Splicing of exon 6 to any of the downstream exons would result in a frameshift and a premature stop codon. Hence, with a successful deletion of those two exons, no catalytically active TET3 can be produced in the TET3KO iNGNs.

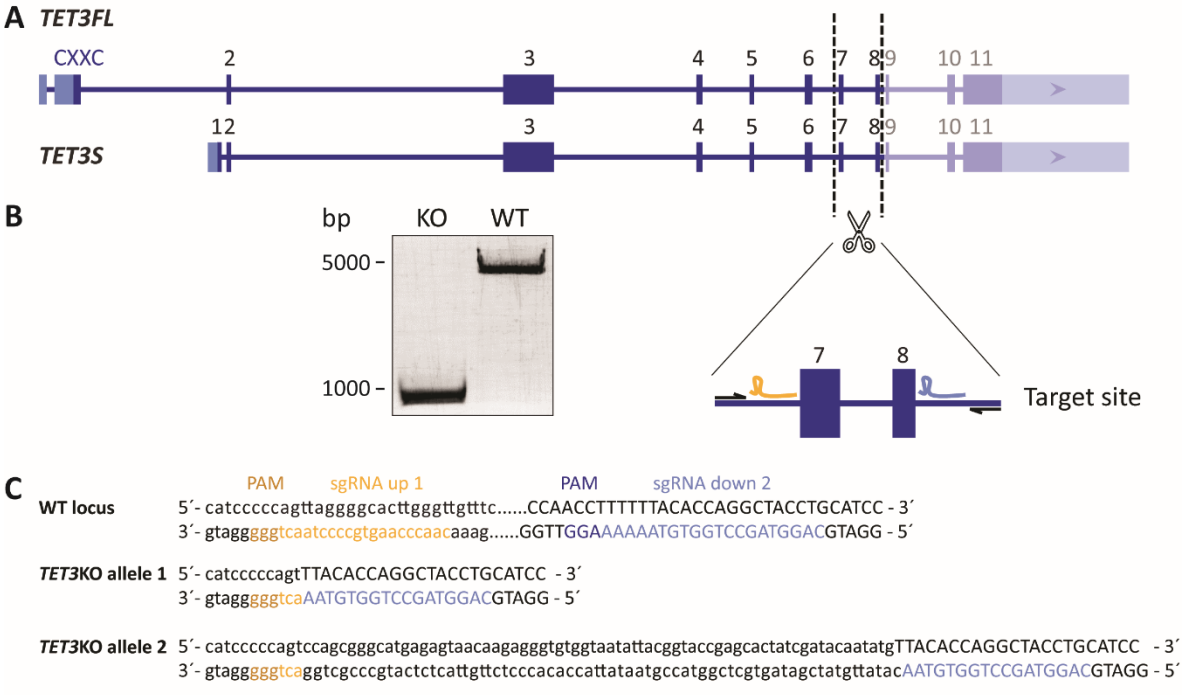


Figure 9 Generation of TET3KO iNGN cell line
(A) For the generation of an iNGN cell line lacking functional TET3, *TET3* exons 7 and 8 partially encoding for the first oxygenase domain, were deleted using CRISPR-Cas9 technology. To achieve this, single guide RNAs (sgRNAs) were designed to bind upstream of exon 7 and downstream of exon 8. Short and full-length *TET3* are equally affected by this deletion that shifts the open reading frame of the downstream coding sequence. ***(B)*** A PCR assay was used to amplify the target region and detect the deletion in the *TET3* locus. WT alleles result in a 4601 bp band while the TET3KO allele yields an ~852 bp product. The scheme on the right-hand side

depicts the position of the sgRNAs in yellow and blue as well as the primers used in the PCR assay as black arrows. **(C)** Sanger sequencing of the PCR bands confirmed the deletion of exon 7 and 8 in the TET3KO iNGN cell line resulting in slightly different sequences at the two TET3 alleles. The target sequences of upstream sgRNA 1 and downstream sgRNA 2 are depicted in yellow and light blue, respectively. The protospacer adjacent motif (PAM) next to each sgRNA is colored in darker shades.

For the deletion, WT iNGN cells were nucleofected with a mixture of ribonuleoprotein complexes consisting of recombinant *Sp*Cas9 protein and one of four sgRNAs (two upstream and two downstream of the target exons). After screening of 21 edited clones, one clone was identified with a successful deletion of exon 7 and 8 on both alleles and used for further experiments. The designed genotyping assay utilizing a primer pair flanking the deletion site yields a small PCR product in case of the deletion and a longer product when exon 7 and 8 are amplified (Figure 9B). Sanger sequencing of the PCR products was performed to exactly understand the edited genomic sequence. On both alleles, a desired deletion within the DNA region complementary to the two outermost sgRNAs occurred. On one allele, the DSB was repaired seamlessly, while on the other allele, NHEJ resulted in a small insertion of 68 bp in the intron between exon 6 and 9 (Figure 9C).

4.1.3 Characterization of TET3KO iNGN cells

To validate the TET3KO iNGN cell line, qRT-PCR analyses using *TET3* isoform specific primer sets were performed at d0, d4 and d8 of differentiation (Figure 10). Primer pairs were used that amplify either the exon encoding the CXXC domain or the *TET3S* unique exon 1 to specifically detect *TET3FL* or *TET3S* transcript, respectively. While in undifferentiated TET3KO iNGNs (d0) *TET3S* transcript levels are only slightly decreased, *TET3FL* is significantly reduced in comparison to WT iNGNs (Figure 10A, B). In d4 and d8 differentiated TET3KO neurons both, *TET3FL* and *TET3S* transcript levels, are significantly diminished to around 50 % of WT levels. Using a primer pair that binds to the 3´region of *TET3* and thus equally amplifies *TET3FL* and *TET3S* transcript (total *TET3*) yields comparable results. Total *TET3* mRNA levels are also reduced in undifferentiated TET3KO iNGNs (Figure 10C). Upon differentiation, TET3KO neurons show a highly significant decrease in total *TET3* transcript levels compared to WT neurons.

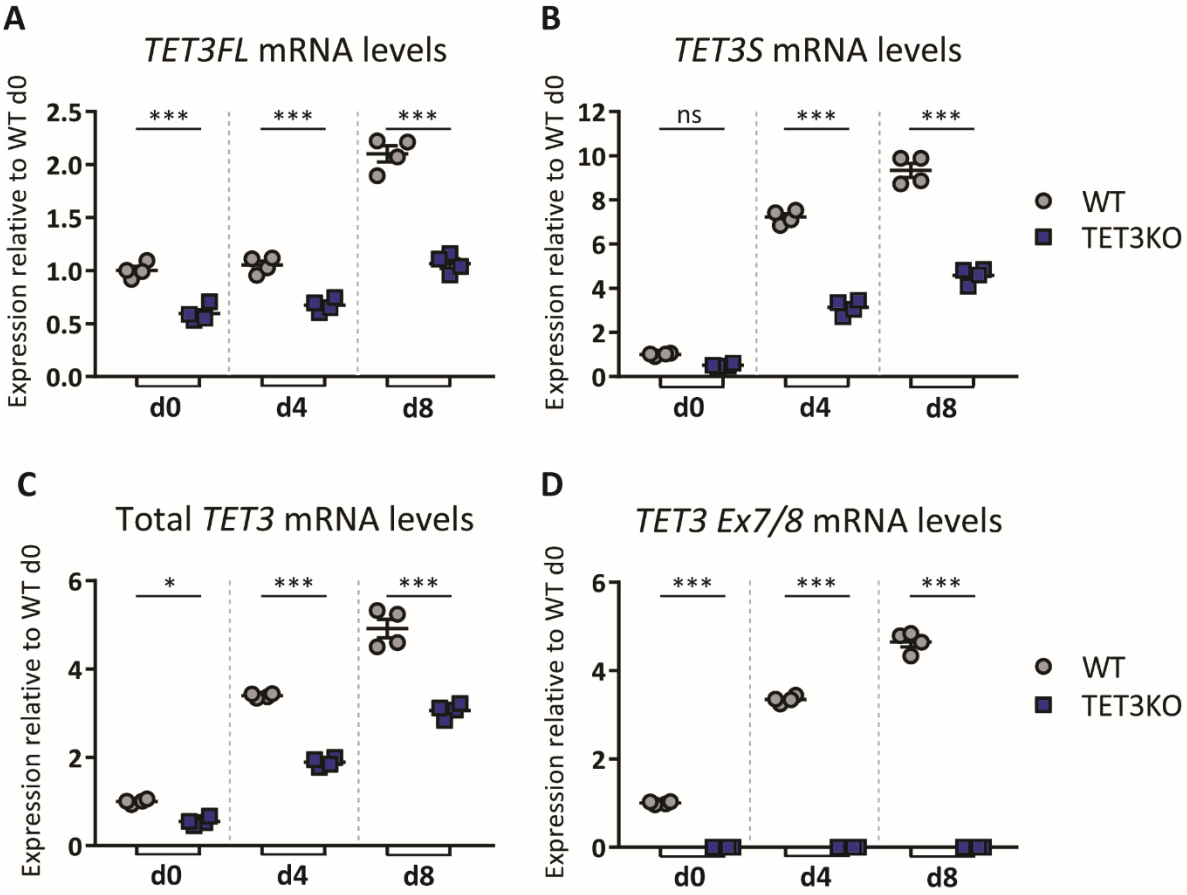


Figure 10 Relative mRNA expression levels of TET3 isoforms in WT and TET3KO iNGN cells

Relative ***(A)*** *TET3* full-length (FL), ***(B)*** *TET3* short (S), ***(C)*** total *TET3* and ***(D)*** *TET3 Ex7/8* transcript levels in WT and TET3KO iNGNs without DOX application (d0) or after 4 (d4) and 8 days (8d) of DOX-induced differentiation were analyzed by qRT-PCR. The TET3 Ex7/8 primer pair binds to the sequence encoded by exons 7 and 8, which are deleted in the TET3KO cell line. Primers used for detection of total TET3 transcript bind to exon 10 and 11 coding sequence. Expression is normalized to the respective mRNA levels in WT iNGN at d0. ns $p > 0.05$, * $p \leq 0.05$, *** $p \leq 0.001$ (Ordinary one-way ANOVA with Sidak´s multiple comparisons test), n=4.

In summary, *TET3* transcript levels are significantly reduced in the TET3KO iNGNs suggesting that the applied CRISPR-Cas9 targeting scheme results in unstable mutated *TET3* mRNA, which is prone to nonsense-mediated mRNA decay. As an extra control for successful deletion of *TET3* exon 7 and 8, cDNA was amplified using a set of primers that bind in these exons. No transcript at all matching this sequence could be detected in TET3KO iNGNs (Figure 10D) supporting the assumption that exon 7 and 8 were erased. Subsequently, it was tested if this reduction in *TET3* expression was somehow compensated by upregulating other TET members in TET3KO iNGNs. Hence, qRT-PCR analysis using *TET1* and *TET2* specific primers was performed.

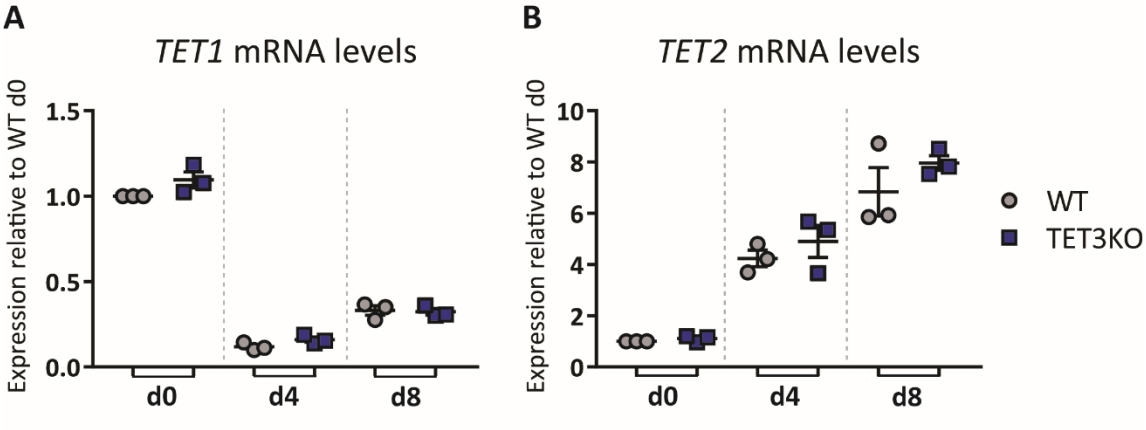


Figure 11 Relative TET1 and TET2 mRNA expression levels in WT and TET3KO iNGN cells
Relative ***(A)*** *TET1* and ***(B)*** *TET2* transcript levels in WT and TET3KO iNGNs without DOX application (d0) or after 4 (d4) and 8 days (8d) of DOX-induced differentiation were analyzed by qRT-PCR. Expression is normalized to the respective mRNA levels in WT iNGN at d0. No significant differences could be detected in TET1 and TET2 transcript expression levels in WT and TET3KO iNGN cells ($p > 0.05$, Ordinary one-way ANOVA with Sidak´s multiple comparisons test, n=3).

Importantly, no significant changes in *TET1* and *TET2* expression could be observed in TET3KO iNGNs when compared to WT iNGNs. In d0 TET3KO iNGN cells as well as d4 and d8 neurons, *TET1* expression was unchanged when compared to WT cells at the respective time point (Figure 11A). Moreover, *TET2* expression was unaffected by the *TET3* decrease in undifferentiated iNGNs as well as d4 and d8 differentiated TET3KO neurons (Figure 11B).

Next, it was analyzed whether the depletion of functional *TET3* also causes aberrations in 5mC removal or 5hmC generation. First, global 5mC and 5hmC levels were quantified in WT and TET3KO iNGNs before and after differentiation using UHPLC-MS/MS. 5mC levels were found to be increased only in undifferentiated (d0) TET3KO iNGNs compared to WT iNGNs (Figure 12A). Moreover, 5hmC levels were significantly lower in TET3KO iNGN-derived d8 neurons than in WT neurons (Figure 12B). With regards to differentiation driven changes in 5hmC, WT d4 and d8 neurons show a significant increase in 5hmC compared to undifferentiated WT d0 iNGNs (5hmC/dN WT d0 vs. d8 $p < 0.0001$, Ordinary one-way ANOVA with Tukey´s multiple comparisons test). Most strikingly, TET3KO neurons (d4 and d8) do not contain more 5hmC than TET3KO d0 iNGNs (5hmC/dN TET3KO d0 vs. d8 $p = 0.4239$, Ordinary one-way ANOVA with Tukey´s multiple comparisons test). From this result, it could be concluded that a substantial proportion of the 5hmC accumulation during neuronal differentiation is indeed mediated by TET3.

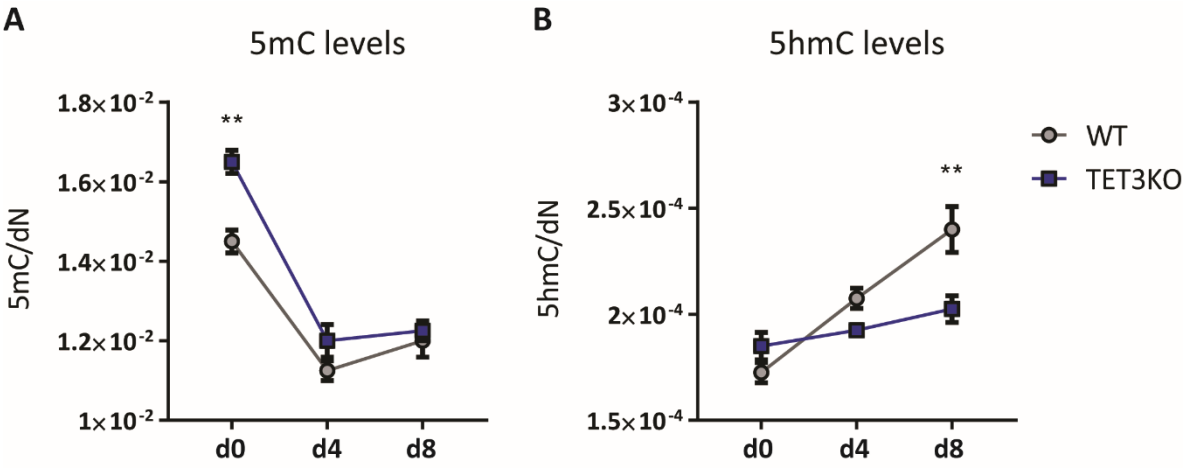


Figure 12 Quantification of global 5mC and 5hmC levels in WT and TET3KO iNGN cells

Global ***(A)*** 5mC and ***(B)*** 5hmC levels per deoxynucleotides (dN) were quantified by UHPLC-MS/MS in the genomic DNA of WT and TET3KO iNGNs without DOX application (d0) or after 4 (d4) and 8 days (d8) of DOX-induced differentiation. ** $p \leq 0.01$ (Ordinary one-way ANOVA with Sidak´s multiple comparisons test), n=4.

4.1.4 Analysis of genome-wide differential methylation

Besides global changes in DNA methylation and hydroxymethylation, site-specific genome-wide changes in 5mC between WT and TET3KO neurons were analyzed using bisulfite conversion together with the Infinium® MethylationEPIC BeadChip covering over 850,000 methylation sites in the human genome.

To assess differential methylation between WT and TET3KO d8 neurons, the mean methylation percentage of a certain methylation site in WT neurons was subtracted from the mean methylation at the same site in TET3KO neurons. A differential methylation of ≥ 20 % was considered meaningful. In addition, statistical analysis was performed to identify statistically significant (adjusted p-value ≤ 0.05) differences in methylation. 2127 methylation sites met both criteria and are considered truly differentially methylated between WT and TET3KO neurons (Figure 13). Among these differentially methylated sites, 1613 CpGs were hypomethylated and 514 CpGs were hypermethylated in TET3KO neurons compared to WT neurons.

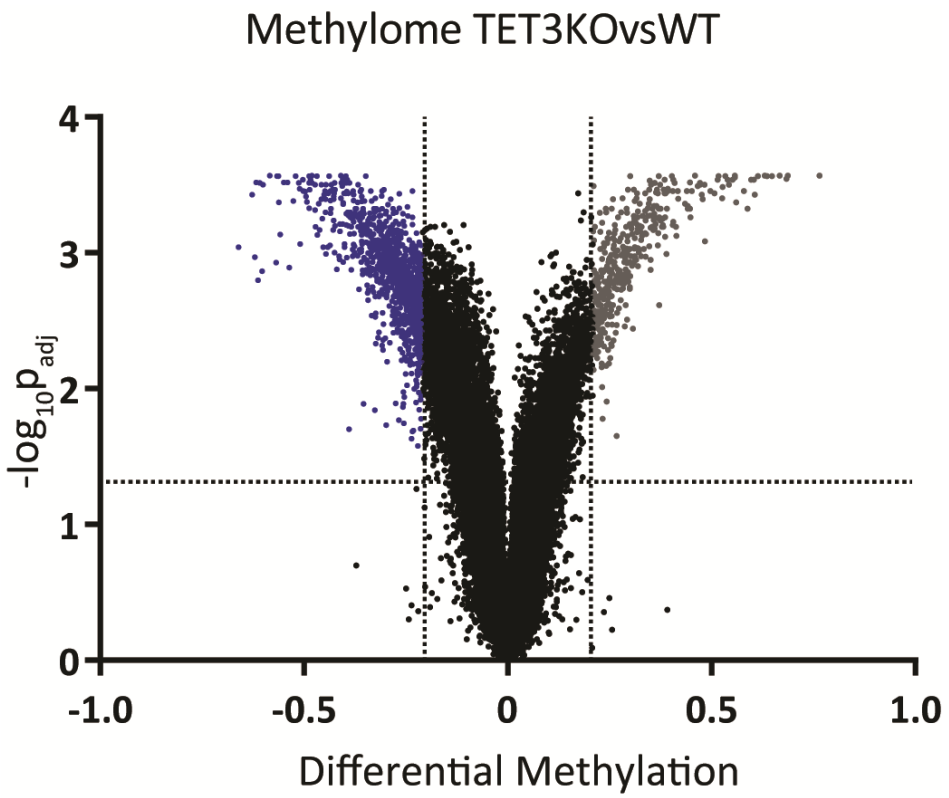


Figure 13 Volcano blot depicting differential methylation between TET3KO and WT neurons

Differential methylation value of each analyzed methylation site is plotted against its $-\log_{10}p_{adj}$. A differential methylation threshold of ±0.2 (20 %) and a $-\log_{10}$(padj) threshold of 1.3 (equals an adjusted p-value of 0.05) were chosen. All methylation sites with significant differential methylation above 0.2 are shown in gray while all methylation sites with significant differential methylation below -0.2 are depicted in blue. A positive differential methylation value marks methylation sites that are hypermethylated in TET3KO neurons (right-hand side of the graph), while a negative differential methylation identifies hypomethylated sites (left-hand side).

For validation of the Methylation BeadChip results, bisulfite pyrosequencing as an independent method was performed. In detail, WT and TET3KO gDNA was bisulfite-converted and pyrosequencing of two regions was performed to detect any differential methylation as proposed by the methylation chip. The methylation status of eight CpGs each in the promoter region of the *family with sequence similarity 122C* (*FAM122C*) and the *protocadherin alpha 6* (*PCDHA6*) locus was quantified. The difference between the methylation percentages obtained in WT and TET3KO neurons was calculated for each CpG and compared to the differential methylation values obtained in the BeadChip experiment.

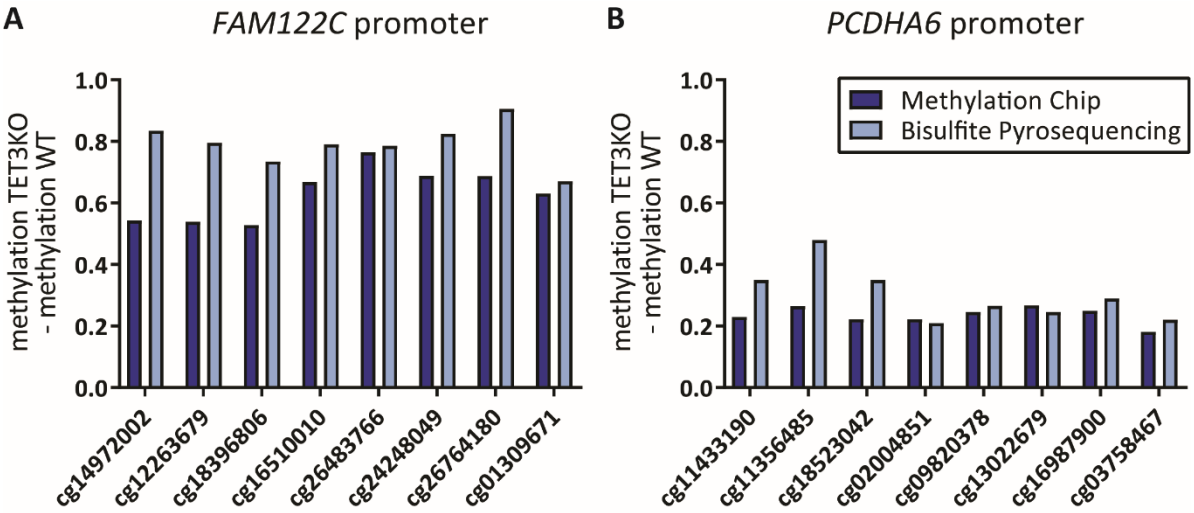


Figure 14 Validation of selected differentially methylated CpGs by bisulfite pyrosequencing

Methylation differences of selected CpGs between TET3KO and WT iNGN-derived d8 neurons as determined by methylation chip and bisulfite pyrosequencing are depicted. Eight CpGs each in the promoter region of ***(A)*** *FAM122C* and ***(B)*** *PCDHA6* were investigated and the means of the two biological replicates analyzed per method are plotted.

In general, the differences calculated from the pyrosequencing results of eight CpGs in the promoter of *FAM122C* were similar to those obtained with the methylation chip (Figure 14A). Differential methylation measured by bisulfite pyrosequencing was lower in the *PCDHA6* than in the *FAM122C* promoter which matched the findings from the methylation chip (Figure 14B).

Since hypermethylation of promoter regions is believed to be associated with a repression of gene expression, transcriptional changes resulting from differential methylation between WT and TET3KO iNGNs are of particular interest. Hence, qRT-PCR analysis was performed using primer pairs specific for four genes which showed several differentially methylated CpGs on the methylation chip. In TET3KO neurons, CpGs in the *FAM122C*, *PCDHA6* and *zinc finger protein 577* (*ZNF577*) gene were found to have an increased methylation level by on average 63.79 %, 23.82 % and 51.70 %, respectively (Figure 15A). Concurrently, transcript levels of *FAM122C*, *PCDHA6* and *ZNF577* were significantly decreased in TET3KO neurons compared to WT (Figure 15B). Vice versa, *retinoic early transcript 1L* (*RAET1L*) was picked as a gene, which contains hypomethylated CpGs in TET3KO neurons. The mean methylation of these CpGs was decreased by 47.76 % (Figure 15A) and resulted in significantly elevated *RAET1L* expression in TET3KO neurons (Figure 15B).

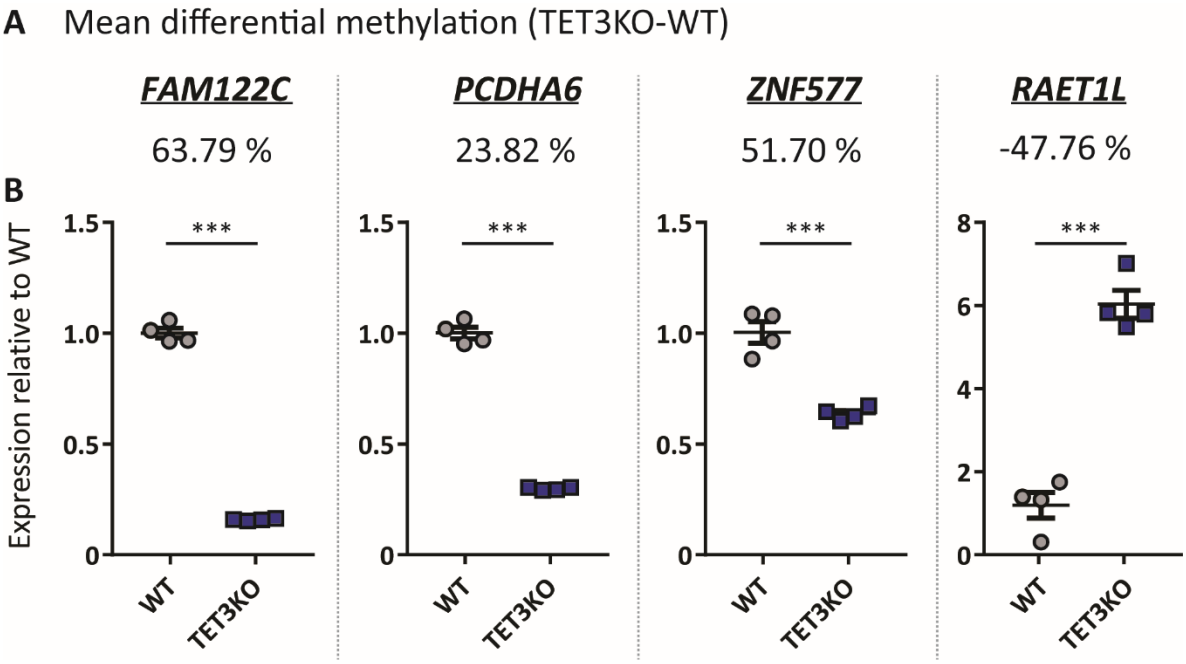


Figure 15 Transcriptional analysis of selected differentially methylated genes
(A) Mean differential methylation between TET3KO and WT iNGN-derived d8 neurons was calculated from the methylation differences of all CpGs in the *FAM122C*, *PCDHA6*, *ZNF577* and *RAET1L* genes that show differential methylation above the threshold of 0.2 or below -0.2, respectively. ***(B)*** Relative transcript levels of *FAM122C*, *PCDHA6*, *ZNF577* and *RAET1L* in WT and TET3KO iNGN-derived d8 neurons were analyzed by qRT-PCR. Expression is normalized to the respective mRNA levels in WT neurons. *** $p \leq 0.001$ (Unpaired t test, two-tailed), n=4.

To better understand the possible effects of the differentially methylated sites, gene ontology (GO) analysis was performed. To identify an enrichment of differentially methylated genes in a certain biological process, Ingenuity® Pathway Analysis software was used. The gene set that is hypomethylated in TET3KO neurons compared to WT showed a highly significant enrichment for genes involved in neuronal function and signaling comprising diverse receptors, such as acetylcholine, glutamate and γ-aminobutyric acid (GABA) receptors (Figure 16A). In contrast, Ingenuity® Pathway Analysis with the hypermethylated gene set mainly resulted in canonical pathways involved in metabolism (Figure 16B).

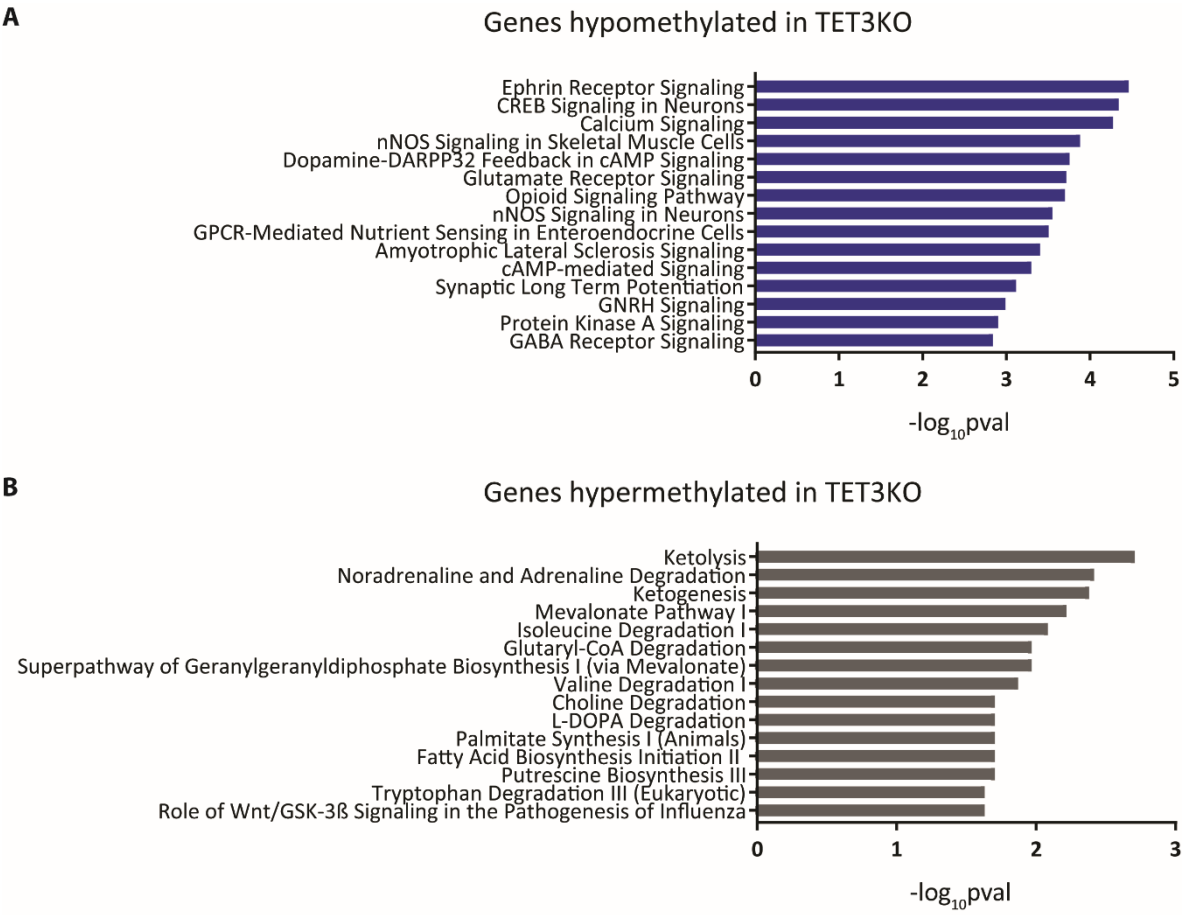


Figure 16 Significantly enriched canonical pathways of hypo- and hypermethylated genes in TET3KO neurons
The top 15 significantly enriched canonical pathways resulting from Ingenuity® Pathway Analysis on the ***(A)*** hypomethylated and ***(B)*** hypermethylated genes in TET3KO neurons compared to WT neurons are shown. The -$\log_{10}$(p-value) indicates the significance of enrichment of genes in the respective pathway.

4.1.5 Global analysis of differential gene expression

To get a more conclusive insight into the transcriptional changes caused by depletion of functional TET3, the complete transcriptome of d8 TET3KO and WT iNGN-derived neurons was analyzed by RNA-seq.

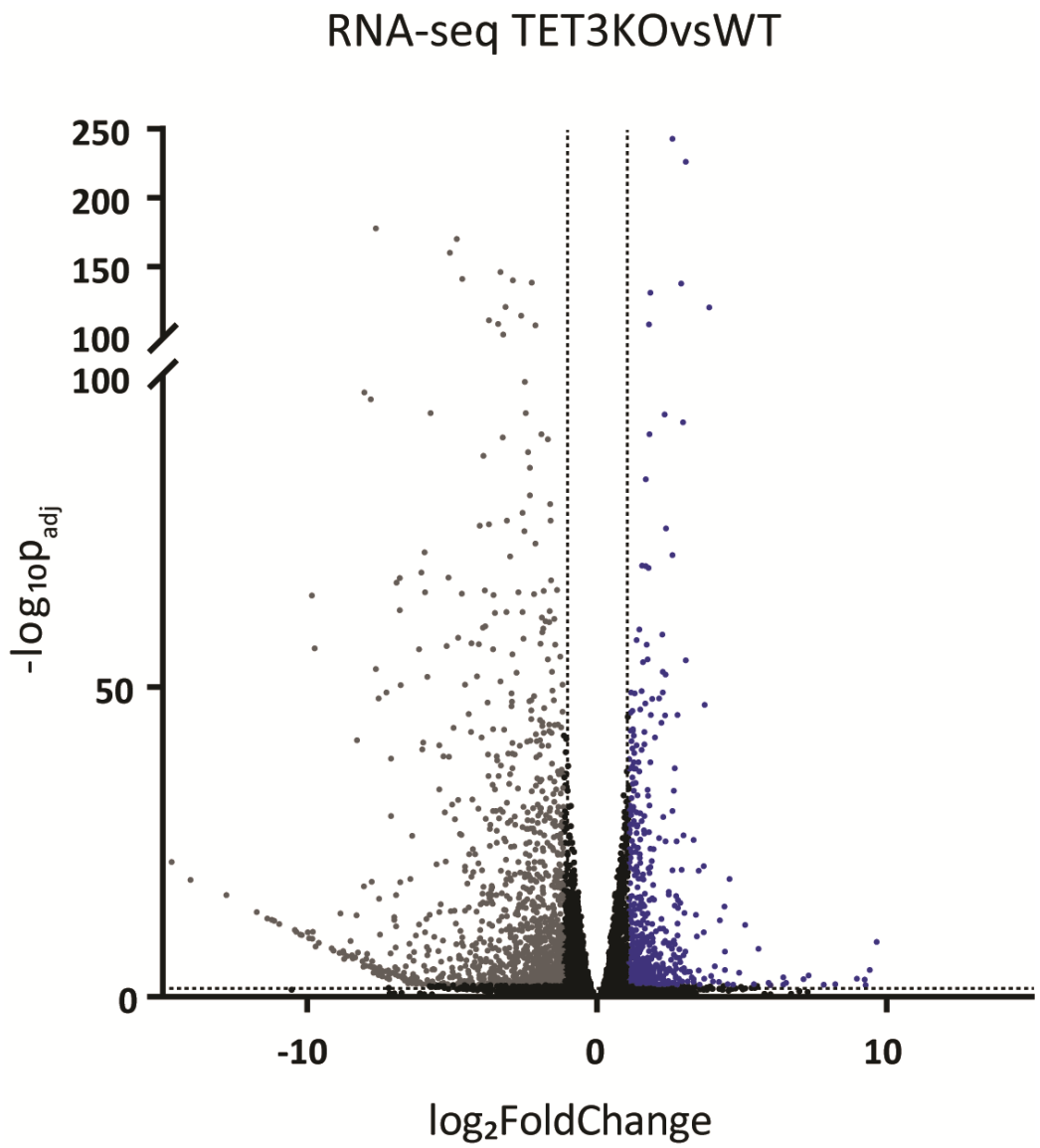


Figure 17 Volcano blot depicting transcriptomic changes in TET3KO neurons compared to WT neurons

Each transcript is plotted according to expression fold change between TET3KO and WT d8 neurons (log_2FoldChange) and adjusted p-value ($-log_{10}p_{adj}$). The threshold value for the expression fold change was set to log_2FoldChange ≥ 1 and ≤ -1 (equals a fold change of 2 and 0.5, respectively). For the significance level, $-log_{10}p_{adj} \leq 1.3$ (equals a $p_{adj} \leq 0.05$) was chosen as cut-off value. Significantly upregulated transcripts are depicted on the right-hand side of the graph in blue, while downregulated transcripts are shown in gray on the left-hand side.

As evident from the volcano blot (Figure 17), a substantial number of genes (2486 genes) are differentially expressed in TET3KO d8 neurons compared to WT d8 neurons. Interestingly, more transcripts are downregulated (1630 genes) in TET3KO neurons than upregulated (856 genes) when compared to WT neurons.

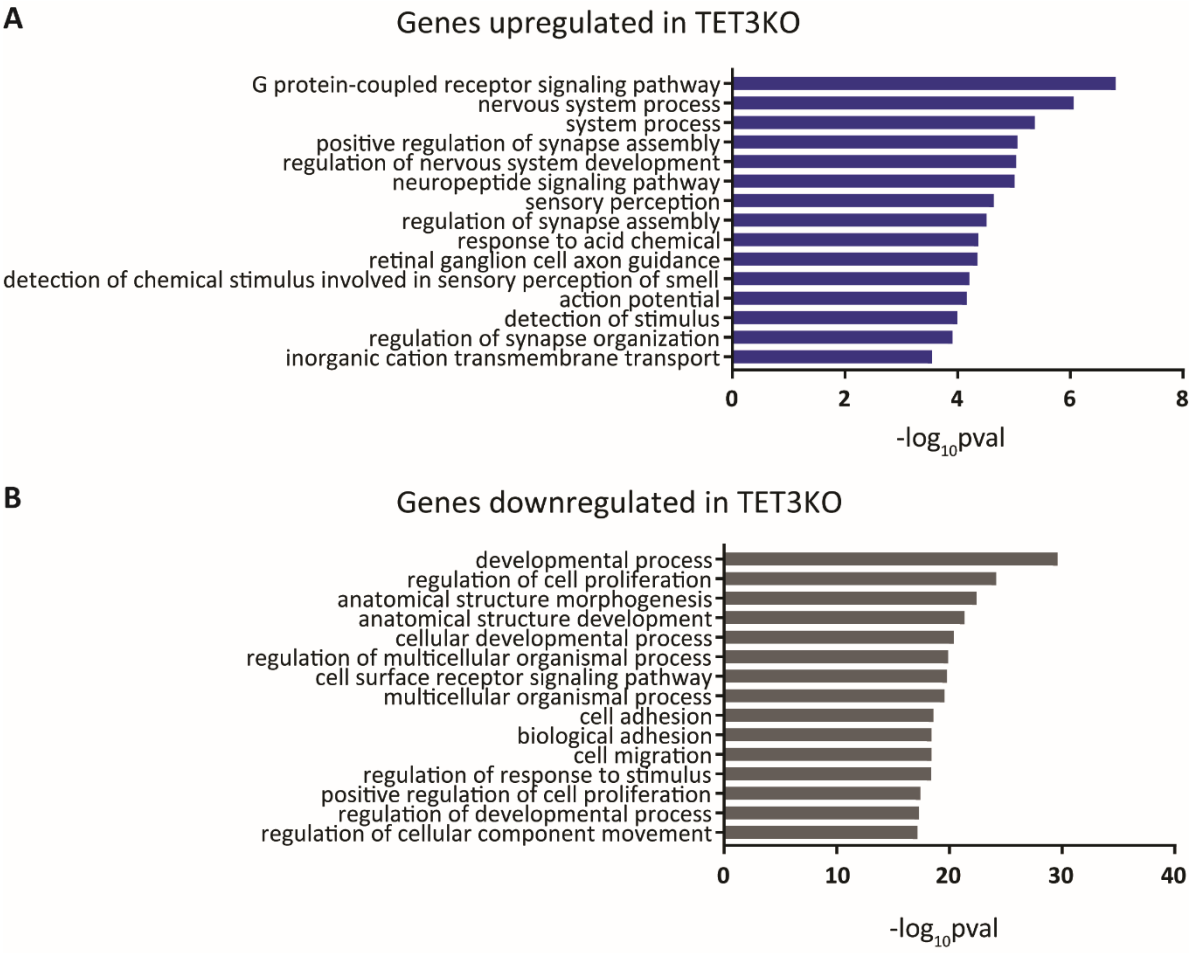


Figure 18 Gene Ontology processes enriched in differentially regulated genes in TET3KO neurons

The top 15 Gene Ontology (GO) processes enriched in a list of ***(A)*** significantly upregulated or ***(B)*** significantly down-regulated genes in TET3KO neurons compared to WT neurons (DOX d8) are shown. The GOrilla tool was applied for GO enrichment analysis and GO terms are ranked according to the significance of their enrichment (-log10pval).

GO analyses revealed that a statistically significant percentage of the genes that were upregulated in TET3KO neurons in comparison to WT neurons are associated with neuronal function including synapse formation and receptor signaling (Figure 18A). In contrast, a significant enrichment of genes involved in development and proliferation as well as cell adhesion and migration was found in the set of genes that are downregulated in the TET3KO (Figure 18B). For validation of the RNA-seq data, qRT-PCR analysis of selected transcripts was performed in d8 WT and TET3KO neurons (Figure 19).

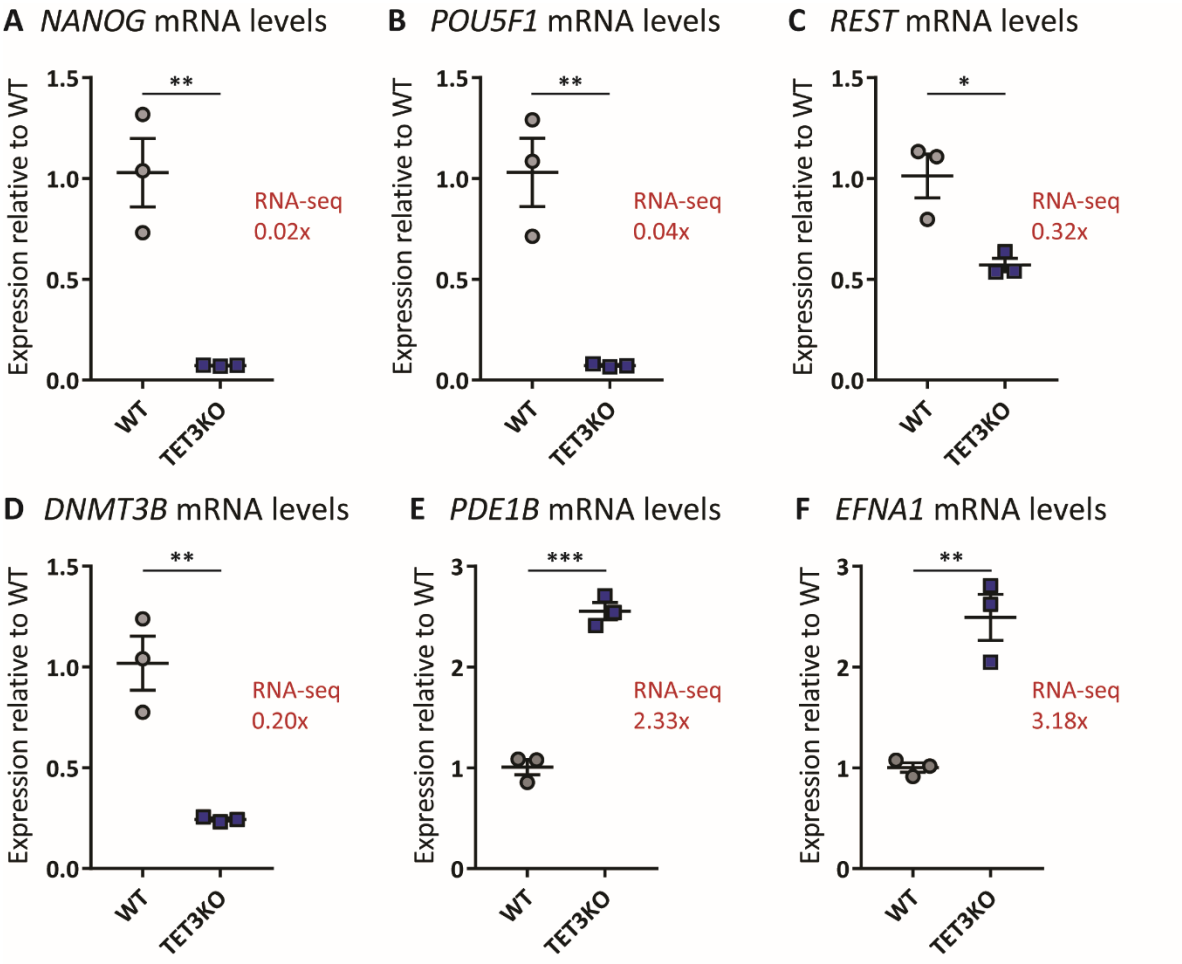


Figure 19 Validation of selected RNA-seq results by qRT-PCR
Relative ***(A)*** *NANOG homeobox*, ***(B)*** *POU class 5 homeobox 1* (*POU5F1*), ***(C)*** *REST*, ***(D)*** *DNMT3B*, ***(E)*** *PDE1B* and ***(F)*** *ephrin A1* (EFNA1) transcript levels in WT and TET3KO d8 neurons were analyzed by qRT-PCR. Expression is normalized to the respective mRNA levels in WT neurons. In red the fold change for the respective transcript measured between WT and TET3KO neurons in the RNA-seq experiment is noted. * $p \leq 0.05$, ** $p \leq 0.01$, p *** $p \leq 0.001$ (Unpaired t test, two-tailed), n=3.

Interestingly, after eight days of DOX treatment, WT iNGN-derived neurons still express the pluripotency markers NANOG and OCT4 (*POU5F1*), while their expression is almost completely lost in TET3KO neurons at the same timepoint (Figure 19A, B). Moreover, *REST* and *DNMT3B* expression is also significantly lower in TET3KO compared to WT at d8 of iNGN differentiation (Figure 19C, D). In contrast, *PDE1B* and *ephrin A1* (*EFNA1*) show higher transcript levels in TET3KO neurons compared to WT (Figure 19E, F). For all analyzed transcripts, the direction of expression changes detected in the qRT-PCR matched the fold change measured by RNA-seq.

4.1.6 Neuronal differentiation

Based on the observed transcriptional changes, differences in WT and TET3KO iNGN differentiation or neuronal function would be expected. To compare the differentiation process of WT and TET3KO iNGN cells in detail, images were taken from d1 to d8 of DOX application. Surprisingly, the time-lapse analysis did not reveal any prominent differences between the two genotypes. Just like WT cells, TET3KO iNGN cells were able to differentiate with similar kinetics and the resulting neurons showed a comparable cell morphology (Figure 20).

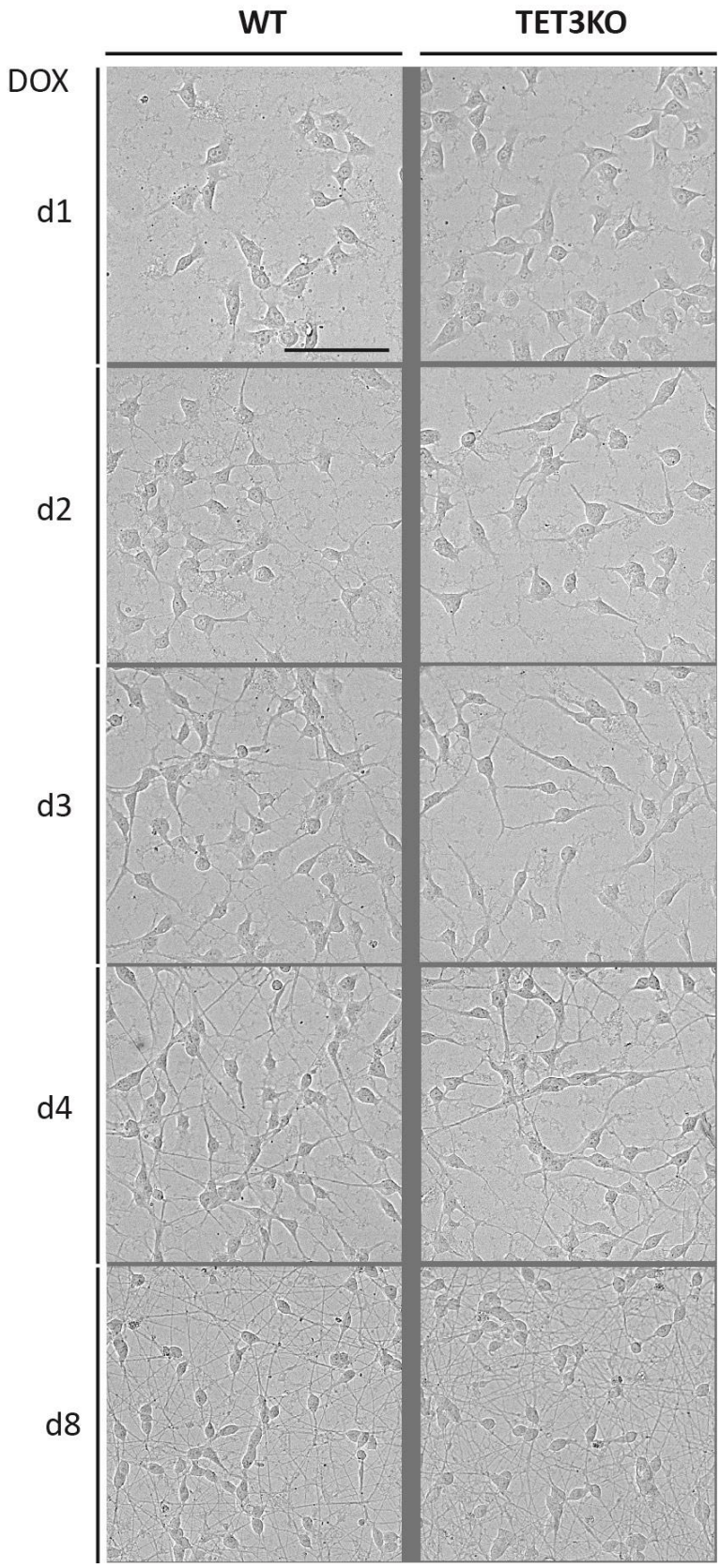


Figure 20 Comparison of WT and TET3KO iNGN differentiation
Representative images of WT (left column) and TET3KO (right column) iNGN cells after 1, 2, 3,4 and 8 days of DOX application are shown. Scale bar 100 µm.

4.1.7 Off-target analysis in TET3KO iNGN cells

It is widely known that off-target effects can be a big problem when working with CRISPR-Cas9. Hence, an off-target analysis in the TET3KO iNGN genome was performed on the five most likely regions per used sgRNA. The analysis aimed at identifying potential off-target mutations resulting from Cas9-mediated DSB generation and subsequent incorrect repair. Off target sequences and corresponding genomic loci analyzed for off-target mutations are listed in Table 9.

Table 9 Off-target analysis in TET3KO iNGN cells
The sgRNA sequence as well as 5 off-target sequences each in the human genome are given. Mismatches of the off-target sequence to the sgRNA target sequence are marked in red. Genomic location and position in the gene are shown. A green tick indicates that no changes in the off-target sequence were found in TET3KO iNGNs.

sgRNA sequence	off-target sequence	genomic location	position in gene	intact?
TET3 up 1	TATAGGGGTTTACGAAAGTAAGG	chr8:64013974-64013996	intron: lncRNA LOC102724623	✓
CACATGGGTTTACGGAAGTAAGG	AATATTGGTTTACAGAAGTATGG	chr4:149354957-149354979	intron: IQCM	✓
	GACATGGGATGACGAAAGTAGGG	chr14:95748692-95748714	intron: lncRNA LOC107984703	✓
	TACATGGGCTTATGGAAGCAAGG	chr5:66181238-66181260	exon: SREK1	✓
	CACATGGGTTAATGGAGTTAGGG	chr4:105631409-105631431	exon: ARHGEF38	✓
TET3 up 2	CAACCTAAATGACCCAAACTAGG	chr3:135858460-135858482	intergenic: LOC105374124-PPP2R3A	✓
CAACCCAAGTGCCCCTAACTGGG	GAACCTAAGGGACCCTAACTGGG	chr14:98205258-98205280	lncRNA LOC107987207	✓
	AAGCCCAAGTGCCTCCAACTTGG	chr9:126003580-126003602	intergenic: PBX3-LOC105376273	✓
	AAACCCAGGAGGCCCTAACTTGG	chr20:5994402-5994424	exon: MCM8	✓
	CAACACAACTGCACCAAACTGGG	chr11:95979520-95979542	exon: MAML2	✓
TET3 down 1	CAAGTAATCTGATGTAAAAATGG	chr19:34652460-34652482	intron: SCGB1B2P	✓
CAGGTAGCCTGGTGTAAAAAAGG	CAGATAAAGTGGTGTAAAAAAGG	chr17:71070839-71070861	intergenic: LOC100131241-CASC17	✓
	CAGGAAGCTTGATGTAAAAGAGG	chr11:39714220-39714242	intron: LOC105376637	✓
	CGGGTAGCCAGCTATAAAAAAGG	chr12:14613172-14613194	exon:GUCY2C/PLBD1-AS1	✓
	CAGGTCGACTAGAGTAAAAAAGG	chr10:73801780-73801802	exon:ZSWIM8/ZSWIM8-AS1/NDST2	✓
TET3 down 2	TATCCCTTCACTTTTGAACTAGG	chr6:19729084-19729106	intergenic: LOC105374959-LOC100506885	✓
TATCGCTTTACTTCTGGACTAGG	TCTGGATCTACTTCTGGACTTGG	chr2:118640735-118640757	intergenic: LOC101927709-EN1	✓
	TATTGCCTTACTTCAGAACTTGG	chr4:87472401-87472423	intergenic: : LOC105377321-SPARCL1	✓
	CATCGCTGGACTTCTGGACATGG	chr3:52481846-52481868	exon: NISCH	✓
	TAGCTCTTTACTTCTGTATTTGG	chr10:96324263-96324285	exon: DNTT	✓

Importantly, none of the 20 tested loci showed any differences in the off-target sequence in the TET3KO clone as compared to the human reference genome indicating low off-target activity by the used Cas9 RNP complex.

4.2 Transactivation

4.2.1 Transcriptional activation of *Cnga1*

In the long run, the transactivation project is aiming at developing novel gene therapies for inherited retinal degenerative diseases caused by various mutations in genes essential for photoreceptor function. The basic assumption is that homologous genes in rod and cone photoreceptors can functionally compensate for each other when they are ectopically expressed. Hence, dCas9-based transcriptional activators could be used to trigger the ectopic expression of the functionally similar gene in photoreceptors affected by a disease-causing mutation in the homologous gene. The homologous protein might take over the task of the mutated one thereby restoring photoreceptor function and preventing degeneration. To initially test the feasibility of this idea, rod-specific *Cnga1* was chosen as a first target and its transcriptional activation in 661W cells was explored. 661W cells represent a suitable murine *in vitro* model of cone photoreceptors as they are easy to cultivate and express many cone-specific genes, while rod-specific proteins are absent (Tan et al., 2004). For proper function, the transcriptional activators need to be located to the promoter region of the target genes. Thus, three sgRNAs were designed that target the transcriptional activators to 104, 270 or 309 bp upstream of the transcription start site (TSS) of *Cnga1* (Figure 21A).

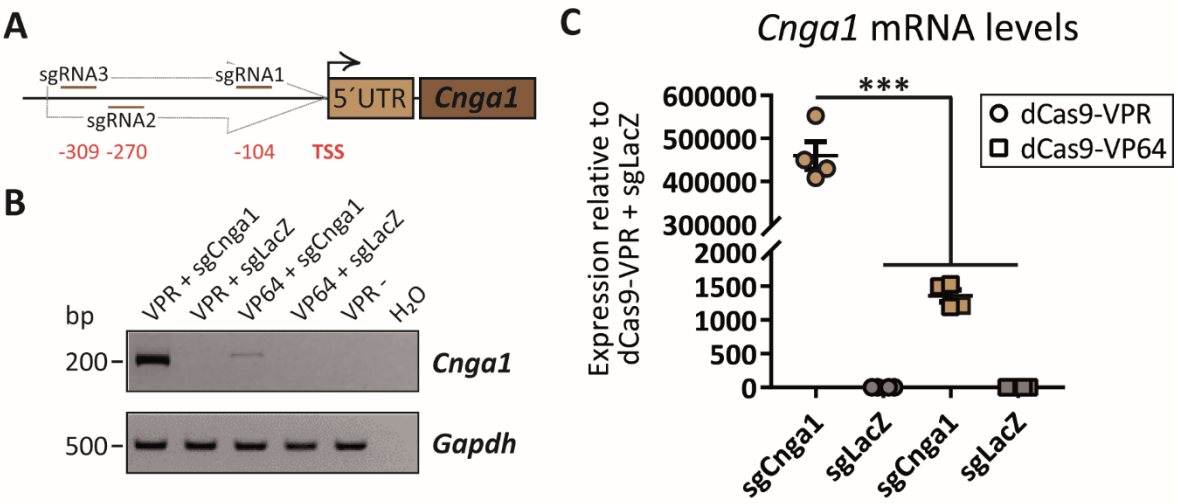


Figure 21 Transcriptional activation of Cnga1 in 661W cells
(A) Schematic representation of the position of the three sgRNAs used for targeting dCas9-based transcriptional activators to the promoter of the *Cnga1* gene. The relative distance of each sgRNA to the transcription start site (TSS) of the target gene is given in bp. ***(B)*** RT-PCR results from 661W cells co-transfected with either dCas9-VPR or dCas9-VP64 in combination with either *Cnga1* specific sgRNAs or a LacZ control sgRNA. For amplification of cDNA, specific primers for *Cnga1* and *Glyceraldehyde-3-phosphatase* (*Gapdh*) transcript as a loading control were used. ***(C)***

qRT-PCR results using *Cnga1* specific primers of co-transfected 661W cells. Expression levels were normalized to 661W cells co-transfected with dCas9-VPR and LacZ sgRNA. Data points of dCas9-VP64 transfection are depicted as circle while dCas9-VPR results are depicted as rectangle. *** $p \leq 0.001$ (Ordinary one-way ANOVA with Dunnett´s multiple comparisons test), n=4.

The two published transcriptional activators dCas9-VP64 (Mali et al., 2013b) and dCas9-VPR (Chavez et al., 2015) were compared in their ability to stimulate *Cnga1* transcription using these three sgRNAs simultaneously. For this, 661W cells were co-transfected with plasmids encoding for the sgRNAs and CMV promoter-driven dCas9-VP64 or dCas9-VPR, respectively. As a negative control, a non-targeting sgRNA specific for the *E.coli LacZ* gene was used in combination with the transcriptional activators. 48 h post transfection, RNA was extracted and reverse transcribed into cDNA that served as template in the subsequent RT-PCR and qRT-PCR analyses. When no or the control *LacZ* sgRNA were co-expressed, no *Cnga1* transcript could be detected in 661W cells while the three *Cnga1* specific sgRNAs could mediate activation of *Cnga1* transcript expression by dCas9-VPR and dCas9-VP64 (Figure 21B). Quantification of transcript levels confirmed a highly significant increase (> 460,000 x) in *Cnga1* mRNA by dCas9-VPR and *Cnga1* specific sgRNAs compared to control sgRNA transfection (Figure 21C). As expected from literature (Chavez et al., 2015) the less potent transcriptional activator dCas9-VP64 could only induce a minor increase in *Cnga1* mRNA levels. Hence, dCas9-VPR was chosen for all further transactivation experiments.

To further characterize and validate the transcriptional activation of *Cnga1*, 661W cells were generated that enable an inducible expression of dCas9-VPR. For this, PiggyBac technology was applied to generate two 661W cell lines that contain the dCas9-VPR coding sequence under control of the tetracycline response element (TRE), a reverse tetracycline-controlled transactivator (rtTA) expression cassette and either the three *Cnga1* specific sgRNAs or the *LacZ* control sgRNA in their genome (Figure 22A). Those cells are referred to as 661W-pb cells in the following. Only in the presence of doxycycline, a tetracycline derivate, rtTA can bind to the TRE and thereby induce the expression of dCas9-VPR (Figure 22B), while the sgRNAs are continuously expressed from human U6 promoters. Based on the protocol used for iNGN cell differentiation, 500 ng/ml DOX were chosen for initiation of dCas9-VPR expression in the initial experiments. In detail, 661W-pb cells expressing *Cnga1* specific sgRNAs or *LacZ* sgRNA were cultured with DOX containing medium for 48 h and subsequently, *Cnga1* transcript levels were analyzed by qRT-PCR (Figure 22C).

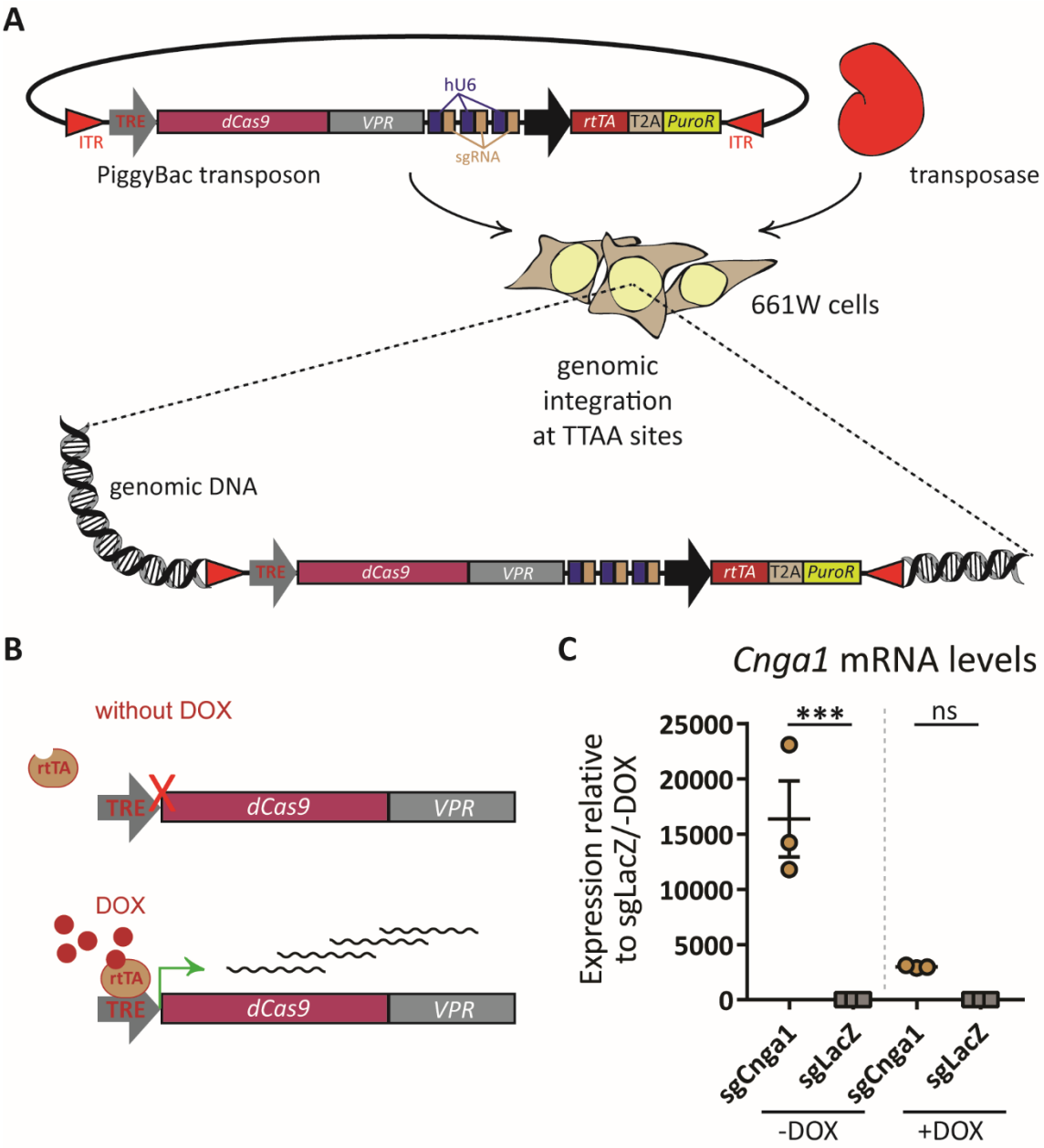


Figure 22 Generation of 661W-pb cells with inducible dCas9-VPR expression
(A) PiggyBac (pb) technology was used for the generation of 661W-pb cells expressing inducible dCas9-VPR in combination with either *Cnga1* specific sgRNAs or a LacZ control sgRNA. The PiggyBac transposase catalyzes the integration of the transposon sequence between the inverted terminal repeats (ITR) into TTAA sites in the genome. ***(B)*** Expression of dCas9-VPR in the 661W cells is controlled by the Tet-On system. The reverse tetracycline-controlled transactivator (rtTA) can only bind to the tetracycline responsive element (TRE) and initiate transcription when doxycycline (DOX, a tetracycline derivate) is present. ***(C)*** qRT-PCR results using *Cnga1* specific primers of 661W-pb cells without DOX (-DOX) and with 500 ng/ml DOX (+DOX) are shown. Expression levels were normalized to the control cell line expressing *LacZ* sgRNA without DOX. *** $p \leq 0.001$, ns $p > 0.05$ (Ordinary one-way ANOVA with Sidak´s multiple comparisons test), n=3.

Surprisingly, the highest activation of *Cnga1* could be observed in 661W-pb cells expressing *Cnga1* specific sgRNAs in the absence of DOX. In the presence of 500 ng/ml DOX, *Cnga1* expression was increased, but not to the extent seen in the absence of DOX. Irrespective of DOX, 661W-pb cells with *LacZ* sgRNA did not show expression of *Cnga1*. The activation of *Cnga1* in the absence of DOX hints to a leaky expression of *dCas9-VPR*, even without DOX. To further investigate the dependencies between *dCas9-VPR* expression and *Cnga1* activation levels, the dose-response relationship for DOX-induced expression was determined using four different DOX concentrations (Figure 23).

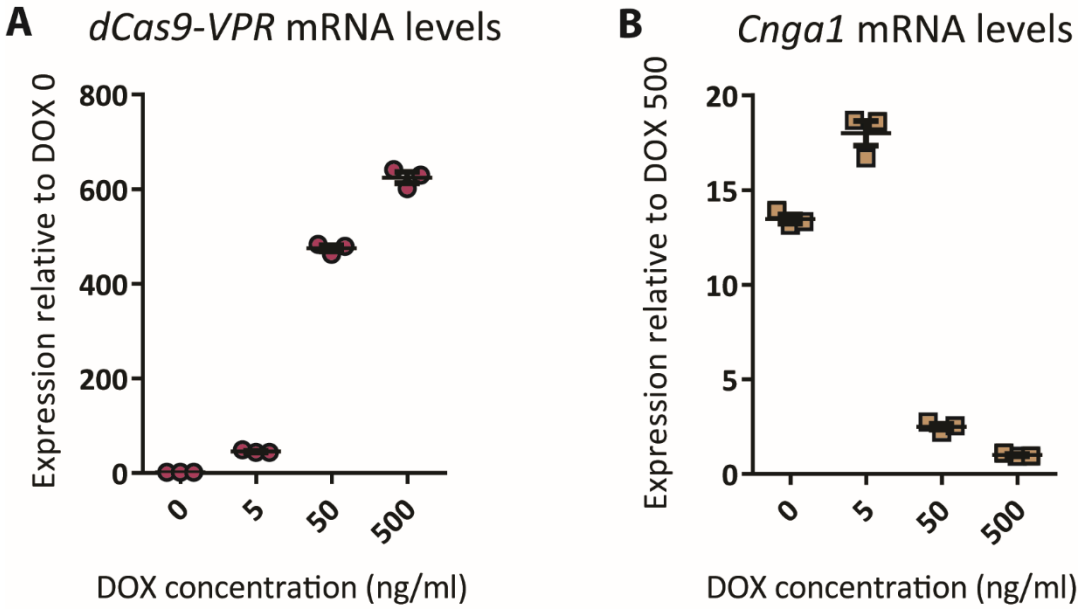


Figure 23 Inverse correlation between Cnga1 activation efficiency and dCas9-VPR expression levels

(A) Relative expression of *dCas9-VPR* in 661W-pb cells as measured by qRT-PCR using *Cas9* specific primers is shown for a DOX concentration of 0, 5, 50 and 500 ng/ml. *dCas9-VPR* expression levels are normalized to 0 ng/ml DOX. ***(B)*** Relative expression of *Cnga1* in 661W-pb cells as measured by qRT-PCR using *Cnga1* specific primers is shown for a DOX concentration of 0, 5, 50 and 500 ng/ml. *Cnga1* expression levels are normalized to 500 ng/ml DOX.

Using *Cas9* specific primers, the leaky expression of *Cas9-VPR* transcript without DOX could be confirmed, while the expected positive correlation between DOX concentration and *dCas9-VPR* expression level was observed (Figure 23A). In contrast, a negative correlation between *dCas9-VPR* and *Cnga1* expression levels was revealed (Figure 23B). At high *dCas9-VPR* levels (50-500 ng/ml DOX), *Cnga1* activation was found to be reduced compared to lower *dCas9-VPR* expression (0-5 ng/ml DOX). Since the highest *Cnga1* activation could be achieved at a DOX concentration of 5 ng/ml, this concentration was used in all further 661W-pb experiments.

To further validate the 661W-pb cell line, immunocytochemistry was performed using Cas9 and Cnga1 specific antibodies (Figure 24).

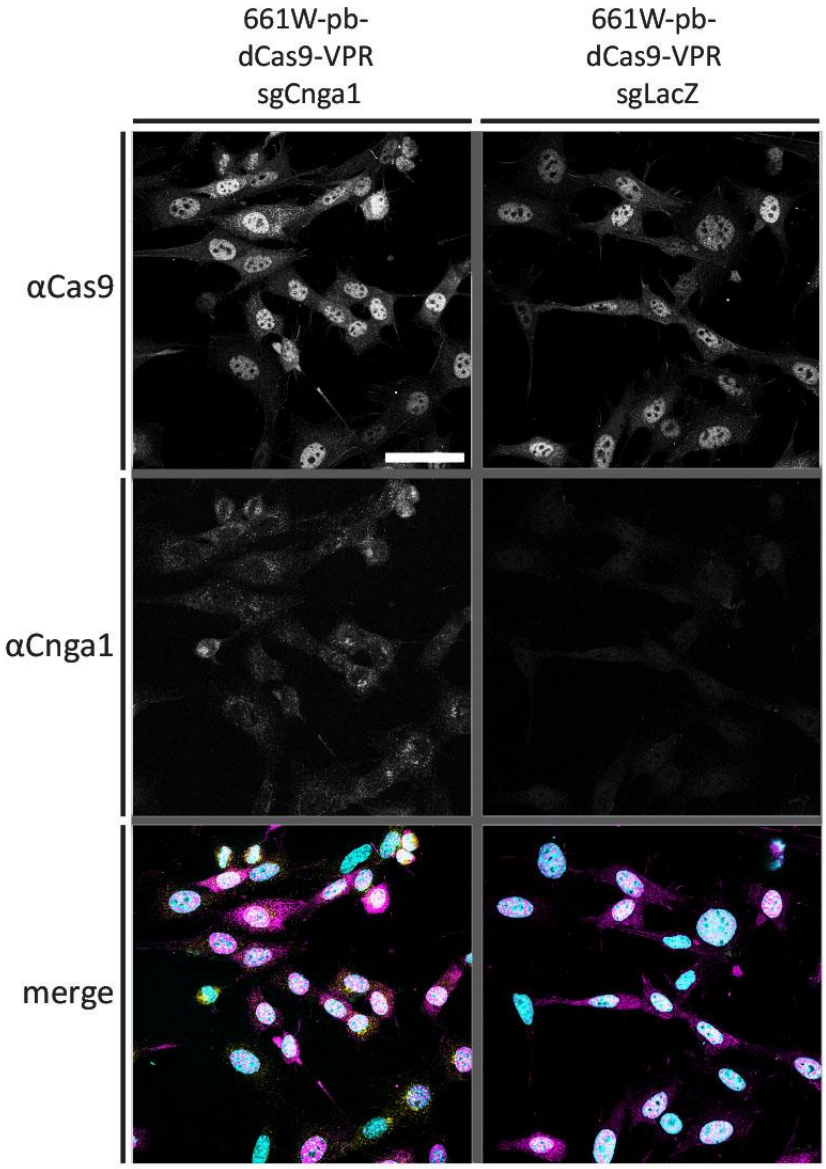


Figure 24 Immunocytochemistry on 661W-pb cells

661W cells expressing *Cnga1* (left column) or *LacZ* targeting sgRNAs (right column) were incubated with 5 ng/ml DOX for 48 h to induce dCas9-VPR expression. Immunostainings using αCas9 antibody are shown in the first row. The second row depicts an αCnga1 immunostaining. Merge images (third row) are composed of αCas9 (magenta), αCnga1 (yellow) and Hoechst nuclear (cyan) staining. Scale bar 50 µm.

Immunostainings using a Cas9 specific antibody illustrate the nuclear localization of dCas9-VPR in 661W-pb cells treated with DOX. In the absence of DOX, 661W-pb cells incubated with the same αCas9 antibody exclusively show unspecific staining in the cytoplasm (data not shown). Importantly, staining of Cnga1 protein could be observed in the membranes of 661W-pb cells expressing *Cnga1* specific sgRNAs,

while Cnga1 protein was not detectable when *LacZ* sgRNA was co-expressed with dCas9-VPR.

Furthermore, it was already shown that Cnga1 alone can form a functional CNG channel when heterologously expressed (Kaupp and Seifert, 2002). This raised the question if functional CNG channels are also present in the 661W-pb cell line. Therefore, inside-out voltage-clamp measurements were performed on patches excised from 661W-pb cells expressing dCas9-VPR together with *Cnga1* or *LacZ* specific sgRNAs at 5 ng/ml DOX.

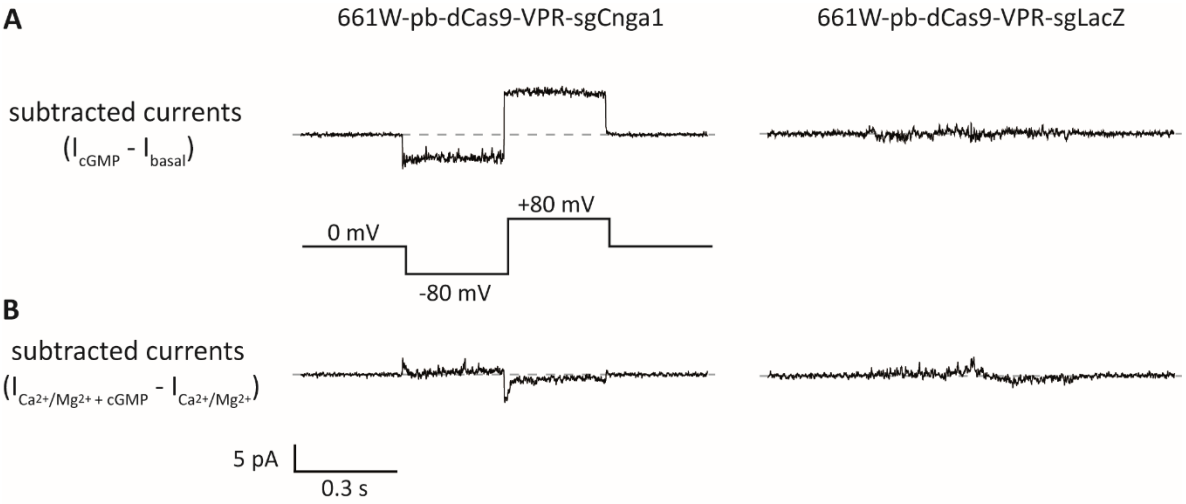


Figure 25 Representative traces of subtracted currents measured by inside-out voltage-clamp on patches of 661W-pb cells
Representative traces of subtracted currents from a 661W-pb cell expressing dCas9-VPR together with sgCnga1 (left column) or sgLacZ (right column) at 5 ng/ml DOX are depicted. ***(A)*** Currents measured under basal conditions were subtracted from the currents evoked by addition of cGMP. ***(B)*** Currents measured under calcium/magnesium blockage (Ca^{2+}/Mg^{2+}) were subtracted from currents evoked by cGMP addition during prevailing calcium/magnesium blockage (Ca^{2+}/Mg^{2+} + cGMP).

To specifically activate CNG channel-mediated currents, cGMP was added to the extracellular solution. Upon cGMP application, a significantly increased current could be observed in the sgCnga1 661W-pb cells only, while no effect on the current was evoked in the control sgLacZ cell line (Figure 25A, Figure 26). Another characteristic feature of CNG channels is the blockage by bivalent ions. Addition of cGMP during simultaneous calcium/magnesium blockage (+ Ca^{2+}/Mg^{2+} + cGMP) did not have any effect on the current in both, sgCnga1 and sgLacZ 661W-pb cells (Figure 25B, Figure 26).

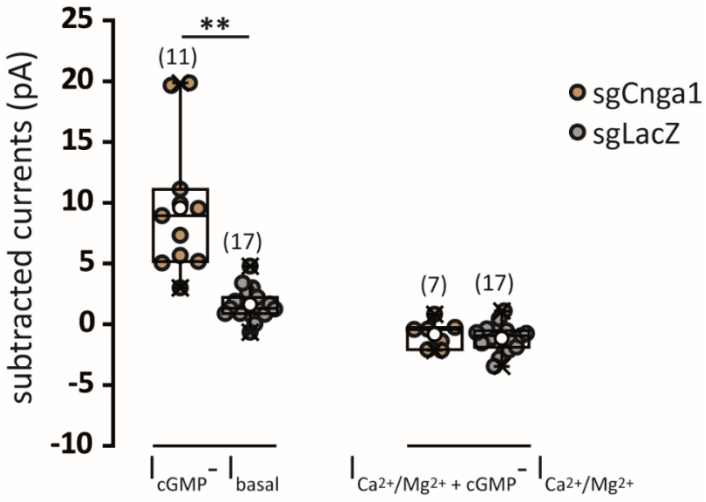


Figure 26 Quantification of subtracted currents measured by inside-out voltage-clamp on patches of 661W-pb cells
Currents measured under basal conditions were subtracted from the currents evoked by addition of cGMP. On the left-hand side these subtracted currents for patches from 661W-pb cells expressing dCas9-VPR in combination with *Cnga1* specific sgRNAs or control *LacZ* sgRNA at 5 ng/ml DOX are compared. Moreover, currents measured under calcium/magnesium blockage (Ca^{2+}/Mg^{2+}) were subtracted from currents evoked by cGMP addition during prevailing calcium/magnesium blockage (Ca^{2+}/Mg^{2+} + cGMP). Those subtracted currents for patches from 661W-pb cells expressing dCas9-VPR in combination with *Cnga1* specific sgRNAs or control *LacZ* sgRNA at 5 ng/ml DOX are shown on the right-hand side of the graph. ** $p \leq 0.005$ (Unpaired t test, two-tailed), number of measured patches is noted in graph.

In summary, these data indicate that the successful activation of *Cnga1* expression in the cone-like 661W cells results in high levels of Cnga1 protein, which is in fact capable of forming functional CNG channels. These results represent an important milestone on the way of developing a transactivation-based treatment of CNG channelopathies.

4.2.2 Development of split dCas9-VPR rAAV vectors for gene therapy

One big obstacle in applying this concept for gene therapy lies in the large size of dCas9-VPR (5.8 kb). The gold standard gene delivery tool in this area are rAAVs, which can only fit a cargo size of ~4.7 kb. To enable rAAV-mediated delivery of dCas9-VPR and *Cnga1* specific sgRNAs, the separation of the dCas9-VPR coding sequence into two parts combined with split intein-mediated protein *trans*-splicing was tested. First, *Sp*Cas9 reconstitution efficiencies using two published split sites were compared. For this, *Rhodothermus marinus (Rma)* N-intein and C-intein sequences were incorporated after amino acid position V713 or E573 of *Sp*Cas9 (Figure 27A).

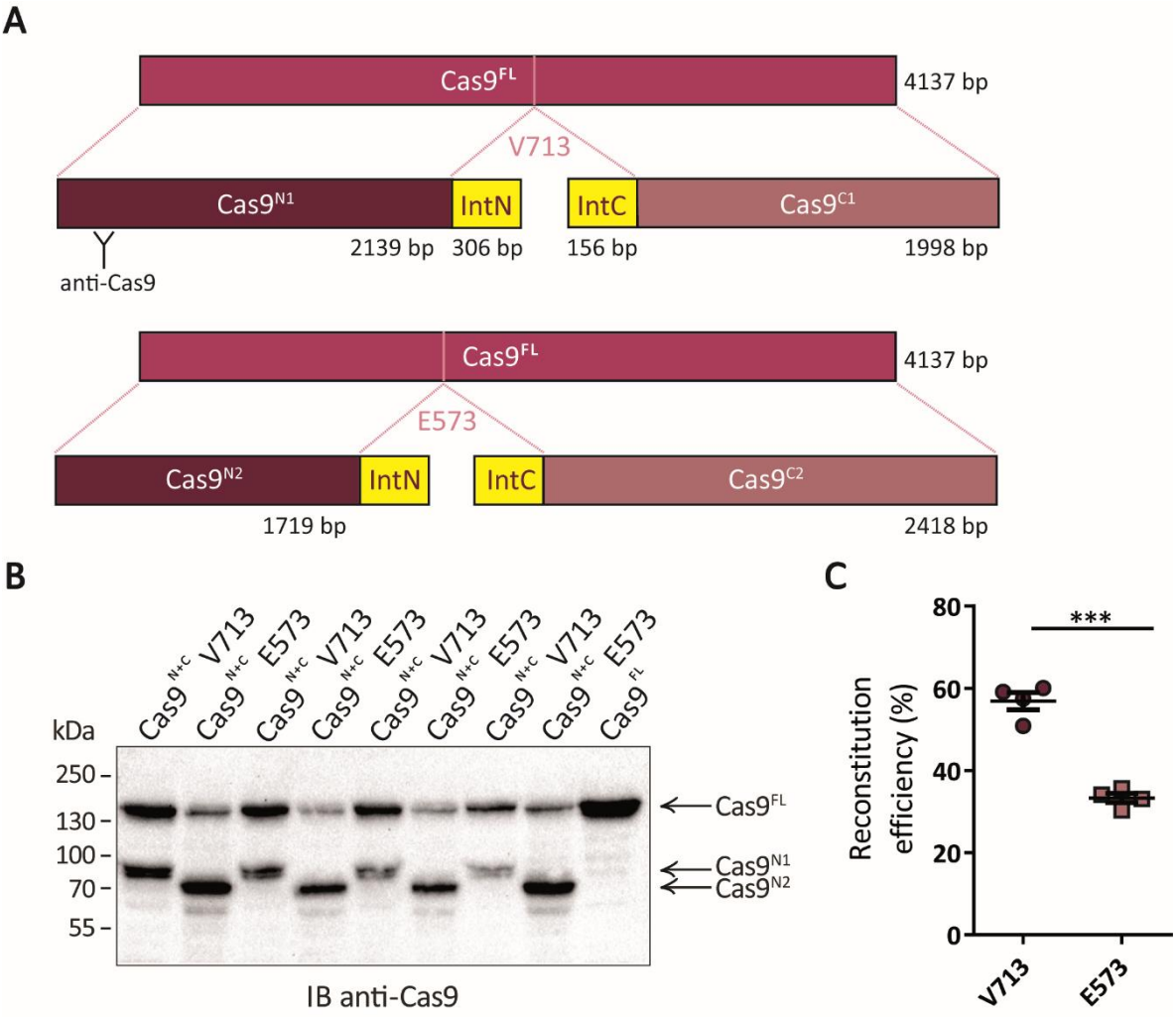


Figure 27 Comparison of V713 and E573 split Cas9 reconstitution efficiency
(A) N-intein and C-intein sequences of *Rhodothermus marinus* (*Rma*) were incorporated after amino acid position V713 or E573 of *Sp*Cas9. ***(B)*** Anti-Cas9 immunoblot (IB) using TX lysates from HEK293 cells co-transfected with the respective Cas9 halves. ***(C)*** Quantification of reconstitution efficiency by measuring the ratios of $Cas9^{FL}$ and $Cas9^{N}$ band intensities. *** $p \leq 0.001$ (Unpaired t test, two-tailed), n=4.

HEK293 cells were co-transfected with plasmids expressing the respective Cas9 halves fused to the intein moieties and whole cell protein was isolated after 48 h. For detection of full-length Cas9 ($Cas9^{FL}$) protein in the subsequent western blotting analysis an antibody was used that specifically binds to the N-terminus of *Sp*Cas9. With both split intein constructs successful protein *trans*-splicing could be achieved as indicated by the ~160 kDa $Cas9^{FL}$ band on the western blot (Figure 27B). By quantifying the band intensities of $Cas9^{FL}$ and non-reconstituted N-terminal SpCas9 ($Cas9^{N}$) fragments, reconstitution efficiencies were calculated. While 33 % of E573 split Cas9 halves were efficiently reconstituted, the V713 constructs mediated a higher reconstitution efficiency of 57 % (Figure 27C). Based on these results, *Rma* split intein sequences were also inserted after V713 of dCas9-VPR. Two rAAV

constructs were cloned containing on the one hand, the *Cnga1* or *LacZ* specific sgRNAs as well as the N-terminal part of dCas9 fused to N-intein and on the other hand, the C-intein fused to the C-terminal part of dCas9-VPR (Figure 28A). Both dCas9-VPR halves were expressed from a CMV promoter. To test if split dCas9-VPR is also capable of activating *Cnga1* transcription, both rAAV vectors were co-transfected into 661W cells and 48 h later *Cnga1* transcript levels were analyzed.

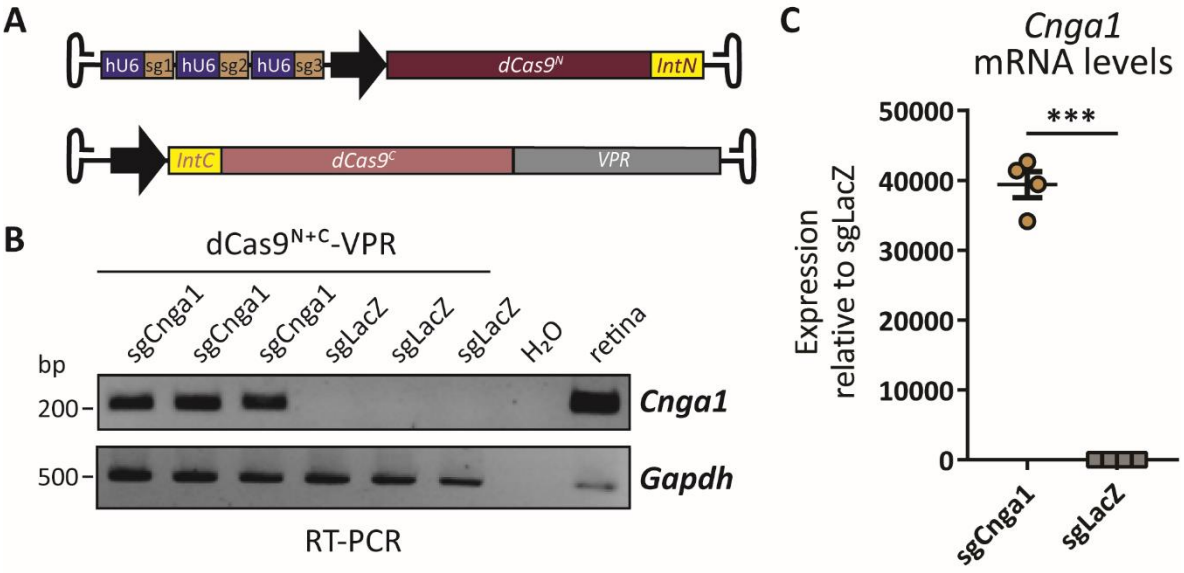


Figure 28 Transcriptional activation of Cnga1 in 661W cells using split dCas9-VPR

(A) 661W cells were co-transfected with rAAV-sgRNA-dCas9N and rAAV-dCas9C-VPR. A rAAV plasmid encoding either for the *Cnga1* specific sgRNAs or the *LacZ* control sgRNA was used. ***(B)*** RT-PCR results using specific primers for *Cnga1* and *Glyceraldehyde-3-phosphatase* (*Gapdh*) as a loading control are shown. As a positive control cDNA synthesized from a retinal RNA sample was employed as template. ***(C)*** qRT-PCR results using *Cnga1* specific primers are illustrated. Expression levels were normalized to 661W cells transfected with split dCas9-VPR and control *LacZ* sgRNAs *** $p \leq 0.001$ (Unpaired t test, two-tailed), n=4.

Importantly, split dCas9-VPR expression with *Cnga1* specific sgRNAs resulted in robust *Cnga1* transcription (Figure 28B). qRT-PCR was used to detect a highly significant increase in *Cnga1* mRNA levels compared to 661W cells co-transfected with rAAV vectors encoding split dCas9-VPR and control *LacZ* sgRNA (Figure 28C).

4.2.3 Transcriptional activation of M- and S-opsin

For validation of the concept in another cell type, transcriptional activation of cone-specific M- and S-opsin in rod photoreceptors was chosen. If successful, the ectopic activation of cone opsins in rods degenerating due to RHO mutations would be of great therapeutic potential. Owing to the lack of a suitable rod-like cell line, the activation of cone opsins was investigated in MEF cells, as it was assumed that they neither express cone nor rod opsins.

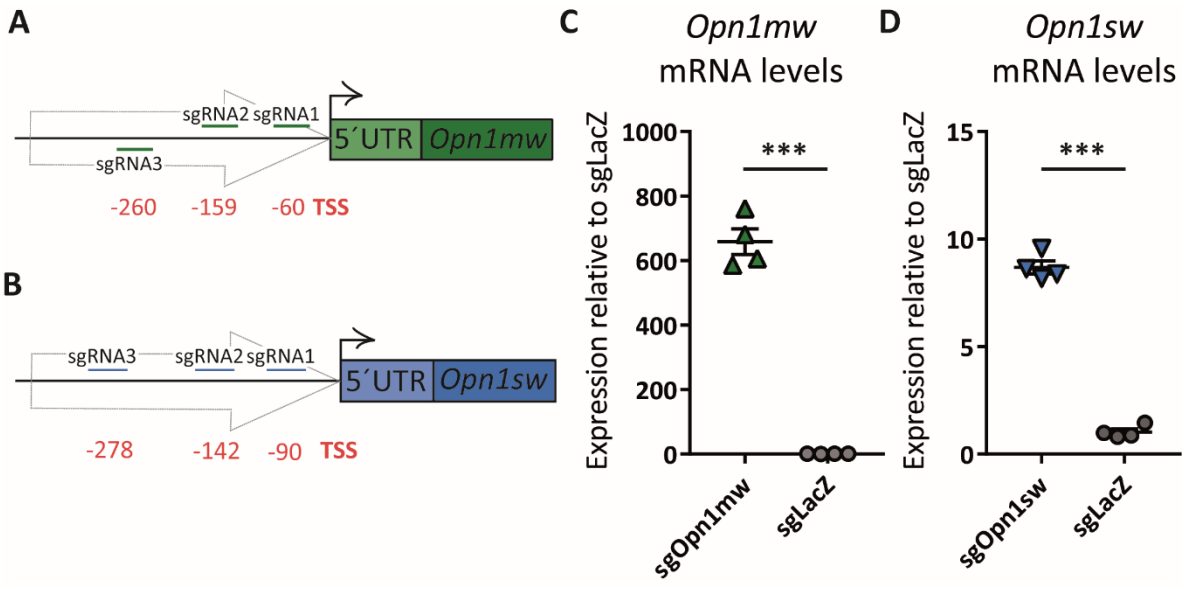


Figure 29 Transcriptional activation of Opn1mw and Opn1sw in MEF cells
Schematic representation of the position of the 3 sgRNAs used for targeting dCas9-VPR to the promoter of the ***(A)*** *Opn1mw* and ***(B)*** *Opn1sw* gene. The relative distance of each sgRNA to the transcription start site (TSS) of the targeted gene is given in bp. MEF cells were co-transfected with full-length dCas9-VPR and sgRNAs targeting *Opn1mw*, *Opn1sw* or control *LacZ*. ***(C)*** qRT-PCR results using *Opn1mw* specific primers are depicted. *Opn1mw* expression levels are normalized to MEF cells transfected with control *LacZ* sgRNA. *** $p \leq 0.001$ (Unpaired t test, two-tailed), n=4. ***(D)*** qRT-PCR results using *Opn1sw* specific primers are shown. *Opn1sw* expression levels are normalized to MEF cells transfected with control *LacZ* sgRNA *** $p \leq 0.001$ (Unpaired t-test, two-tailed), n=4.

To target dCas9-VPR to the promoter of the *Opn1mw* gene encoding for M-opsin, three sgRNAs located 60, 159 and 260 bp upstream of the TSS were designed (Figure 29A). For the *Opn1sw* promoter controlling S-opsin expression, three sgRNAs 90, 142 and 278 bp upstream of the TSS were chosen (Figure 29B). Co-transfection of MEF cells with plasmids encoding FL dCas9-VPR and *Opn1mw* or *Opn1sw* specific sgRNAs resulted in significant activation of *Opn1mw* or *Opn1sw* transcription, respectively. Unexpectedly, this first qRT-PCR analysis using *Opn1mw* specific primers revealed that a certain amount of *Opn1mw* transcript was already present in untreated or sgLacZ transfected MEF cells (data not shown). However, compared to the LacZ control a more than 657-fold increase in *Opn1mw* levels could be obtained by dCas9-VPR (Figure 29C), whereas *Opn1sw* levels could only be increased about 8-fold (Figure 29D). Based on these results further experiments focused on *Opn1mw* transactivation.

To clarify if the inverse correlation between dCas9-VPR expression levels and target gene activation efficiency is a *Cnga1* specific phenomenon, two MEF-pb cell lines

were generated encoding inducible dCas9-VPR and either *Opn1mw* specific sgRNAs or a *LacZ* control sgRNA.

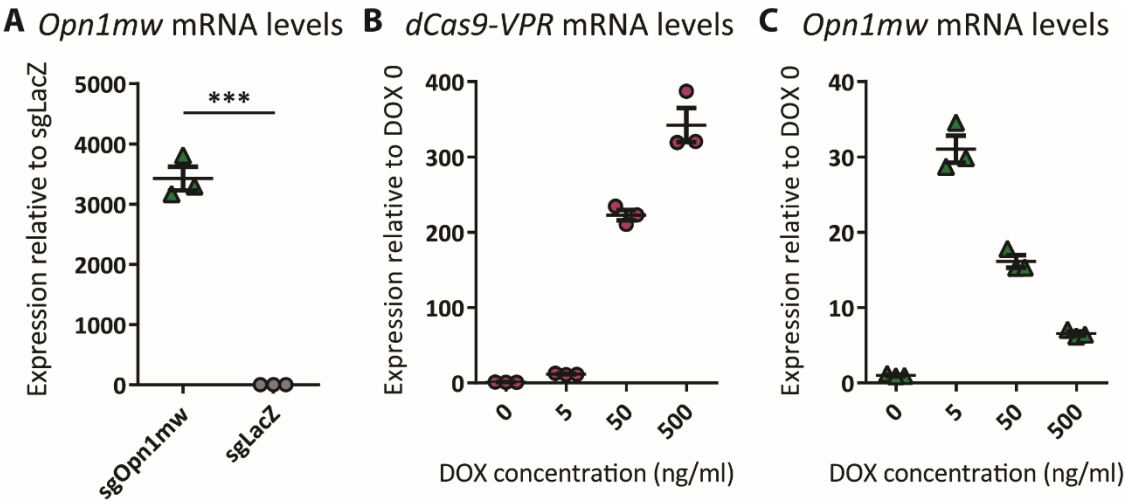


Figure 30 MEF-pb cells stably expressing dCas9-VPR and sgRNAs
(A) PiggyBac technology was used for generation of MEF-pb cells expressing inducible dCas9-VPR in combination with either *Opn1mw* specific sgRNAs or a LacZ control sgRNA. qRT-PCR results using *Opn1mw* specific primers of MEF-pb cells at a DOX concentration of 5 ng/ml are shown. Expression levels were normalized to the control cell line expressing LacZ sgRNA. *** $p \leq 0.001$ (Unpaired t test, two-tailed), n=3. ***(B)*** The relative expression of *dCas9-VPR* as measured by qRT-PCR using *Cas9* specific primers is shown for a DOX concentration of 0, 5, 50 and 500 ng/ml. dCas9-VPR expression levels are normalized to 0 ng/ml DOX. ***(C)*** Relative expression of *Opn1mw* as measured by qRT-PCR using *Opn1mw* specific primers is shown for a DOX concentration of 0, 5, 50 and 500 ng/ml. *Opn1mw* expression levels are normalized to 0 ng/ml DOX.

First, using the previously established concentration of 5 ng/ml DOX, the transcriptional activation of *Opn1mw* in the MEF-pb cells expressing the target sgRNAs could be boosted to almost a 3500-fold compared to control sgRNA expression (Figure 30A). Similar to 661W-pb cells, also this MEF-pb cell line showed a more efficient transcriptional activation at lower levels of *dCas9-VPR* expression (Figure 30B/C). 5 ng/ml was confirmed as the optimal DOX concentration resulting in maximal *Opn1mw* transcript levels. In contrast to the 661W-pb cell line expressing *Cnga1* specific sgRNAs, the MEF-pb-sgOpn1mw cell line did not exhibit such substantial leakiness of dCas9-VPR expression and resulting *Opn1mw* activation without DOX (Figure 30C).

Finally, the activation of *Opn1mw* expression by split dCas9-VPR was tested. For this, rAAV vectors were cloned: one encoding sgRNAs as well as N-terminal dCas9 plus *Rma* N-intein and another one carrying the C-intein fused to the C-terminal dCas9-VPR sequence (Figure 31A).

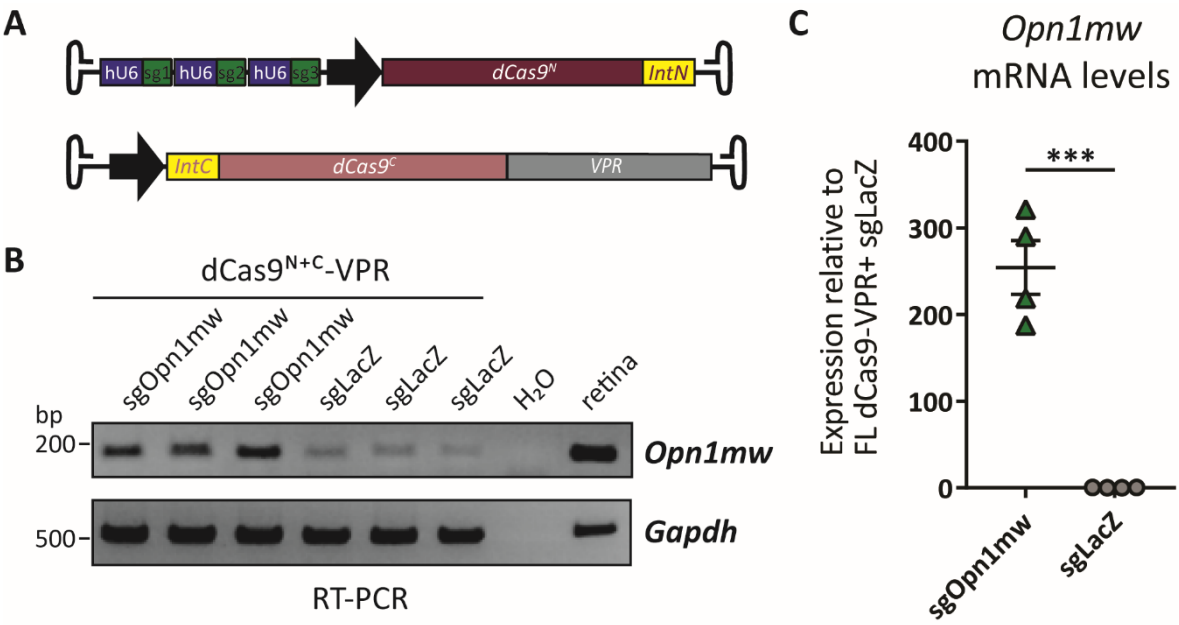


Figure 31 Transcriptional activation of Opn1mw using split dCas9-VPR
(A) MEF cells were co-transfected with rAAV-sgRNA-dCas9N and rAAV-dCas9C-VPR. A rAAV plasmid encoding either for the *Opn1mw* sgRNAs or the *LacZ* control sgRNA was used. ***(B)*** RT-PCR results using specific primers for *Opn1mw* and *Glyceraldehyde-3-phosphatase* (*Gapdh*) as a loading control are shown. As a positive control cDNA synthesized from a retinal RNA sample was employed as template. ***(C)*** qRT-PCR results using *Opn1mw* specific primers are illustrated. Expression levels were normalized to MEF cells transfected with FL dCas9-VPR and control *LacZ* sgRNAs *** $p \leq 0.001$ (Unpaired t test, two-tailed), n=4.

As already noted in the qRT-PCR analysis, WT MEF cells endogenously express *Opn1mw* at low levels. This is also the case in MEF cells co-transfected with split dCas9-VPR and sgLacZ as indicated by amplification of *Opn1mw* cDNA in the RT-PCR (Figure 31B). Nevertheless, co-expression of split dCas9-VPR and *Opn1mw* specific sgRNAs resulted in more intense bands in the RT-PCR providing evidence for higher amounts of *Opn1mw* transcript. Importantly, in the qRT-PCR analysis a significant activation of *Opn1mw* expression could be demonstrated in MEF cells co-transfected with rAAVs expressing split *dCas9-VPR* and *Opn1mw* sgRNAs compared to *LacZ* control sgRNA (Figure 31C).

5 Discussion

5.1 TET3KO iNGN cell line

5.1.1 Validation of TET3KO iNGN cells

The aim was to generate an iNGN cell line that lacks catalytically active TET3 to investigate the effects of the KO on neuronal differentiation. To achieve this, Cas9 protein and four *TET3* targeting sgRNAs were co-delivered into WT iNGN cells. In detail, two of the sgRNAs were positioned upstream of exon 7 and the other two sgRNAs downstream of exon 8 (Figure 9A). The positioning was chosen according to a published conditional TET3KO mouse model (Amouroux et al., 2016). Exon 7 encodes for the beginning of the first oxygenase domain of TET3 and thus, a deletion of exon 7 and 8 by Cas9-mediated DSB generation on both sides represents a partial deletion of the catalytic domain of TET3. Importantly, this deletion renders the reading frame if exon 6 might be spliced to one of the 3´exons (exon 9-11) leading to almost immediate stop codon occurrence. Hence, also the remaining part of the first as well as the second oxygenase domain cannot be translated into functional protein.

After isolation of numerous single clones from the edited iNGN cell pool via FACS only one iNGN clone was identified in a PCR assay to show a successful bi-allelic deletion of exon 7 and 8 (Figure 9B). This low success rate is most likely not due to inefficient genome editing of the TET3 locus but more to the extreme difficulty of culturing iPS cells as single cells. From 120 living (propidium iodide negative) cells that were seeded in singe wells less than 20% (21 cells) survived and formed colonies that could be genotyped. Nevertheless, a TET3KO iNGN cell line could be successfully established from the single positive clone and was further characterized. In the next step, the modifications of the TET3 locus were examined in detail by Sanger sequencing (Figure 9C). Interestingly, this analysis revealed that the outcome of CRISPR-Cas9 genome editing was different on the two alleles. On both alleles, the two outermost sgRNAs were efficient in guiding Cas9 to induce DSBs. On one of the alleles, the DNA was joined seamlessly at the two Cas9 cut sites 3 bp upstream of the PAM, while on the other one a random nucleotide sequence of 68 bp length was inserted. The Nucleotide Basic Local Alignment Search Tool (blastn suite, BLAST, NCBI) could not identify any significant sequence similarity of the inserted sequence to another locus in the human genome ruling out the possibility of homologous recombination. Since this insertion is in the intronic sequence

between exon 6 and 9, it is most likely inconsiderable and characterization of the cell line continued.

Global *TET3* as well as *TET3* isoform specific transcript levels were compared in WT and TET3KO undifferentiated iNGN cells and derived neurons (Figure 10). TET3KO iNGN cells as well as derived neurons at d4 and d8 of DOX application constantly show significantly decreased total TET3, TET3S and TET3FL expression. This observation is more prominent in the neuronal than in the pluripotent state, most probably due to the different expression levels of TET3 in undifferentiated iNGNs and derived neurons. The reduced transcript levels are most likely due to nonsense-mediated mRNA decay eliminating transcripts that contain premature stop codons (Brogna and Wen, 2009). Nevertheless, the expression of TET3 mRNA is not completely abolished in the TET3KO cells. However, this is expected since the deletion is rather late in the sequence and the changes made to the sequence do not primarily affect transcription but translation.

Probably, the biggest weakness of this work, is the missing validation of the TET3KO on protein level. Unfortunately, all attempts of generating WB or immunofluorescence data using αTET3 antibodies failed. More than five different commercial and custom-made antibodies were tested but all of them either yielded unspecific non-nuclear staining in WT iNGN cells (immunocytochemistry) or unspecific bands of the wrong size on the WB.

5.1.2 DNA Methylation in TET3KO Neurons

It is known from literature that neurons contain remarkably high 5hmC levels when compared to non-neuronal tissues (Globisch et al., 2010; Khare et al., 2012; Szwagierczak et al., 2010). So far it was also noted that this 5hmC accumulation occurring with neuronal differentiation correlates with an increase in TET3 expression levels (Hahn et al., 2013). In this study, the quantification of global 5hmC levels in WT and TET3KO iNGN cells at different timepoints of DOX induction revealed that TET3KO neurons do not show higher 5hmC levels than TET3KO iPS cells (Figure 12B). This proves that TET3 indeed is the enzyme catalyzing 5mC oxidation to 5hmC during the process of neuronal differentiation. Surprisingly, global 5mC levels were found to be significantly increased in undifferentiated TET3KO iNGN cells compared to WT iNGNs (Figure 12A). Since TET1 and TET2 transcript levels are unchanged in TET3KO iNGN cells (Figure 11), the effect cannot be attributed to them. Because TET3 expression is relatively low in pluripotent d0 iNGN cells (Figure 8A), catalytic function of TET3 was initially thought to be rather insignificant at this time point. However, considering the increase in 5mC at d0, it

might as well be worth having a closer look into the role of TET3 in the undifferentiated cells.

Besides quantifying global levels of 5mC and 5hmC, also site-specific methylation changes were analyzed using a methylation beadchip. For this, genomic DNA of WT and TET3KO neurons (d8 DOX) was bisulfite-treated. Bisulfite (BS, $NaHSO_3$) treatment of DNA combined with various sequencing methods is probably the most common way to identify 5mC and 5hmC in the DNA at single-base resolution. Importantly, bisulfite treatment of DNA leads to the conversion of unmodified cytosines, 5fC and 5caC to uracil, while 5mC and 5hmC are unaffected by the treatment. After PCR amplification of bisulfite converted DNA, both 5mC and 5hmC are read as cytosines, while unmodified cytosines result in thymine in the sequencing state. Thus, in the performed methylation chip (Figure 13) as well as in the bisulfite-pyrosequencing validation (Figure 14), it is not possible to discriminate between 5mC and 5hmC. Nevertheless, in this data the measured percentages mainly reflect DNA methylation since 60-80 % of CpGs in the genome are methylated while 5hmC only accounts for a comparably small amount of cytosine modifications (Smith and Meissner, 2013). qRT-PCR analysis of four target genes confirmed an inverse correlation between promoter methylation and gene expression level supporting the previous assumption of 5mC based effects (Figure 15). To distinguish between 5mC and 5hmC, the bisulfite sequencing method has been developed further to oxidative bisulfite sequencing (oxBS-seq) or Tet-assisted bisulfite sequencing (TAB-seq). OxBS relies on the specific oxidation of 5hmC to 5fC by potassium perruthenate and the fact that the subsequent bisulfite treatment only leaves 5mC unaffected (Booth et al., 2012, 2013). So, all cytosines read in the final sequencing reaction exclusively represent 5mC on the gDNA level. The levels of 5hmC can then be inferred by subtracting the percentages obtained by oxBS-seq from the ones obtained by traditional BS-seq ((5mC+5hmC)-5mC). Vice versa, TAB-seq analysis exclusively provides information on the percentage of 5hmC in the gDNA. For this method, gDNA is treated with β-glucosyltransferase to specifically glucosylate 5hmC. In the subsequent step, recombinant Tet1 enzyme is used to oxidize 5mC to 5fC, whereas the glucosylation protects 5hmC from oxidation (Yu et al., 2012). The subsequent bisulfite treatment converts 5caC and 5fC (including former 5mC) but not β-glucosyl-5hmC. By comparison to the traditional BS-seq result, the percentage of 5mC can be inferred.

To validate the results obtained by the methylation chip, two loci containing multiple highly differentially methylated CpGs were chosen and CpG methylation was quantified in WT and TET3KO samples using a combination of bisulfite treatment and pyrosequencing. At both loci, the *FAM122C* and *PCDHA6* promoter,

the differential methylation percentages obtained by bisulfite pyrosequencing matched the values or at least the trend of the methylation chip analysis (Figure 14). Alternatively, enabling a cost-effective validation of a broad range of CpGs, reduced representation bisulfite sequencing (RRBS) could be performed. This method relies on the enrichment for CpG-dense regions in gDNA commonly by digesting the genome with the methylation-insensitive restriction enzyme *Msp*I (recognition site C^CGG). Subsequently, only fragments of a certain size range representing most CpG islands, promoter and enhancer areas are extracted, bisulfite treated and finally sequenced using next generation sequencing (NGS) techniques (Meissner et al., 2005; Smith et al., 2009). To obtain reliable single bp resolution data of these CpG-rich restriction fragments a comparably low sequencing depth is required, which in turn means reduced costs compared to whole genome BS-seq (Smith et al., 2009).

5.1.3 Effect of TET3KO on neuronal differentiation and function

Importantly, TET3KO iNGN cells show no obvious phenotypic change and retain the ability of differentiating into bipolar neurons upon DOX application (Figure 20). However, the effects of TET3 absence during the differentiation process and the resulting differential transcriptome might go unnoticed in the early differentiation phase and rather manifest in mature neuronal function or only in the context of a neuronal network. Since the RNA-seq data implies that genes involved in synaptic function and signaling are upregulated in TET3KO neurons and genes involved in proliferation are downregulated (Figure 18), TET3KO d8 neurons might be considered more mature than WT neurons at the same stage of differentiation. To unmask such possible differences in neuronal activity and signaling behavior, electrophysiological properties of WT and TET3KO neurons will be measured.

5.1.4 Specificity of genome editing

When it was discovered that there can be substantial off-target genome editing using the CRISPR-Cas9 technology (Hsu et al., 2013; Pattanayak et al., 2013), the issue of CRISPR-Cas9 specificity became probably the most heavily discussed topic in the field. The term off-target refers to any unwanted or unintended change to the gDNA sequence at positions other than targeted in the experiment. Obviously, even the possibility of uncontrolled modifications to the genome represents a major risk and hence, would hamper the application of the technology for clinical therapy.

To define the specificity of a certain sgRNA, different studies have coined the term seed region of a sgRNA. The seed region is the part of the sgRNA complementary to the protospacer fraction directly 5´of the PAM motif. In the literature the length of the seed region varies between 7 and 13 bp (Cong et al., 2013; Hsu et al., 2013; Jiang

et al., 2013; Jinek et al., 2012; Semenova et al., 2011; Sternberg et al., 2014; Wiedenheft et al., 2011). Single mismatches in this seed region completely abolish genomic cleavage by Cas9 while mismatches 5´of it, so more distal to the PAM, still allow DNA DSB formation and hence, off-target DNA editing.

Quite some efforts have been made to increase the specificity of CRISPR-Cas9 genome editing. The ways to achieve this range from the modification of Cas9 endonuclease itself over altering the sgRNA design to different delivery methods of CRISPR-Cas9 components. For example, structure-guided protein engineering was used to generate novel high-fidelity *Sp*Cas9 variants that show either reduced non-specific DNA interactions *(Sp*Cas9-HF1 (Kleinstiver et al., 2016)), weakened binding of the non-target DNA strand (e*Sp*Cas9 (Slaymaker et al., 2016)) or enhanced proofreading (HypaCas9 (Chen et al., 2017)). Additionally, a directed evolution approach in *E. coli* resulted in Sniper-Cas9 showing increased specificity and on-target activities equal to WT Cas9 (Lee et al., 2018). Moreover, the use of cooperative nickases for more precise genome editing was introduced. In this approach, the co-expression of two Cas9 nickases (D10A) together with two specific sgRNAs can generate an off-set DSB (Mali et al., 2013b; Ran et al., 2013). Since nicked genomic DNA is typically repaired seamlessly or by HDR (Cong et al., 2013), this approach exhibits an increased likelihood for successful HDR-mediated modification as well as a decreased probability of off-target editing. The latter is due to the strict necessity of two sgRNAs to bind the DNA in a certain distance to each other, which is rather unlikely to occur anywhere other than the target locus in the genome. Taking this even further, fusions of dCas9 to *Fok*I nucleases were created that can only mediate enzymatic DNA DSB formation in pair and are supposed to be even more specific than cooperative nickases (Guilinger et al., 2014; Tsai et al., 2014). Furthermore, truncating the sgRNA at its 5´end to create only 17 or 18 nucleotides of complementarity between sgRNA and protospacer was proposed to decrease the rate of off-target genome editing without sacrificing on-target editing frequencies. In addition, combining these truncated sgRNAs with the paired Cas9 variants mentioned above can further reduce off-target effects (Fu et al., 2014). Moreover, titrating the amount of Cas9 and sgRNA as well as generally reducing the amount of active Cas9 also proved very beneficial in minimizing off-target modifications. (Davis et al., 2015; Hsu et al., 2013; Zetsche et al., 2015b). By directly delivering Cas9 mRNA or even Cas9 protein instead of Cas9 encoding plasmids the presence and thereby activity of the endonuclease can be limited drastically, which in turn goes along with significantly decreased off-target effects (Kim et al., 2014; Ramakrishna et al., 2014; Zuris et al., 2015). Another way to limit the duration of Cas9 expression is the use of a “self-inactivating” expression system. This can be achieved by co-

expressing at least two sgRNAs, one of them targeting the gene of interest and the other one targeting the Cas9 gene itself. Hence, also the Cas9 expression cassette would be cut abolishing further Cas9 expression (Chen et al., 2016; Merienne et al., 2017).

For the generation of the TET3KO cell line in this work, iNGN cells were nucleofected with a previously assembled ribonucleoprotein complex of recombinant Cas9 protein and specific sgRNAs. In line with the assumption of an increased specificity of this delivery approach, the analysis of 20 predicted off-target sites (see Table 9) could not detect any changes in the DNA sequence at the respective loci. Nevertheless, to make an absolute statement about the absence of off-target modifications NGS techniques need to be utilized.

Recently, it was even shown that single Cas9-mediated DSB and subsequent repair can also cause quite extensive on-target effects like large deletions (up to several kb) and complex rearrangements that are undetectable by short range PCR assays (Kosicki et al., 2018). Such unpredictable and complex on-target editing is potentially problematic, for example when translocation events may activate dormant oncogenes and therefore, needs to be carefully monitored in any clinical trial (Lee and Kim, 2018). Also, in the TET3KO iNGN cell line, the deletion was assessed by PCR using primers that are flanking the sgRNA binding sites. Much larger deletions would not be detectable with this PCR assay. Nevertheless, the used PCR assay yielded two different products, one for each allele and hence, assuming diploidy, larger deletions can be ruled out in the TET3KO iNGN cell line. However, screening of possible off-target sites was also performed by PCR amplification and subsequent Sanger sequencing. With the used primer pairs yielding bands of around 800 bp size, large deletions on one of the alleles might go undetected.

5.2 Transactivation

5.2.1 Efficiency of Transcriptional Activation

In general, the efficiency of transactivation heavily depends on which transcriptional activation domains or combinations thereof are being guided to the promoter. Every domain functions via different mechanisms that mainly involve the recruitment of co-activators and key protein complexes involved in transcription initiation at the promoter.

In eukaryotes, all protein-coding genes are transcribed by RNA polymerase II (Pol II) (Soutourina, 2018). The initiation of transcription involves the assembly of the pre-initiation complex (PIC) containing Pol II and the so-called general transcription

factors (TF) TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. Furthermore, Mediator of Pol II transcription, short Mediator, is an essential co-activator complex that transmits regulatory signals from enhancer bound specific transcription factors to the core promoter by facilitating the recruitment and/or stability of different components of the PIC. Moreover, it regulates Pol II phosphorylation and thereby the transition from the transcription initiation to the elongation step (Soutourina, 2018).

Depending on which transcriptional activation domain was used, the degree of activation differs. VP16 was one of the first domains that was applied for transcription stimulation and thus, the underlying mechanisms have been studied extensively (Ikeda et al., 2002). VP16 was shown to recruit histone acetyltransferases (HAT) and SWItch/Sucrose Non-Fermentable (SWI/SNF) components that catalyze chromatin decondensation (Memedula and Belmont, 2003; Neely et al., 1999). Moreover, it is known that VP16 promotes PIC assembly by directly interacting with the general TFs as well as a multitude of co-activators including the Mediator complex, positive cofactor 4 (PC4) and members of the cAMP response element-binding protein (CREB)-binding protein (CBP)/p300 co-activator family (Ge and Roeder, 1994; Ikeda et al., 2002; Konermann et al., 2015; Mittler et al., 2003). The use of multiple VP16 domains, such as in VP64 (4xVP16), results in an increase of the degree of transactivation. In addition to VP64, the dCas9-VPR construct contains p65 and Rta activation domains. The NF-κB transactivating domain p65 was found to have two modes of transcriptional activation. On the one hand, direct interaction of p65 with the Mediator complex was shown to be essential for the expression of some NF-κB target genes, while on the other hand p65 recruits secondary transcription factors, including activating transcription factors (ATF)/CREB proteins, AP-1 and Sp1 to target gene promoters (van Essen et al., 2009). Also the Epstein-Barr virus transcription factor Rta was shown to form a complex with Sp1 via a direct interaction with MBD1-containing chromatin-associated factor 1 (MCAF1) (Chang et al., 2005). Sp1 is a ubiquitous zinc finger transcription factor that binds to G-rich elements such as the GC box in many promoters and enhancers. It mediates transcriptional activation by binding general transcription factors like TATA-box binding protein, which represents a subunit of TFIID (Fujita et al., 2003).

In natural context, endogenous transcription factors act in synergy with co-activators (Lemon and Tjian, 2000), so it is not surprising that the combination of several activation domains results in increased transcriptional activation (Konermann et al., 2015). This is especially efficient when distinct subsets of transcription factors and chromatin remodelers are recruited by each

transactivation domain (Konermann et al., 2015). The superiority of dCas9-VPR over dCas9-VP64 in activating *Cnga1* in 661W cells (Figure 21B-C) clearly proves the synergistic effects obtained by combining several activation domains.

By now, many different second-generation dCas9 activators are published that show robust transactivation efficiencies. Since these efficiencies depend on various factors, such as the used cell lines and sgRNAs as well as the basal expression level of the target gene, Chavez et al. set out to directly compare their performance in a single experiment (Chavez et al., 2016). They concluded that across a range of target genes and species, dCas9-VPR (Chavez et al., 2015), SunTag (Tanenbaum et al., 2014) and SAM (Konermann et al., 2015) are superior to all other tested activators. The fold-changes in gene expression achieved with those three strategies were within an order of magnitude of each other (Chavez et al., 2016). Among them, the dCas9-VPR system is the only transactivation system with a single activator component in addition to the sgRNA and thus, most suited for *in vivo* delivery using rAAVs (Lin et al., 2015).

In this study, dCas9-VPR activation was much more efficient for *Cnga1* in 661W cells than for *Opn1mw* or *Opn1sw* in MEF cells, which is most likely due to the different basal expression levels of the target genes in the used cell lines. Since in WT MEF cells already a substantial amount of M-opsin or S-opsin transcript was expressed, the fold changes in gene expression achieved by dCas9-VPR are comparably low (Figure 31B-C). With *Cnga1* not being expressed at all in 661W cells, very high fold changes of activation can be reached (Figure 28B-C). It seems to be a general phenomenon that weakly expressed genes show a high fold change in activation and, vice versa, with already higher expressed genes the reached fold change is much lower. This negative correlation between basal gene expression level and relative expression gain induced by the transcriptional activator was already reported in literature (Chavez et al., 2015; Konermann et al., 2015). They suggest that these systems might only be capable of inducing gene expression to some static upper limit that is dependent on the activator used (Chavez et al., 2016). The use of more sgRNAs was suggested as a strategy to push the expression levels to maximal induction (Chavez et al., 2016) which could be worth testing for *Opn1mw* and *Opn1sw* in the MEF setup but should be less relevant for the *in vivo* application with *Opn1mw* expression being absent in rods. Another reason for the different activation efficiencies could be that the chosen sets of *Opn1mw/sw* targeting sgRNAs were less efficient than the *Cnga1* sgRNA set. However, since testing of another three *Opn1mw* targeting sgRNAs resulted in an even lower *Opn1mw* activation efficiency (data not shown), this possibility can probably be ruled out. Nevertheless, suboptimal sgRNA efficiency could be a reason for the low activation efficiency of *Opn1sw*. The dCas9-

VPR-mediated activation of *Opn1sw* was much lower than of *Opn1mw*, although both genes show similar basal expression levels in MEF cells.

Using multiple sgRNAs to target one locus was shown to have synergistic effects (Chavez et al., 2015, 2016) and thus, a set of three sgRNAs was used per promoter in this transactivation study. These sets of sgRNAs were pairing to DNA sequences between 60 and 309 bp upstream of the transcription start site of the respective gene (Figure 21A, Figure 29A, Figure 29B). Positioning them too close or even after the TSS was avoided to rule out any hindering effects of the big dCas9-VPR complex for transcription initiation as known for early CRISPRi approaches (Qi et al., 2013).

5.2.2 Effects of dCas9-VPR levels on target gene expression

The aim of creating a stable cell line expressing *dCas9-VPR* under the control of a TRE fused to a minimal CMV promoter was to gain control over the time frame and amount of *dCas9-VPR* expression. PiggyBac transposon technology was applied, which results in random integration of the sequence encoded in the PiggyBac transposon into TTAA sites in the genome (Figure 22A). The transfected ratio of PiggyBac transposase to transposon vector determines how many molecules are integrated into the genome. Here, a ratio of 1:2.5 was used aiming at a rather low integration rate. Application of DOX to the culture medium results in the translocation of the rtTA to the TRE and subsequent expression of *dCas9-VPR* (Figure 22B). The outcome of the initial experiments using a condition without and with 500 ng/ml DOX were very surprising because a significant increase in *Cnga1* transcript levels could only be detected without DOX (Figure 22C). This means that even in the absence of DOX a leaky expression of dCas9-VPR is given. To better understand the connection between dCas9-VPR expression levels and the efficiency of *Cnga1* activation, the experiment was repeated using additional DOX concentrations ranging between no DOX and the previously used high concentration (500 ng/ml). As expected, the levels of *dCas9-VPR* expression were positively correlating with the DOX concentration (Figure 23A). In contrast, an inverse correlation between *dCas9-VPR* expression and stimulated *Cnga1* transcript levels could be detected (Figure 23B). In detail, a maximal Cnga1 activation could be achieved at 5 ng/ml DOX and increasing DOX levels were inhibiting the activation. This finding could also be replicated in MEF cells stably expressing dCas9-VPR and *Opn1mw* targeting sgRNAs (Figure 30). Although there was not such a pronounced leaky expression of dCas9-VPR without DOX, low DOX concentrations (5 ng/ml) proved to cause most efficient activation, too. Based on the observation that after incubation with high DOX concentrations cell densities were lower compared to cells exposed to no or low DOX, this could be due to adverse effects of high amounts

of dCas9-VPR. Whether high amounts of dCas9-VPR are simply toxic, inhibit growth or are associated with increased off-target gene activation is unclear. RNA-seq analysis must be performed to assess possible off-target effects on the expression of genes other than *Cnga1* or *Opn1mw*, respectively. Although observed for two different genes, it cannot be ruled out that high levels of both *Cnga1* or *Opn1mw* result in fast onset cell death.

Due to the above described effects, a concentration of 5 ng/ml was used for all further experiments. In the 661W-pb cell line the expression of Cnga1 could also be detected on protein level by immunocytochemistry (Figure 24). Labeling of dCas9-VPR using αCas9 antibody resulted in some unspecific background staining in the cytoplasm, but since a nuclear staining was only observed with DOX application, this signal in considered to be specific. Moreover, electrophysiological analysis could even show the existence of functional CNG channels in the membrane of 661W-pb cells co-expressing dCas9-VPR and sgCnga1, while CNG-specific currents could not be measured in 661W cells co-expressing a control sgRNA (sgLacZ) (Figure 25 and Figure 26). These results collectively prove the efficiency of Cnga1 transactivation using dCas9-VPR and the designed set of sgRNAs.

5.2.3 Split Cas9

To enable *in vivo* application of this validated transactivation approach, the dCas9-VPR sequence had to be split and distributed into two rAAV vectors. For subsequent reconstitution into functional FL dCas9-VPR, split intein-mediated protein *trans*-splicing was tested.

Protein splicing requires sequential nucleophilic displacement reactions (Mills et al., 2014). The stepwise mechanism begins with the assembly of the N-intein and C-intein thereby forming a catalytic core next to the extein sequence, i.e. dCas9-VPR. Proper folding of the intein attached to Cas9 fragments is essential to align the nucleophilic residues and residues that assist catalysis (Mills et al., 2014). Upon proper formation, an initial N>O/S acyl shift at the N-terminal serine or cysteine residue of the C-extein takes place. As a result of this, an ester or thioester bond, respectively, is formed between the N-terminal side-chain of the C-extein and the peptide backbone of the N-extein. In a next step, the two exteins are linked via a trans(thio)esterification resulting in a covalent bond between the C-terminus of the N-extein and the N-terminal side-chain of the C-extein. At this stage the C-extein is still connected with a peptide bond to the C-intein. So, in the subsequent step this branched intermediate is resolved by cyclization of the C-terminal asparagine of the intein that results in breakage of the peptide bond between intein and extein. As a

final step, an O/S>N acyl rearrangement leaves the exteins scarlessly linked by a peptide bond (iGEM Team Heidelberg, 2014; Mills et al., 2014).

Based on this mechanism, a certain protein cannot be simply split after every possible aa position. Several points need to be considered when deciding where to split a protein of interest. Inteins require the first (N-terminal) aa of the C-extein to contain either a hydroxyl or thiol reactive group for the self-catalytic splicing reaction, i.e. the first aa has to be a cysteine, serine or threonine (Fine et al., 2015; Shah et al., 2011). Furthermore, the tertiary structure of the protein of interest has to be analyzed in detail to find a split site that is surface-exposed due to the sterical need for protein *trans*-splicing (Schwartz et al., 2007; Truong et al., 2015). Moreover, it should be avoided to expose hydrophobic core regions to solvent by splitting as this might result in aggregate formation of the split proteins (Dill, 1985; Truong et al., 2015). In some cases, protein subdomains need to fold in a cooperative manner and random splitting would result in misfolded protein halves that cannot be reconstituted into functional protein (Lee et al., 2012; Truong et al., 2015). The publication of the crystal structure of *Sp*Cas9 (Nishimasu et al., 2014) was very helpful for identifying optimal split sites. The *Sp*Cas9 tertiary structure comprises two lobes with a positively charged and hydrophilic groove in-between and hence appears ideal for splitting (Truong et al., 2015). Chew et al. hypothesized that splitting SpCas9 in the disordered linker connecting the two lobes (after V713) would maintain protein folding for each lobe and hence allow the reconstitution of functional *Sp*Cas9 (Chew et al., 2016). In this study, the reconstitution efficiencies of *Rma* intein fused split *Sp*Cas9 with either this split site between V713-S714 or the also published E573-C574 split site (Truong et al., 2015; Zetsche et al., 2015b) were directly compared in HEK293 cells (Figure 27A). Almost 60 % of *Sp*Cas9 fragments split after V713 were reconstituted into FL Cas9, while only 33 % of E573 split Cas9 parts were successfully reconstituted (Figure 27B-C). Based on this result, V713 split dCas9-VPR rAAV vectors were cloned and validated *in vitro* (Figure 28, Figure 31). However, a study published in 2018 also demonstrates successful *in vivo* transactivation using a dual rAAV system encoding for E573 split *dCas9-VPR* (Moreno et al., 2018).

Studies on Cas9 structure revealed that the binding of the sgRNA leads to a major conformational rearrangement of the Cas9 NUC lobe respective to the REC lobe to generate a cleft for DNA binding (Jinek et al., 2014; Nishimasu et al., 2014). Based on this knowledge and the finding that there are almost no direct protein-protein interactions between the lobes (Jinek et al., 2014), the group of Jennifer A. Doudna took a different approach to splitting Cas9 (Wright et al., 2015). They rationally designed a split Cas9 by simply expressing the NUC and REC lobe as separate

polypeptides. Interestingly, providing a sgRNA is entirely sufficient to assemble the two lobes into a functional Cas9 complex that is still capable of catalyzing site-specific DNA cleavage. While the DNA cleavage efficiency of this split Cas9 variant matched that of WT Cas9 in the reaction tube, in human cell culture this split Cas9 showed reduced activity compared to WT Cas9 (Wright et al., 2015). Another study reported as well that there is auto-assembly of Cas9 fragments without fusion to dimerization domains or inteins. Zetsche and colleagues even suggest a stable low-copy expression of split Cas9 fragments for genome editing to reduce off-target editing frequencies (Zetsche et al., 2015b). Notably, these studies highlight the important role of the sgRNA in structural assembly and conformational activation of Cas9 (Wright et al., 2015). Nevertheless, split intein-mediated reconstitution of FL Cas9 is reported to be more efficient than non-covalent heterodimerization approaches (Chew et al., 2016).

With the possible application of CRISPR-Cas9 in gene therapy, the efficacy and safety for the patient have top priority. Hence, also intensive screening for possible adverse effects caused by dCas9-VPR delivery must be performed. Both, Cas9 and inteins are of bacterial origin and hence, might be immunogenic. The first report of transcriptional activation by split AAV9-Cas9-VPR in postnatal mice was published in 2016 (Chew et al., 2016). Indeed, RNA-seq analysis revealed many upregulated genes. Many of them showed an increased expression not because of off-target binding of Cas9-VPR but due to an immune response to Cas9. Nevertheless, after analyzing tissue histology, they concluded that AAV-Cas9-VPR does not evoke extensive cellular damage unlike other delivery methods (electroporation of FL Cas9) (Chew et al., 2016).

In this study, the co-transfection of cone-like cells with two rAAV vectors encoding for V713 split dCas9-VPR and *Cnga1* targeting sgRNAs resulted in robust transcriptional activation of *Cnga1* (Figure 28C). The transactivation efficiency was lower when compared to 661W cells co-transfected with FL Cas9-VPR and the same set of *Cnga1*-specific sgRNAs (Figure 21C). Also, the activation of *Opn1mw* was more efficient when FL dCas9-VPR was provided compared to split dCas9-VPR (Figure 29 and Figure 31). This is most likely because reconstitution of functional dCas9-VPR by split intein-mediated protein *trans*-splicing never reaches 100 % efficiency resulting in reduced levels of dCas9-VPR available at the promoter. Interestingly, Ma et al. also described reduced transcriptional activation by V713 split dCas9-VPR compared to the FL control in HEK293 cells, but showed transactivation by dCas9-VPR split after aa 203, 468 and 1153 to be as efficient as with the FL dCas9-VPR construct (Ma et al., 2016). In contrast, for the application in genome editing Chew et al. reported no significant difference between the mutation frequency by V713

split Cas9 or FL Cas9 (Chew et al., 2016). In general, since the co-delivery of two rAAV vectors might be less efficient than delivery of a single rAAV vector, high transduction efficiency of both fragments is crucial to obtain high genome editing or transactivation efficiencies (Fine et al., 2015).

An alternative way of delivering dCas9-VPR via rAAVs would be a dual rAAV expression of the two fragments and subsequent reconstitution by mRNA *trans*-splicing. To achieve this, a donor splice site needs to be placed 3´of the first dCas9-VPR half while 5´ of the other dCas9-VPR half a splice acceptor needs to be added. When both AAVs infect the same cell, their genomes can form intermolecular concatamers enabling the transcription of one chimeric pre-mRNA. Subsequently, the cell´s endogenous spliceosome can remove the intermediate ITR sequence resulting in one mature mRNA encoding for FL dCas9-VPR (Reich et al., 2003; Wally et al., 2012). One study published in 2018 already demonstrated successful *in vivo* therapy of Duchenne muscular dystrophy by supplying an adenine base editing Cas9 fusion construct via dual AAV delivery and mRNA *trans*-splicing (Ryu et al., 2018). Another alternative would be a reconstitution by homologous recombination if both expressed dCas9-VPR halves contain a substantial sequence overlap. Unfortunately, this approach was proven very inefficient in post-mitotic photoreceptors. Additionally, a hybrid technique combining the *trans*-splicing elements with a recombinogenic region has been established, which showed a higher reconstitution efficiency (Carvalho et al., 2017; Dyka et al., 2014; Trapani et al., 2014). The *in vivo* reconstitution efficiencies of these alternative techniques range between 5-10 % when compared to expression from a single vector (Carvalho et al., 2017). How effective a mRNA *trans*-splicing or hybrid approach works for the transactivation of *Cnga1* or *Opn1mw* by dCas9-VPR would need thorough testing.

5.2.4 Outlook

The concept of transcriptional activation of endogenous genes has several advantages over classical gene supplementation therapy and therefore, shows great potential of being used for gene therapy soon. Cas9-mediated transactivation results in target gene activation at physiologically relevant levels which represents an advantage when compared to the artificial overexpression of genes. Moreover, transactivation is more feasible for genes that are difficult to clone because of their large size or diverse splicing outcomes. Also, multiplexing can be facilitated by supplying multiple sgRNAs rather than delivering multiple expression vectors (Chavez et al., 2015).

Interestingly, sgRNAs with a truncated spacer of 14-15 nucleotides still enable efficient binding of Cas9 to the genomic locus but do not allow for DNA cleavage by

Cas9. Hence, depending on sgRNA spacer length catalytically active Cas9 fused to VPR can be used either for gene editing or transcriptional activation (Chew et al., 2016; Dahlman et al., 2015; Kiani et al., 2015). The application of such a multi-functional Cas9 enzyme could be of special interest when it comes the therapy of inherited retinal degenerative diseases caused by dominant-negative mutations. In detail, co-delivery of Cas9-VPR with a standard *Rho* targeting sgRNA and truncated sgRNAs specific to the *Opn1mw* promoter could result in the simultaneous ablation of mutated rhodopsin and the compensatory transactivation of M-opsin.

The *in vitro* results presented in this study are very promising and the next steps involve the *in vivo* testing of the *Cnga1* and *Opn1mw* transactivation approach. At first, rAAV vectors expressing split dCas9-VPR under control of the cone-specific murine short wavelength opsin (mSWS) promoter and the tested *Cnga1* specific sgRNAs as well as rAAV vectors encoding for split dCas9-VPR under the rod-specific human rhodopsin (hRHO) promoter and the *Opn1mw* specific sgRNAs were cloned and virus will be produced. It is planned to subretinally inject these virus suspensions into WT mice as well as suitable mouse models for ACHM and RP, respectively. Several weeks after the injection into the eyes of WT mice the expression levels of either Cnga1 or Opn1mw must be analyzed by qRT-PCR and immunohistochemistry. If a successful activation of both genes could be achieved *in vivo*, the therapeutic potential of the transactivation must be tested in the respective mouse models of retinal degeneration like Cnga3$^{-/-}$ and Rho$^{-/-}$ mice. Several weeks or months after the subretinal injection the degree of photoreceptor degeneration can be inferred from the thickness of the outer nuclear layer of the retina, which can be measured by optical coherence tomography. The functional readout would comprise electrophysiological measurement of light-driven signals in the retina and behavioral tests, e.g. light-dark-test, that evaluate if the treatment resulted in improved vision.

6 Appendix

6.1 Primer sequences

Table 10 Sequences of primer pairs used for off-target analysis

Off-target locus	**Primer pair sequences 5´-3´**
intron: lncRNA LOC102724623	fw AGCTGCTATGATTCTCCTGGC rev TGCTCCTCAAACAAGCTACCA
intron: IQCM	fw TGAAGTATCAGAGTCATGTCTGTTT rev CAGGGAAAGGCAACCACAGA
intron: lncRNA LOC107984703	fw AGCAAACCTAGGTGAAGCCC rev ACCCTGGATTGCCCATAAGC
exon: SREK1	fw GACGTGGAGAGTGCATGTTTG rev AGCATACGTTGCTGAAAGGC
exon: ARHGEF38	fw TGTGAGATAACACGTGGCCC rev GCTACAGGATCAGACGGGTG
intergenic: LOC105374124-PPP2R3A	fw CCCGGCCAATTTTTAAAGCACTA rev AAGGGATGCTCTACATCGTTAGTC
lncRNA LOC107987207	fw AATCCTCGGTCAAGCCCTCA rev AAGATGCTCTAGCACACACCC
intergenic: PBX3-LOC105376273	fw GATGACATCCCGGGGAGTTG rev CCCACCATACTCCAACTCAGG
exon: MCM8	fw GCACTTCAAACATGGGCTGTTA rev TTAACACATGGCTACCACTACA
exon: MAML2	fw AGCCCTGGAACTTTGGTTGG rev ACGGGACAAGAATGCCATCA
intron: SCGB1B2P	fw AAGCACAGGCAACAAAAGCA rev GCCTGGATGTGTACCCAGTAA
intergenic: LOC100131241-CASC17	fw CCACTAGGACCTTAGGATGGC rev CCGGATACCTCCGTGACCA
intron: LOC105376637	fw ACACCAGTATTCCGTTCCTTT rev GTCCTTTGCAGGGACATGGA
exon: GUCY2C/PLBD1-AS1	fw TTTCAGGACACTTGAGGTCGC rev CTTGGAGTGTTCCAGGTCCA
exon:ZSWIM8/ZSWIM8-AS1/NDST2	fw TGCGGACTACGACGACTTTG rev CTCTTCATCCGCAGCAGTGT
intergenic: LOC105374959-LOC100506885	fw AACCACTTGGGAATCACGGA rev GCAGAGCAGCTTTCTAACTGG

intergenic: LOC101927709-EN1	fw CAGGGTCAACTGCTGGAAGA rev GACAGCTCTCAGCAGCAACA
intergenic: LOC105377321-SPARCL1	fw GGGCATGTTCTCTCAGTGACTA rev AAGAGGTGCTCATTAATGTCCTTCA
exon: NISCH	fw GGGTTAGGTGCCAGGATGAT rev GCGGATAGAGCGGAGACATC
exon: DNTT	fw ACAGGCAGGCCGTCAATTAG rev CAGACACAGTGCTGAGGCAT

Table 11 Sequences of primer pairs used for pyrosequencing assay

F and R represent the forward and reverse primer, respectively, used for locus-specific PCR amplification, while S marks the sequencing primer for the pyrosequencing run. All reverse primers contain a 5´ Biotin [BIO] tag.

Target gene	Assay#	Primer name	Sequence 5´-3´
FAM122C	1	S1	GGTAGGATATGGGGATAAA
	2	S2	TTGAAGTGAAGTTATTGTTTGTAT
	1_2	F1	AGGAAGGAGGGTAGGATAT
		R1	[BIO]-TCTCTCAACCCTAACAACTCATAAT
PCDHA6	1	F1	GTTGAATTGGTTGATATTAATGTTGTAAA
		S1	TTATTTTAAATTAATAAAAGAAGT
	2	S1	TGTTTTAAAAATATATGGTATAG
	3	S2	ATTGTATTATTTTTTGATTGTTGA
	2_3	F2	GGAGGGTTTATTATAAAAGTAGAGAAAAGA
	1_2_3	R1	[BIO]-AAATAAACACCATATCAAATACCTCTAAT

Table 12 Sequences of primer pairs used for qRT-PCR

	Primer name	**Nucleotide sequence 5´-3´**
Alas (housekeeping)	mAlas fw mAlas rev	TCGCCGATGCCCATTCTTATC GGCCCCAACTTCCATCATCT
ACTB (housekeeping)	hACTB fw hACTB rev	GCCGCCAGCTCACCAT CACGATGGAGGGGAAGACG
Cnga1	mCnga1 fw mCnga1rev	AACGAGCCATTTGTGCTGC TGGTTAGTTTAATATCTGCGCTTGT
Opn1mw	mOpn1mw fw mOpn1mw rev	GGAGCAGGTACTGGCCTTATG GGAGGTAGCAGAGCACGATG

Opn1sw	mOpn1sw fw mOpn1sw rev	ACAAAAAGTTGCGACAGCCC CCATCCTGTCACTAGACCTGC
Cas9	Cas9 fw Cas9 rev	AGTACAAGGTGCCGAGCAAA CCGTGCTGTTCTTTTGAGCC
TET1	hTET1 fw hTET1 rev	GCTCTCATGGGTGTCCAATTGCT ATGAGCACCACCATCACAGCAG
TET2	hTET2 fw hTET2 rev	AAGGCTGAGGGACGAGAACGA TGAGCCCATCTCCTGCTTCCA
TET3S	hTET3 short Exon 1 fw hTET3 Exon 3 rev	CGATGCACCAGAGGAGGC CATGGTACACTGGCCCTGAG
TET3FL	hTET3CXXC Exon 1 fw hTET3CXXC Exon 2 rev	AAACTGCGAAAATGTGAGGTGC TTGACAGCCGCTCCTTGTCC
Total *TET3*	hTET3 Exon10_11 fw hTET3 Exon 11 rev	GACCCAGGCCTGTCTCTGA ACCGAGTAGCTCTCCACCAC
TET3 Ex7/8	hTET3 fw_VSp hTET3 rev_VSp	GCAAGACACCTCGCAAGTTC CCTCGTTGGTCACCTGGTTC
FAM122C	hFAM122C fw hFAM122C rev	TAGGTTTCAAGTCGCTGCCG AACATGTCAGCTTGCAACACC
PCDHA6	hPCDHA6 fw hPCDHA6 rev	ATTATGATGGGTAAGGCGGAGAA GATACTGTTGGCCACTGCTGA
RAET1L	hRAET1L fw hRAET1L rev	TTCATCTTCCAGGATCCACCTT GTGAGGGTCGTCTCGCC
ZNF577	hZNF577 fw hZNF577 rev	GGAGGAGTGGCAGTTTTTGG CCTGGACAGATTTGACTGTGGG

7 References

Aapola, U., Kawasaki, K., Scott, H.S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Kawasaki, K., Minoshima, S., Krohn, K., et al. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. Genomics *65*, 293–298.

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science *353*, aaf5573.

Abudayyeh, O.O., Gootenberg, J.S., Essletzbichler, P., Han, S., Joung, J., Belanto, J.J., Verdine, V., Cox, D.B.T., Kellner, M.J., Regev, A., et al. (2017). RNA targeting with CRISPR-Cas13. Nature *550*, 280–284.

Agne, M., Blank, I., Emhardt, A.J., Gäbelein, C.G., Gawlas, F., Gillich, N., Gonschorek, P., Juretschke, T.J., Krämer, S.D., Louis, N., et al. (2014). Modularized CRISPR/dCas9 effector toolkit for target-specific gene regulation. ACS Synth. Biol. *3*, 986–989.

Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. Proc. Natl. Acad. Sci. U. S. A. *51*, 786–794.

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

Arnould, S., Chames, P., Perez, C., Lacroix, E., Duclert, A., Epinat, J.-C., Stricher, F., Petit, A.-S., Patin, A., Guillier, S., et al. (2006). Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. J. Mol. Biol. *355*, 443–458.

Ashworth, J., Havranek, J.J., Duarte, C.M., Sussman, D., Monnat, R.J., Stoddard, B.L., and Baker, D. (2006). Computational redesign of endonuclease DNA binding and cleavage specificity. Nature *441*, 656–659.

Ashworth, J., Taylor, G.K., Havranek, J.J., Quadri, S.A., Stoddard, B.L., and Baker, D. (2010). Computational reprogramming of homing endonuclease specificity at multiple adjacent base pairs. Nucleic Acids Res. *38*, 5601–5608.

Assenov, Y., Müller, F., Lutsik, P., Walter, J., Lengauer, T., and Bock, C. (2014). Comprehensive analysis of DNA methylation data with RnBeads. Nat. Methods *11*, 1138–1140.

Balboa, D., Weltner, J., Eurola, S., Trokovic, R., Wartiovaara, K., and Otonkoski, T. (2015). Conditionally Stabilized dCas9 Activator for Controlling Gene Expression in Human Cell Reprogramming and Differentiation. Stem Cell Rep. *5*, 448–459.

Banin, E., Gootwine, E., Obolensky, A., Ezra-Elia, R., Ejzenberg, A., Zelinger, L., Honig, H., Rosov, A., Yamin, E., Sharon, D., et al. (2015). Gene Augmentation Therapy

Restores Retinal Function and Visual Behavior in a Sheep Model of CNGA3 Achromatopsia. Mol. Ther. J. Am. Soc. Gene Ther. *23*, 1423–1433.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Bareil, C., Hamel, C.P., Delague, V., Arnaud, B., Demaille, J., and Claustres, M. (2001). Segregation of a mutation in CNGB1 encoding the beta-subunit of the rod cGMP-gated channel in a family with autosomal recessive retinitis pigmentosa. Hum. Genet. *108*, 328–334.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.

Basner-Tschakarjan, E., and Mingozzi, F. (2014). Cell-Mediated Immunity to AAV Vectors, Evolving Concepts and Potential Solutions. Front. Immunol. *5*, 350.

Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.-L. (2013). Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep. *4*, 220–228.

Becker, P.B., Ruppert, S., and Schütz, G. (1987). Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. Cell *51*, 435–443.

Beerli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F. (1998). 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. U. S. A. *95*, 14628–14633.

Belfort, M., and Bonocora, R.P. (2014). Homing Endonucleases: From Genetic Anomalies to Programmable Genomic Clippers. Methods Mol. Biol. Clifton NJ *1123*, 1–26.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes Dev. *23*, 781–783.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA Particles in Living Yeast. Mol. Cell *2*, 437–445.

Bestor, T.H. (2000). The DNA methyltransferases of mammals. Hum. Mol. Genet. *9*, 2395–2402.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J. Mol. Biol. *203*, 971–983.

Biel, M., and Michalakis, S. (2007). Function and dysfunction of CNG channels: insights from channelopathies and mouse models. Mol. Neurobiol. *35*, 266–277.

Biel, M., Seeliger, M., Pfeifer, A., Kohler, K., Gerstner, A., Ludwig, A., Jaissle, G., Fauser, S., Zrenner, E., and Hofmann, F. (1999). Selective loss of cone function in mice lacking

the cyclic nucleotide-gated channel CNG3. Proc. Natl. Acad. Sci. U. S. A. *96*, 7553–7557.

Bird, A. (1992). The essentials of DNA methylation. Cell *70*, 5–8.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. Science *326*, 1509–1512.

Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W., and Balasubramanian, S. (2012). Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science *336*, 934–937.

Booth, M.J., Ost, T.W.B., Beraldi, D., Bell, N.M., Branco, M.R., Reik, W., and Balasubramanian, S. (2013). Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. Nat. Protoc. *8*, 1841–1851.

Bourc'his, D., Xu, G.-L., Lin, C.-S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the Establishment of Maternal Genomic Imprints. Science *294*, 2536–2539.

Boyes, J., and Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell *64*, 1123–1134.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254.

Brogna, S., and Wen, J. (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. Nat. Struct. Mol. Biol. *16*, 107–113.

Buschhausen, G., Wittig, B., Graessmann, M., and Graessmann, A. (1987). Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. Proc. Natl. Acad. Sci. *84*, 1177–1181.

Busskamp, V., Lewis, N.E., Guye, P., Ng, A.H.M., Shipman, S.L., Byrne, S.M., Sanjana, N.E., Murn, J., Li, Y., Li, S., et al. (2014). Rapid neurogenesis through transcriptional activation in human stem cells. Mol. Syst. Biol. *10*, 760.

Carvalho, L.S., Xu, J., Pearson, R.A., Smith, A.J., Bainbridge, J.W., Morris, L.M., Fliesler, S.J., Ding, X.-Q., and Ali, R.R. (2011). Long-term and age-dependent restoration of visual function in a mouse model of CNGB3-associated achromatopsia following gene therapy. Hum. Mol. Genet. *20*, 3161–3175.

Carvalho, L.S., Turunen, H.T., Wassmer, S.J., Luna-Velez, M.V., Xiao, R., Bennett, J., and Vandenberghe, L.H. (2017). Evaluating Efficiencies of Dual AAV Approaches for Retinal Targeting. Front. Neurosci. *11*, 503.

Cedar, H. (1988). DNA methylation and gene activity. Cell *53*, 3–4.

Chakraborty, S., Ji, H., Kabadi, A.M., Gersbach, C.A., Christoforou, N., and Leong, K.W. (2014). A CRISPR/Cas9-based system for reprogramming cell lineage specification. Stem Cell Rep. *3*, 940–947.

Chang, H.H.Y., Pannunzio, N.R., Adachi, N., and Lieber, M.R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat. Rev. Mol. Cell Biol. *18*, 495–506.

Chang, L.-K., Chung, J.-Y., Hong, Y.-R., Ichimura, T., Nakao, M., and Liu, S.-T. (2005). Activation of Sp1-mediated transcription by Rta of Epstein-Barr virus via an interaction with MCAF1. Nucleic Acids Res. *33*, 6528–6539.

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.-W., and Xi, J.J. (2013). Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell Res. *23*, 465–472.

Charpentier, E., and Marraffini, L.A. (2014). Harnessing CRISPR-Cas9 immunity for genetic engineering. Curr. Opin. Microbiol. *19*, 114–119.

Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. Nat. Methods *12*, 326–328.

Chavez, A., Tuttle, M., Pruitt, B.W., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., Haque, S.J., Cecchi, R.J., Kowal, E.J.K., Buchthal, J., et al. (2016). Comparative Analysis of Cas9 Activators Across Multiple Species. Nat. Methods *13*, 563–567.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell *155*, 1479–1491.

Chen, J.S., Dagdas, Y.S., Kleinstiver, B.P., Welch, M.M., Sousa, A.A., Harrington, L.B., Sternberg, S.H., Joung, J.K., Yildiz, A., and Doudna, J.A. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. Nature *550*, 407–410.

Chen, L., MacMillan, A.M., Chang, W., Ezaz-Nikpay, K., Lane, W.S., and Verdine, G.L. (1991). Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. Biochemistry *30*, 11018–11025.

Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S.C., Falck, J.R., Peng, J., Gu, W., and Zhao, Y. (2007). Lysine propionylation and butyrylation are novel post-translational modifications in histones. Mol. Cell. Proteomics MCP *6*, 812–819.

Chen, Y., Liu, X., Zhang, Y., Wang, H., Ying, H., Liu, M., Li, D., Lui, K.O., and Ding, Q. (2016). A Self-restricted CRISPR System to Reduce Off-target Effects. Mol. Ther. J. Am. Soc. Gene Ther. *24*, 1508–1510.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res. *23*, 1163–1171.

Chew, W.L., Tabebordbar, M., Cheng, J.K.W., Mali, P., Wu, E.Y., Ng, A.H.M., Zhu, K., Wagers, A.J., and Church, G.M. (2016). A multi-functional AAV-CRISPR-Cas9 and its host response. Nat. Methods *13*, 868–874.

Choe, H.-W., Kim, Y.J., Park, J.H., Morizumi, T., Pai, E.F., Krauß, N., Hofmann, K.P., Scheerer, P., and Ernst, O.P. (2011). Crystal structure of metarhodopsin II. Nature *471*, 651–655.

Choudhury, S.R., Cui, Y., Lubecka, K., Stefanska, B., and Irudayaraj, J. (2016). CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. Oncotarget *7*, 46545–46556.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. Genetics *186*, 757–761.

Colquitt, B.M., Allen, W.E., Barnea, G., and Lomvardas, S. (2013). Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. Proc. Natl. Acad. Sci. U. S. A. *110*, 14682–14687.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Cosloy, S.D., and Oishi, M. (1973). Genetic transformation in Escherichia coli K12. Proc. Natl. Acad. Sci. U. S. A. *70*, 84–87.

Dahlman, J.E., Abudayyeh, O.O., Joung, J., Gootenberg, J.S., Zhang, F., and Konermann, S. (2015). Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. Nat. Biotechnol. *33*, 1159–1161.

Davies, B., Davies, G., Preece, C., Puliyadi, R., Szumska, D., and Bhattacharya, S. (2013). Site specific mutation of the Zic2 locus by microinjection of TALEN mRNA in mouse CD1, C3H and C57BL/6J oocytes. PloS One *8*, e60216.

Davis, K.M., Pattanayak, V., Thompson, D.B., Zuris, J.A., and Liu, D.R. (2015). Small molecule-triggered Cas9 protein with improved genome-editing specificity. Nat. Chem. Biol. *11*, 316–318.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature *471*, 602–607.

Deng, W.-T., Sakurai, K., Liu, J., Dinculescu, A., Li, J., Pang, J., Min, S.-H., Chiodo, V.A., Boye, S.L., Chang, B., et al. (2009). Functional interchangeability of rod and cone transducin alpha-subunits. Proc. Natl. Acad. Sci. U. S. A. *106*, 17681–17686.

Deng, W.-T., Sakurai, K., Kolandaivelu, S., Kolesnikov, A.V., Dinculescu, A., Li, J., Zhu, P., Liu, X., Pang, J., Chiodo, V.A., et al. (2013). Cone phosphodiesterase-6α' restores rod function and confers distinct physiological properties in the rod phosphodiesterase-6β-deficient rd10 mouse. J. Neurosci. Off. J. Soc. Neurosci. *33*, 11745–11753.

DeWitt, M.A., Magis, W., Bray, N.L., Wang, T., Berman, J.R., Urbinati, F., Heo, S.-J., Mitros, T., Muñoz, D.P., Boffelli, D., et al. (2016). Selection-free genome editing of the

sickle mutation in human adult hematopoietic stem/progenitor cells. Sci. Transl. Med. *8*, 360ra134.

Dill, K.A. (1985). Theory for the folding and stability of globular proteins. Biochemistry *24*, 1501–1509.

Ding, Q., Strong, A., Patel, K.M., Ng, S.-L., Gosis, B.S., Regan, S.N., Cowan, C.A., Rader, D.J., and Musunuru, K. (2014). Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. Circ. Res. *115*, 488–492.

Ding, X.-Q., Harry, C.S., Umino, Y., Matveev, A.V., Fliesler, S.J., and Barlow, R.B. (2009). Impaired cone function and cone degeneration resulting from CNGB3 deficiency: down-regulation of CNGA3 biosynthesis as a potential mechanism. Hum. Mol. Genet. *18*, 4770–4780.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. *34*, 184–191.

Doerfler, W. (1983). DNA methylation and gene activity. Annu. Rev. Biochem. *52*, 93–124.

Dreos, R., Ambrosini, G., Périer, R.C., and Bucher, P. (2015). The Eukaryotic Promoter Database: expansion of EPDnew and new promoter analysis tools. Nucleic Acids Res. *43*, D92-96.

Dreos, R., Ambrosini, G., Groux, R., Cavin Périer, R., and Bucher, P. (2017). The eukaryotic promoter database in its 30th year: focus on non-vertebrate organisms. Nucleic Acids Res. *45*, D51–D55.

Dryja, T.P., McGee, T.L., Reichel, E., Hahn, L.B., Cowley, G.S., Yandell, D.W., Sandberg, M.A., and Berson, E.L. (1990). A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. Nature *343*, 364–366.

Dryja, T.P., Finn, J.T., Peng, Y.W., McGee, T.L., Berson, E.L., and Yau, K.W. (1995). Mutations in the gene encoding the alpha subunit of the rod cGMP-gated channel in autosomal recessive retinitis pigmentosa. Proc. Natl. Acad. Sci. U. S. A. *92*, 10177–10181.

Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W.A., Hou, L., and Lin, S.M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics *11*, 587.

Dyka, F.M., Boye, S.L., Chiodo, V.A., Hauswirth, W.W., and Boye, S.E. (2014). Dual adeno-associated virus vectors result in efficient in vitro and in vivo expression of an oversized gene, MYO7A. Hum. Gene Ther. Methods *25*, 166–177.

Dzeja, C., Hagen, V., Kaupp, U.B., and Frings, S. (1999). Ca2+ permeation in cyclic nucleotide-gated channels. EMBO J. *18*, 131–144.

East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H.D., Tjian, R., and Doudna, J.A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature *538*, 270–273.

Eden, E., Lipson, D., Yogev, S., and Yakhini, Z. (2007). Discovering motifs in ranked lists of DNA sequences. PLoS Comput. Biol. *3*, e39.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

van Essen, D., Engist, B., Natoli, G., and Saccani, S. (2009). Two modes of transcriptional activation at native promoters by NF-kappaB p65. PLoS Biol. *7*, e73.

Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., and Church, G.M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat. Methods *10*, 1116–1121.

Fesenko, E.E., Kolesnikov, S.S., and Lyubarsky, A.L. (1985). Induction by cyclic GMP of cationic conductance in plasma membrane of retinal rod outer segment. Nature *313*, 310–313.

Fine, E.J., Appleton, C.M., White, D.E., Brown, M.T., Deshmukh, H., Kemp, M.L., and Bao, G. (2015). Trans-spliced Cas9 allows cleavage of HBB and CCR5 genes in human cells using compact expression cassettes. Sci. Rep. *5*, 10777.

Firth, A.L., Menon, T., Parker, G.S., Qualls, S.J., Lewis, B.M., Ke, E., Dargitz, C.T., Wright, R., Khanna, A., Gage, F.H., et al. (2015). Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs. Cell Rep. *12*, 1385–1390.

Fischer, M.D., Wilhelm, B., Michalakis, S., Zobor, D., Kohl, S., Seeliger, M., Zrenner, E., Ueffing, M., Wissinger, B., Biel, M., et al. (2016). Successful delivery of rAAV8.CNGA3 in a patient with CNGA3 achromatopsia. Invest. Ophthalmol. Vis. Sci. *57*.

Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiácovo, M.P., Church, G.M., and Calarco, J.A. (2013). Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nat. Methods *10*, 741–743.

Frings, S., Seifert, R., Godde, M., and Kaupp, U.B. (1995). Profoundly different calcium permeation and blockage determine the specific function of distinct cyclic nucleotide-gated channels. Neuron *15*, 169–179.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. *32*, 279–284.

Fujita, N., Watanabe, S., Ichimura, T., Ohkuma, Y., Chiba, T., Saya, H., and Nakao, M. (2003). MCAF mediates MBD1-dependent transcriptional repression. Mol. Cell. Biol. *23*, 2834–2843.

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas

bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67–71.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U. S. A. *109*, E2579-2586.

Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature *551*, 464–471.

Ge, H., and Roeder, R.G. (1994). Purification, cloning, and characterization of a human coactivator, PC4, that mediates transcriptional activation of class II genes. Cell *78*, 513–523.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442–451.

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

Guilinger, J.P., Thompson, D.B., and Liu, D.R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat. Biotechnol. *32*, 577–582.

Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., et al. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol. *17*, 148.

Hahn, M.A., Qiu, R., Wu, X., Li, A.X., Zhang, H., Wang, J., Jui, J., Jin, S.-G., Jiang, Y., Pfeifer, G.P., et al. (2013). Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. Cell Rep. *3*, 291–300.

Hai, T., Teng, F., Guo, R., Li, W., and Zhou, Q. (2014). One-step generation of knockout pigs by zygote injection of CRISPR/Cas system. Cell Res. *24*, 372–375.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development *129*, 1983–1993.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303–1307.

Hecht, S. (1930). The Development of Thomas Young's Theory of Color Vision. JOSA *20*, 231–270.

Heidenreich, M., and Zhang, F. (2016). Applications of CRISPR-Cas systems in neuroscience. Nat. Rev. Neurosci. *17*, 36–44.

Heitz, E. (1928). Das Heterochromatin der Moose (Jahrb Wiss Botanik).

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. *33*, 510–517.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat. Biotechnol. *29*, 731–734.

Holliday, R., and Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. Science *187*, 226–232.

Holtzman, L., and Gersbach, C.A. (2018). Editing the Epigenome: Reshaping the Genomic Landscape. Annu. Rev. Genomics Hum. Genet.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262–1278.

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

Huang, Y.-H., Su, J., Lei, Y., Brunetti, L., Gundry, M.C., Zhang, X., Jeong, M., Li, W., and Goodell, M.A. (2017). DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A. Genome Biol. *18*, 176.

Humphries, M.M., Rancourt, D., Farrar, G.J., Kenna, P., Hazel, M., Bush, R.A., Sieving, P.A., Sheils, D.M., McNally, N., Creighton, P., et al. (1997). Retinopathy induced in mice by targeted disruption of the rhodopsin gene. Nat. Genet. *15*, 216–219.

Hüttl, S., Michalakis, S., Seeliger, M., Luo, D.-G., Acar, N., Geiger, H., Hudl, K., Mader, R., Haverkamp, S., Moser, M., et al. (2005). Impaired channel targeting and retinal degeneration in mice lacking the cyclic nucleotide-gated channel subunit CNGB1. J. Neurosci. Off. J. Soc. Neurosci. *25*, 130–138.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. *31*, 227–229.

iGEM Team Heidelberg (2014). A detailed description of the trans-splicing reaction. http://2014.igem.org/Team:Heidelberg/Project/Background.

Ikeda, K., Stuehler, T., and Meisterernst, M. (2002). The H1 and H2 regions of the activation domain of herpes simplex virion protein 16 stimulate transcription through distinct molecular mechanisms. Genes Cells Devoted Mol. Cell. Mech. *7*, 49–58.

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

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

Ivanov, E.L., Sugawara, N., Fishman-Lobell, J., and Haber, J.E. (1996). Genetic requirements for the single-strand annealing pathway of double-strand break repair in Saccharomyces cerevisiae. Genetics *142*, 693–704.

Iwan, K., Rahimoff, R., Kirchner, A., Spada, F., Schröder, A.S., Kosmatchev, O., Ferizaj, S., Steinbacher, J., Parsa, E., Müller, M., et al. (2018). 5-Formylcytosine to cytosine conversion by C-C bond cleavage in vivo. Nat. Chem. Biol. *14*, 72–78.

Jacobs, G.H. (2018). Photopigments and the dimensionality of animal color vision. Neurosci. Biobehav. Rev. *86*, 108–130.

Jacobs, G.H., Neitz, J., and Deegan, J.F. (1991). Retinal receptors in rodents maximally sensitive to ultraviolet light. Nature *353*, 655–656.

Jacoby, K., Metzger, M., Shen, B.W., Certo, M.T., Jarjour, J., Stoddard, B.L., and Scharenberg, A.M. (2012). Expanding LAGLIDADG endonuclease scaffold diversity by rapidly surveying evolutionary sequence space. Nucleic Acids Res. *40*, 4954–4964.

Jacquier, A., and Dujon, B. (1985). An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell *41*, 383–394.

Jang, H.S., Shin, W.J., Lee, J.E., and Do, J.T. (2017). CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. Genes *8*.

Jao, L.-E., Wente, S.R., and Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc. Natl. Acad. Sci. U. S. A. *110*, 13904–13909.

Jat, P.S., Cepko, C.L., Mulligan, R.C., and Sharp, P.A. (1986). Recombinant retroviruses encoding simian virus 40 large T antigen and polyomavirus large and middle T antigens. Mol. Cell. Biol. *6*, 1204–1217.

Jiang, F., and Doudna, J.A. (2017). CRISPR-Cas9 Structures and Mechanisms. Annu. Rev. Biophys. *46*, 505–529.

Jiang, D.J., Xu, C.L., and Tsang, S.H. (2018). Revolution in Gene Medicine Therapy and Genome Surgery. Genes *9*.

Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). CRISPR-assisted editing of bacterial genomes. Nat. Biotechnol. *31*, 233–239.

Jin, Z., and Liu, Y. (2018). DNA methylation in human diseases. Genes Dis. *5*, 1–8.

Jin, S.-G., Zhang, Z.-M., Dunwell, T.L., Harter, M.R., Wu, X., Johnson, J., Li, Z., Liu, J., Szabó, P.E., Lu, Q., et al. (2016). Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. Cell Rep. *14*, 493–505.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science *343*, 1247997.

Johnson, T.B., and Coghill, R.D. (1925). Researches on pyrimidines. C111. The discovery of 5-methyl-cytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus1. J. Am. Chem. Soc. *47*, 2838–2844.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat. Genet. *19*, 187–191.

Kaeppel, C., Beattie, S.G., Fronza, R., van Logtenstein, R., Salmon, F., Schmidt, S., Wolf, S., Nowrouzi, A., Glimm, H., von Kalle, C., et al. (2013). A largely random AAV integration profile after LPLD gene therapy. Nat. Med. *19*, 889–891.

Kareta, M.S., Botello, Z.M., Ennis, J.J., Chou, C., and Chédin, F. (2006). Reconstitution and Mechanism of the Stimulation of de Novo Methylation by Human DNMT3L. J. Biol. Chem. *281*, 25893–25902.

Kaupp, U.B., and Seifert, R. (2002). Cyclic nucleotide-gated ion channels. Physiol. Rev. *82*, 769–824.

Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nat. Methods *12*, 401–403.

Kebede, A.F., Schneider, R., and Daujat, S. (2015). Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. FEBS J. *282*, 1658–1674.

Khare, T., Pai, S., Koncevicius, K., Pal, M., Kriukiene, E., Liutkeviciute, Z., Irimia, M., Jia, P., Ptak, C., Xia, M., et al. (2012). 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. Nat. Struct. Mol. Biol. *19*, 1037–1043.

Kiani, S., Chavez, A., Tuttle, M., Hall, R.N., Chari, R., Ter-Ovanesyan, D., Qian, J., Pruitt, B.W., Beal, J., Vora, S., et al. (2015). Cas9 gRNA engineering for genome editing, activation and repression. Nat. Methods *12*, 1051–1054.

Kim, E., Koo, T., Park, S.W., Kim, D., Kim, K., Cho, H.-Y., Song, D.W., Lee, K.J., Jung, M.H., Kim, S., et al. (2017). *In vivo* genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. Nat. Commun. *8*, 14500.

Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. *24*, 1012–1019.

Kim, Y.G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc. Natl. Acad. Sci. U. S. A. *93*, 1156–1160.

Kingston, R.E., Chen, C.A., and Rose, J.K. (2003). Calcium phosphate transfection. Curr. Protoc. Mol. Biol. *Chapter 9*, Unit 9.1.

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature *529*, 490–495.

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

Kohl, S., Marx, T., Giddings, I., Jägle, H., Jacobson, S.G., Apfelstedt-Sylla, E., Zrenner, E., Sharpe, L.T., and Wissinger, B. (1998). Total colourblindness is caused by mutations in the gene encoding the alpha-subunit of the cone photoreceptor cGMP-gated cation channel. Nat. Genet. *19*, 257–259.

Kohl, S., Baumann, B., Broghammer, M., Jägle, H., Sieving, P., Kellner, U., Spegal, R., Anastasi, M., Zrenner, E., Sharpe, L.T., et al. (2000). Mutations in the CNGB3 gene encoding the beta-subunit of the cone photoreceptor cGMP-gated channel are responsible for achromatopsia (ACHM3) linked to chromosome 8q21. Hum. Mol. Genet. *9*, 2107–2116.

Kohli, R.M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. Nature *502*, 472–479.

Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M.D.C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat. Biotechnol. *32*, 267–273.

Komáromy, A.M., Alexander, J.J., Rowlan, J.S., Garcia, M.M., Chiodo, V.A., Kaya, A., Tanaka, J.C., Acland, G.M., Hauswirth, W.W., and Aguirre, G.D. (2010). Gene therapy rescues cone function in congenital achromatopsia. Hum. Mol. Genet. *19*, 2581–2593.

Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature *533*, 420–424.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583–588.

Konermann, S., Lotfy, P., Brideau, N.J., Oki, J., Shokhirev, M.N., and Hsu, P.D. (2018). Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. Cell *173*, 665-676.e14.

Koonin, E.V., Makarova, K.S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. *37*, 67–78.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868–871.

Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. Nat. Biotechnol. *36*, 765.

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

Labun, K., Montague, T.G., Gagnon, J.A., Thyme, S.B., and Valen, E. (2016). CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res. *44*, W272-276.

Lam, R.S., Töpfer, F.M., Wood, P.G., Busskamp, V., and Bamberg, E. (2017). Functional Maturation of Human Stem Cell-Derived Neurons in Long-Term Cultures. PloS One *12*, e0169506.

Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat. Protoc. *8*, 2180–2196.

Lee, H., and Kim, J.-S. (2018). Unexpected CRISPR on-target effects. Nat. Biotechnol. *36*, 703–704.

Lee, J.K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y.-H., Lee, K., Jung, I., Kim, D., Kim, S., et al. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. Nat. Commun. *9*, 3048.

Lee, Y.-T., Su, T.-H., Lo, W.-C., Lyu, P.-C., and Sue, S.-C. (2012). Circular permutation prediction reveals a viable backbone disconnection for split proteins: an approach in identifying a new functional split intein. PloS One *7*, e43820.

Lem, J., Krasnoperova, N.V., Calvert, P.D., Kosaras, B., Cameron, D.A., Nicolò, M., Makino, C.L., and Sidman, R.L. (1999). Morphological, physiological, and biochemical changes in rhodopsin knockout mice. Proc. Natl. Acad. Sci. U. S. A. *96*, 736–741.

Lemon, B., and Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. Genes Dev. *14*, 2551–2569.

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

Li, H., Ulge, U.Y., Hovde, B.T., Doyle, L.A., and Monnat, R.J. (2012). Comprehensive homing endonuclease target site specificity profiling reveals evolutionary constraints and enables genome engineering applications. Nucleic Acids Res. *40*, 2587–2598.

Li, J., Sun, W., Wang, B., Xiao, X., and Liu, X.-Q. (2008). Protein trans-splicing as a means for viral vector-mediated in vivo gene therapy. Hum. Gene Ther. *19*, 958–964.

Li, M., Zhou, X., Wang, S., Michailidis, I., Gong, Y., Su, D., Li, H., Li, X., and Yang, J. (2017). Structure of a eukaryotic cyclic-nucleotide-gated channel. Nature *542*, 60–65.

Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P., and Yang, B. (2011). TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res. *39*, 359–372.

Li, W., Teng, F., Li, T., and Zhou, Q. (2013). Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. Nat. Biotechnol. *31*, 684–686.

Lin, S., Ewen-Campen, B., Ni, X., Housden, B.E., and Perrimon, N. (2015). In Vivo Transcriptional Activation Using CRISPR/Cas9 in Drosophila. Genetics *201*, 433–442.

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

Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. Science *341*, 1237905.

Liu, N., Wang, M., Deng, W., Schmidt, C.S., Qin, W., Leonhardt, H., and Spada, F. (2013). Intrinsic and extrinsic connections of Tet3 dioxygenase with CXXC zinc finger modules. PloS One *8*, e62755.

Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R.A., and Jaenisch, R. (2016). Editing DNA Methylation in the Mammalian Genome. Cell *167*, 233-247.e17.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods San Diego Calif *25*, 402–408.

Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science *351*, 400–403.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251–260.

Ma, D., Peng, S., and Xie, Z. (2016). Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells. Nat. Commun. *7*.

Ma, J.-L., Kim, E.M., Haber, J.E., and Lee, S.E. (2003). Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. Mol. Cell. Biol. *23*, 8820–8828.

Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichtinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., et al. (2008). Rapid "Open-Source" Engineering of Customized Zinc-Finger Nucleases for Highly Efficient Gene Modification. Mol. Cell *31*, 294–301.

Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013). CRISPR RNA-guided activation of endogenous human genes. Nat. Methods *10*, 977–979.

al-Maghtheh, M., Gregory, C., Inglehearn, C., Hardcastle, A., and Bhattacharya, S. (1993). Rhodopsin mutations in autosomal dominant retinitis pigmentosa. Hum. Mutat. *2*, 249–255.

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

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. *9*, 467–477.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013a). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013b). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. *31*, 833–838.

Mandel, M., and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. J. Mol. Biol. *53*, 159–162.

Margolin, J.F., Friedman, J.R., Meyer, W.K., Vissing, H., Thiesen, H.J., and Rauscher, F.J. (1994). Krüppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. U. S. A. *91*, 4509–4513.

Masland, R.H. (2001). The fundamental plan of the retina. Nat. Neurosci. *4*, 877–886.

McDonald, J.I., Celik, H., Rois, L.E., Fishberger, G., Fowler, T., Rees, R., Kramer, A., Martens, A., Edwards, J.R., and Challen, G.A. (2016). Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. Biol. Open *5*, 866–874.

McGhee, J.D., and Felsenfeld, G. (1980). Nucleosome structure. Annu. Rev. Biochem. *49*, 1115–1156.

Mcghee, J.D., and Ginder, G.D. (1979). Specific DNA methylation sites in the vicinity of the chicken β-globin genes. Nature *280*, 419–420.

McGinty, R.K., and Tan, S. (2015). Nucleosome Structure and Function. Chem. Rev. *115*, 2255–2273.

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

Melamed, P., Yosefzon, Y., David, C., Tsukerman, A., and Pnueli, L. (2018). Tet Enzymes, Variants, and Differential Effects on Function. Front. Cell Dev. Biol. *6*.

Memedula, S., and Belmont, A.S. (2003). Sequential recruitment of HAT and SWI/SNF components to condensed chromatin by VP16. Curr. Biol. CB *13*, 241–246.

Mendes, H.F., and Cheetham, M.E. (2008). Pharmacological manipulation of gain-of-function and dominant-negative mechanisms in rhodopsin retinitis pigmentosa. Hum. Mol. Genet. *17*, 3043–3054.

Merienne, N., Vachey, G., de Longprez, L., Meunier, C., Zimmer, V., Perriard, G., Canales, M., Mathias, A., Herrgott, L., Beltraminelli, T., et al. (2017). The Self-Inactivating KamiCas9 System for the Editing of CNS Disease Genes. Cell Rep. *20*, 2980–2991.

Michalakis, S., Geiger, H., Haverkamp, S., Hofmann, F., Gerstner, A., and Biel, M. (2005). Impaired opsin targeting and cone photoreceptor migration in the retina of mice lacking the cyclic nucleotide-gated channel CNGA3. Invest. Ophthalmol. Vis. Sci. *46*, 1516–1524.

Michalakis, S., Mühlfriedel, R., Tanimoto, N., Krishnamoorthy, V., Koch, S., Fischer, M.D., Becirovic, E., Bai, L., Huber, G., Beck, S.C., et al. (2010). Restoration of cone vision in the CNGA3-/- mouse model of congenital complete lack of cone photoreceptor function. Mol. Ther. J. Am. Soc. Gene Ther. *18*, 2057–2063.

Michalakis, S., Koch, S., Sothilingam, V., Garcia Garrido, M., Tanimoto, N., Schulze, E., Becirovic, E., Koch, F., Seide, C., Beck, S.C., et al. (2014). Gene therapy restores vision and delays degeneration in the CNGB1(-/-) mouse model of retinitis pigmentosa. Adv. Exp. Med. Biol. *801*, 733–739.

Michalakis, S., Becirovic, E., and Biel, M. (2018). Retinal Cyclic Nucleotide-Gated Channels: From Pathophysiology to Therapy. Int. J. Mol. Sci. *19*.

Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. Nat. Biotechnol. *29*, 143–148.

Mills, K.V., Johnson, M.A., and Perler, F.B. (2014). Protein Splicing: How Inteins Escape from Precursor Proteins. J. Biol. Chem. *289*, 14498–14505.

Mingozzi, F., and High, K.A. (2011). Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. Nat. Rev. Genet. *12*, 341–355.

Mittler, G., Stühler, T., Santolin, L., Uhlmann, T., Kremmer, E., Lottspeich, F., Berti, L., and Meisterernst, M. (2003). A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. EMBO J. *22*, 6494–6504.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. J. Mol. Evol. *60*, 174–182.

Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. *42*, W401-407.

Moreno, A.M., Fu, X., Zhu, J., Katrekar, D., Shih, Y.-R.V., Marlett, J., Cabotaje, J., Tat, J., Naughton, J., Lisowski, L., et al. (2018). In Situ Gene Therapy via AAV-CRISPR-Cas9-Mediated Targeted Gene Regulation. Mol. Ther. J. Am. Soc. Gene Ther. *26*, 1818–1827.

Moscou, M.J., and Bogdanove, A.J. (2009). A Simple Cipher Governs DNA Recognition by TAL Effectors. Science *326*, 1501–1501.

Müller, H. (1854). Über die entoptische Wahrnehmung der Netzhautgefäße, insbesondere als Beweismittel fur die Lichtperzeption durch die nach hinten gelegenen Netzhautelemente. Verh. Phys. Med. Ges. Würzburg *5*.

Munger, S.D., Lane, A.P., Zhong, H., Leinders-Zufall, T., Yau, K.W., Zufall, F., and Reed, R.R. (2001). Central role of the CNGA4 channel subunit in Ca2+-calmodulin-dependent odor adaptation. Science *294*, 2172–2175.

Münzel, M., Globisch, D., Brückl, T., Wagner, M., Welzmiller, V., Michalakis, S., Müller, M., Biel, M., and Carell, T. (2010). Quantification of the sixth DNA base hydroxymethylcytosine in the brain. Angew. Chem. Int. Ed Engl. *49*, 5375–5377.

Münzel, M., Globisch, D., and Carell, T. (2011). 5-Hydroxymethylcytosine, the sixth base of the genome. Angew. Chem. Int. Ed Engl. *50*, 6460–6468.

Musselman, C.A., Lalonde, M.-E., Côté, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. Nat. Struct. Mol. Biol. *19*, 1218–1227.

Neely, K.E., Hassan, A.H., Wallberg, A.E., Steger, D.J., Cairns, B.R., Wright, A.P., and Workman, J.L. (1999). Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays. Mol. Cell *4*, 649–655.

Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Castellanos Rivera, R.M., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X., et al. (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science *351*, 403–407.

Nihongaki, Y., Kawano, F., Nakajima, T., and Sato, M. (2015). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. Nat. Biotechnol. *33*, 755–760.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. Cell *156*, 935–949.

Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W., et al. (2014). Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell *156*, 836–843.

Okano, M., Xie, S., and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat. Genet. *19*, 219–220.

Olsson, J.E., Gordon, J.W., Pawlyk, B.S., Roof, D., Hayes, A., Molday, R.S., Mukai, S., Cowley, G.S., Berson, E.L., and Dryja, T.P. (1992). Transgenic mice with a rhodopsin mutation (Pro23His): a mouse model of autosomal dominant retinitis pigmentosa. Neuron *9*, 815–830.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., et al. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. Science *289*, 739–745.

Palfi, A., Chadderton, N., O'Reilly, M., Nagel-Wolfrum, K., Wolfrum, U., Bennett, J., Humphries, P., Kenna, P., Millington-Ward, S., and Farrar, J. (2015). Efficient gene delivery to photoreceptors using AAV2/rh10 and rescue of the Rho(-/-) mouse. Mol. Ther. Methods Clin. Dev. *2*, 15016.

Pannunzio, N.R., Watanabe, G., and Lieber, M.R. (2018). Nonhomologous DNA end-joining for repair of DNA double-strand breaks. J. Biol. Chem. *293*, 10512–10523.

Park, S.H., Lee, C.M., Deshmukh, H., and Bao, G. (2016). Therapeutic Crispr/Cas9 Genome Editing for Treating Sickle Cell Disease. Blood *128*, 4703–4703.

Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat. Biotechnol. *31*, 839–843.

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

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods *10*, 973–976.

Petersen-Jones, S.M., Occelli, L.M., Winkler, P.A., Lee, W., Sparrow, J.R., Tsukikawa, M., Boye, S.L., Chiodo, V., Capasso, J.E., Becirovic, E., et al. (2018). Patients and animal models of CNGβ1-deficient retinitis pigmentosa support gene augmentation approach. J. Clin. Invest. *128*, 190–206.

Pfaffeneder, T., Spada, F., Wagner, M., Brandmayr, C., Laube, S.K., Eisen, D., Truss, M., Steinbacher, J., Hackner, B., Kotljarova, O., et al. (2014). Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. Nat. Chem. Biol. *10*, 574–581.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. *29*, e45.

Porteus, M.H., and Baltimore, D. (2003). Chimeric nucleases stimulate gene targeting in human cells. Science *300*, 763.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173–1183.

R Development Core Team (2008). R: A language and environment for statistical computing. (Vienna, Austria: R Foundation for Statistical Computing).

Ramakrishna, S., Kwaku Dad, A.-B., Beloor, J., Gopalappa, R., Lee, S.-K., and Kim, H. (2014). Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Genome Res. *24*, 1020–1027.

Ramlogan-Steel, C.A., Murali, A., Andrzejewski, S., Dhungel, B., Steel, J.C., and Layton, C.J. (2018). Gene therapy and the adeno associated virus in the treatment of genetic and acquired ophthalmic diseases in humans: trials, future directions and safety considerations. Clin. Experiment. Ophthalmol.

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380–1389.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186–191.

Reich, S.J., Auricchio, A., Hildinger, M., Glover, E., Maguire, A.M., Wilson, J.M., and Bennett, J. (2003). Efficient trans-splicing in the retina expands the utility of adeno-associated virus as a vector for gene therapy. Hum. Gene Ther. *14*, 37–44.

Remmer, M.H., Rastogi, N., Ranka, M.P., and Ceisler, E.J. (2015). Achromatopsia: a review. Curr. Opin. Ophthalmol. *26*, 333–340.

Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. Cytogenet. Cell Genet. *14*, 9–25.

Robertson, K.D. (2005). DNA methylation and human disease. Nat. Rev. Genet. *6*, 597–610.

Robertson, K.D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F.A., and Jones, P.A. (1999). The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. Nucleic Acids Res. *27*, 2291–2298.

Rossmiller, B., Mao, H., and Lewin, A.S. (2012). Gene therapy in animal models of autosomal dominant retinitis pigmentosa. Mol. Vis. *18*, 2479–2496.

Ryu, S.-M., Koo, T., Kim, K., Lim, K., Baek, G., Kim, S.-T., Kim, H.S., Kim, D.-E., Lee, H., Chung, E., et al. (2018). Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. Nat. Biotechnol. *36*, 536–539.

Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. Nature *335*, 563–564.

Sallmyr, A., and Tomkinson, A.E. (2018). Repair of DNA double-strand breaks by mammalian alternative end-joining pathways. J. Biol. Chem. *293*, 10536–10546.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Res. *39*, 9275–9282.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schmitz, M.L., and Baeuerle, P.A. (1991). The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. EMBO J. *10*, 3805–3817.

Schön, C., Biel, M., and Michalakis, S. (2015). Retinal gene delivery by adeno-associated virus (AAV) vectors: Strategies and applications. Eur. J. Pharm. Biopharm. Off. J. Arbeitsgemeinschaft Pharm. Verfahrenstechnik EV *95*, 343–352.

Schultze, M. (1866). Zur Anatomie und Physiologie der Retina. In Arch. Mikrosk. Anat. Entwicklungsmech, pp. 165–286.

Schwartz, E.C., Saez, L., Young, M.W., and Muir, T.W. (2007). Post-translational enzyme activation in an animal via optimized conditional protein splicing. Nat. Chem. Biol. *3*, 50–54.

Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl. Acad. Sci. U. S. A. *108*, 10098–10103.

Shah, N.H., Vila-Perelló, M., and Muir, T.W. (2011). Kinetic control of one-pot trans-splicing reactions by using a wild-type and designed split intein. Angew. Chem. Int. Ed Engl. *50*, 6511–6515.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84–87.

Shalem, O., Sanjana, N.E., and Zhang, F. (2015). High-throughput functional genomics using CRISPR-Cas9. Nat. Rev. Genet. *16*, 299–311.

Shi, G., Yau, K.-W., Chen, J., and Kefalov, V.J. (2007). Signaling properties of a short-wave cone visual pigment and its role in phototransduction. J. Neurosci. Off. J. Soc. Neurosci. *27*, 10084–10093.

Shuart, N.G., Haitin, Y., Camp, S.S., Black, K.D., and Zagotta, W.N. (2011). Molecular mechanism for 3:1 subunit stoichiometry of rod cyclic nucleotide-gated ion channels. Nat. Commun. *2*, 457.

Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. Science *351*, 84–88.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. *14*, 204–220.

Smith, Z.D., Gu, H., Bock, C., Gnirke, A., and Meissner, A. (2009). High-throughput bisulfite sequencing in mammalian genomes. Methods San Diego Calif *48*, 226–232.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. *3*, Article3.

Soutourina, J. (2018). Transcription regulation by the Mediator complex. Nat. Rev. Mol. Cell Biol. *19*, 262–274.

Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature *507*, 62–67.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Strong, A., and Musunuru, K. (2017). Genome editing in cardiovascular diseases. Nat. Rev. Cardiol. *14*, 11–20.

Sundin, O.H., Yang, J.M., Li, Y., Zhu, D., Hurd, J.N., Mitchell, T.N., Silva, E.D., and Maumenee, I.H. (2000). Genetic basis of total colourblindness among the Pingelapese islanders. Nat. Genet. *25*, 289–293.

Sweatt, J.D. (2013). The emerging field of neuroepigenetics. Neuron *80*, 624–632.

Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., Sur, M., and Zhang, F. (2015). In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. Nat. Biotechnol. *33*, 102–106.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W. (1983). The double-strand-break repair model for recombination. Cell *33*, 25–35.

Szulwach, K.E., Li, X., Li, Y., Song, C.-X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., et al. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. Nat. Neurosci. *14*, 1607–1616.

Szwagierczak, A., Bultmann, S., Schmidt, C.S., Spada, F., and Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res. *38*, e181.

Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A., et al. (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science *351*, 407–411.

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

Takeuchi, R., Lambert, A.R., Mak, A.N.-S., Jacoby, K., Dickson, R.J., Gloor, G.B., Scharenberg, A.M., Edgell, D.R., and Stoddard, B.L. (2011). Tapping natural reservoirs of homing endonucleases for targeted gene modification. Proc. Natl. Acad. Sci. U. S. A. *108*, 13077–13082.

Tan, E., Wang, Q., Quiambao, A.B., Xu, X., Qtaishat, N.M., Peachey, N.S., Lem, J., Fliesler, S.J., Pepperberg, D.R., Naash, M.I., et al. (2001). The relationship between opsin overexpression and photoreceptor degeneration. Invest. Ophthalmol. Vis. Sci. *42*, 589–600.

Tan, E., Ding, X.-Q., Saadi, A., Agarwal, N., Naash, M.I., and Al-Ubaidi, M.R. (2004). Expression of cone-photoreceptor-specific antigens in a cell line derived from retinal tumors in transgenic mice. Invest. Ophthalmol. Vis. Sci. *45*, 764–768.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., et al. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell *146*, 1016–1028.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell *159*, 635–646.

Taylor, S.M., and Jones, P.A. (1979). Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. Cell *17*, 771–779.

Tesson, L., Usal, C., Ménoret, S., Leung, E., Niles, B.J., Remy, S., Santiago, Y., Vincent, A.I., Meng, X., Zhang, L., et al. (2011). Knockout rats generated by embryo microinjection of TALENs. Nat. Biotechnol. *29*, 695–696.

Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat. Methods *12*, 1143–1149.

Trapani, I., Colella, P., Sommella, A., Iodice, C., Cesi, G., de Simone, S., Marrocco, E., Rossi, S., Giunti, M., Palfi, A., et al. (2014). Effective delivery of large genes to the retina by dual AAV vectors. EMBO Mol. Med. *6*, 194–211.

Traube, F.R., Schiffers, S., Iwan, K., Kellner, S., Spada, F., Müller, M., and Carell, T. (2019). Isotope-dilution mass spectrometry for exact quantification of noncanonical DNA nucleosides. Nat. Protoc. *14*, 283–312.

Tremethick, D.J. (2007). Higher-order structures of chromatin: the elusive 30 nm fiber. Cell *128*, 651–654.

Truong, D.-J.J., Kühner, K., Kühn, R., Werfel, S., Engelhardt, S., Wurst, W., and Ortiz, O. (2015). Development of an intein-mediated split–Cas9 system for gene therapy. Nucleic Acids Res. *43*, 6450–6458.

Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., and Joung, J.K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat. Biotechnol. *32*, 569–576.

al-Ubaidi, M.R., Font, R.L., Quiambao, A.B., Keener, M.J., Liou, G.I., Overbeek, P.A., and Baehr, W. (1992). Bilateral retinal and brain tumors in transgenic mice expressing simian virus 40 large T antigen under control of the human interphotoreceptor retinoid-binding protein promoter. J. Cell Biol. *119*, 1681–1687.

Ulge, U.Y., Baker, D.A., and Monnat, R.J. (2011). Comprehensive computational design of mCreI homing endonuclease cleavage specificity for genome engineering. Nucleic Acids Res. *39*, 4330–4339.

Urnov, F.D., Miller, J.C., Lee, Y.-L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., and Holmes, M.C. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature *435*, 646–651.

Verbakel, S.K., van Huet, R.A.C., Boon, C.J.F., den Hollander, A.I., Collin, R.W.J., Klaver, C.C.W., Hoyng, C.B., Roepman, R., and Klevering, B.J. (2018). Non-syndromic retinitis pigmentosa. Prog. Retin. Eye Res. *66*, 157–186.

Vojta, A., Dobrinić, P., Tadić, V., Bočkor, L., Korać, P., Julg, B., Klasić, M., and Zoldoš, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic Acids Res. *44*, 5615–5628.

Waddington, C.H. (1957). The strategy of the genes : a discussion of some aspects of theoretical biology (London: Allen and Unwin).

Wally, V., Murauer, E.M., and Bauer, J.W. (2012). Spliceosome-mediated trans-splicing: the therapeutic cut and paste. J. Invest. Dermatol. *132*, 1959–1966.

Wang, H., and Xu, X. (2017). Microhomology-mediated end joining: new players join the team. Cell Biosci. *7*.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell *153*, 910–918.

Weinhold, B. (2006). Epigenetics: The Science of Change. Environ. Health Perspect. *114*, A160–A167.

Weitz, D., Ficek, N., Kremmer, E., Bauer, P.J., and Kaupp, U.B. (2002). Subunit stoichiometry of the CNG channel of rod photoreceptors. Neuron *36*, 881–889.

Weltner, J., Balboa, D., Katayama, S., Bespalov, M., Krjutškov, K., Jouhilahti, E.-M., Trokovic, R., Kere, J., and Otonkoski, T. (2018). Human pluripotent reprogramming with CRISPR activators. Nat. Commun. *9*, 2643.

Wiedenheft, B., van Duijn, E., Bultema, J.B., Bultema, J., Waghmare, S.P., Waghmare, S., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J.R., et al. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc. Natl. Acad. Sci. U. S. A. *108*, 10092–10097.

Witherell, G.W., Gott, J.M., and Uhlenbeck, O.C. (1991). Specific Interaction between RNA Phage Coat Proteins and RNA. In Progress in Nucleic Acid Research and Molecular Biology, W.E. Cohn, and K. Moldave, eds. (Academic Press), pp. 185–220.

Wright, A.V., Sternberg, S.H., Taylor, D.W., Staahl, B.T., Bardales, J.A., Kornfeld, J.E., and Doudna, J.A. (2015). Rational design of a split-Cas9 enzyme complex. Proc. Natl. Acad. Sci. U. S. A. *112*, 2984–2989.

Wright, W.D., Shah, S.S., and Heyer, W.-D. (2018). Homologous recombination and the repair of DNA double-strand breaks. J. Biol. Chem. *293*, 10524–10535.

Wu, T.P., Wang, T., Seetin, M.G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S.D., Mackintosh, S.G., Zhong, M., et al. (2016). DNA methylation on N(6)-adenine in mammalian embryonic stem cells. Nature *532*, 329–333.

Wyatt, G.R. (1951). Recognition and estimation of 5-methylcytosine in nucleic acids. Biochem. J. *48*, 581–584.

Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J.D., and Zhao, Y. (2012). Lysine succinylation and lysine malonylation in histones. Mol. Cell. Proteomics MCP *11*, 100–107.

Xu, J. (2005). Preparation, culture, and immortalization of mouse embryonic fibroblasts. Curr. Protoc. Mol. Biol. *Chapter 28*, Unit 28.1.

Xu, J., Morris, L., Fliesler, S.J., Sherry, D.M., and Ding, X.-Q. (2011). Early-onset, slow progression of cone photoreceptor dysfunction and degeneration in CNG channel subunit CNGB3 deficiency. Invest. Ophthalmol. Vis. Sci. *52*, 3557–3566.

Xu, X., Tao, Y., Gao, X., Zhang, L., Li, X., Zou, W., Ruan, K., Wang, F., Xu, G.-L., and Hu, R. (2016). A CRISPR-based approach for targeted DNA demethylation. Cell Discov. *2*, 16009.

Yang, R.B., Foster, D.C., Garbers, D.L., and Fülle, H.J. (1995). Two membrane forms of guanylyl cyclase found in the eye. Proc. Natl. Acad. Sci. U. S. A. *92*, 602–606.

Yeo, N.C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., Kuo, C.-C., Guo, X., Sharma, S., Tung, A., et al. (2018). An enhanced CRISPR repressor for targeted mammalian gene regulation. Nat. Methods *15*, 611–616.

Yin, H., Xue, W., Chen, S., Bogorad, R.L., Benedetti, E., Grompe, M., Koteliansky, V., Sharp, P.A., Jacks, T., and Anderson, D.G. (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat. Biotechnol. *32*, 551–553.

Yokoyama, S. (2000). Molecular evolution of vertebrate visual pigments. Prog. Retin. Eye Res. *19*, 385–419.

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.-X., Jin, P., Ren, B., and He, C. (2012). Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. Nat. Protoc. *7*, 2159–2170.

Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., et al. (2015). Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. Cell *160*, 339–350.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015a). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell *163*, 759–771.

Zetsche, B., Volz, S.E., and Zhang, F. (2015b). A split-Cas9 architecture for inducible genome editing and transcription modulation.

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E.M., Winblad, N., Choudhury, S.R., Abudayyeh, O.O., Gootenberg, J.S., et al. (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat. Biotechnol. *35*, 31–34.

Zhang, Y., Molday, L.L., Molday, R.S., Sarfare, S.S., Woodruff, M.L., Fain, G.L., Kraft, T.W., and Pittler, S.J. (2009). Knockout of GARPs and the β-subunit of the rod cGMP-gated channel disrupts disk morphogenesis and rod outer segment structural integrity. J. Cell Sci. *122*, 1192–1200.

Zhang, Y., Zhang, F., Li, X., Baller, J.A., Qi, Y., Starker, C.G., Bogdanove, A.J., and Voytas, D.F. (2013). Transcription activator-like effector nucleases enable efficient plant genome engineering. Plant Physiol. *161*, 20–27.

Zhang, Y., Yin, C., Zhang, T., Li, F., Yang, W., Kaminski, R., Fagan, P.R., Putatunda, R., Young, W.-B., Khalili, K., et al. (2015). CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs. Sci. Rep. *5*, 16277.

Zheng, J., Trudeau, M.C., and Zagotta, W.N. (2002). Rod cyclic nucleotide-gated channels have a stoichiometry of three CNGA1 subunits and one CNGB1 subunit. Neuron *36*, 891–896.

Zhong, H., Molday, L.L., Molday, R.S., and Yau, K.-W. (2002). The heteromeric cyclic nucleotide-gated channel adopts a 3A:1B stoichiometry. Nature *420*, 193–198.

Zinn, E., and Vandenberghe, L.H. (2014). Adeno-associated virus: fit to serve. Curr. Opin. Virol. *8*, 90–97.

Zong, X., Zucker, H., Hofmann, F., and Biel, M. (1998). Three amino acids in the C-linker are major determinants of gating in cyclic nucleotide-gated channels. EMBO J. *17*, 353–362.

Zou, Q., Wang, X., Liu, Y., Ouyang, Z., Long, H., Wei, S., Xin, J., Zhao, B., Lai, S., Shen, J., et al. (2015). Generation of gene-target dogs using CRISPR/Cas9 system. J. Mol. Cell Biol. *7*, 580–583.

Zuo, E., Cai, Y.-J., Li, K., Wei, Y., Wang, B.-A., Sun, Y., Liu, Z., Liu, J., Hu, X., Wei, W., et al. (2017). One-step generation of complete gene knockout mice and monkeys by CRISPR/Cas9-mediated gene editing with multiple sgRNAs. Cell Res. *27*, 933–945.

Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.-Y., and Liu, D.R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat. Biotechnol. *33*, 73–80.

Acknowledgments

First of all, I deeply thank my family for their love and continuous support during this demanding time. I know I can always count on you, which gives me great comfort. I dedicate this thesis to all of you.

I especially thank my supervisor, PD Dr. Stylianos Michalakis, for his scientific guidance and always taking the time to discuss new ideas. Thank you, Stelio, for never stopping to believe in me throughout the ups and downs of my projects.

I am very grateful to Prof. Dr. Martin Biel for the opportunity to join his lab and his scientific support as one of my TAC members. Thank you to Dr. Irina Solovei for being my third TAC member and the helpful scientific advice. A big thank you also goes to Dr. Elvir Becirovic for making me feel like an honorary member of his group.

Thank you to all the collaborators at the Carell, Vollmar and Walter groups for all the work you have put into these projects and the pleasant communication. I would like to particularly thank Franziska Traube from the Carell lab, Themistoklis Zisis from the Vollmar group and Gilles Gasparoni from the Walter lab in Saarbrücken.

I especially thank the GSN for welcoming me into their excellent program and this extraordinary community. I really appreciate the informative courses and excursions supporting me in my personal development. Thank you also for the financial support. Without it, I would have missed out on a lot of valuable experiences.

I am extremely thankful to my colleagues who made the daily work so much easier. I will always cherish the joined lunch breaks, hat crafting events as well as after-work activities. Special thanks to Anna, Constanze, Johanna, Sybille and Lisa for becoming such good friends over the past years. I could not have finished this PhD without your support, your comforting and encouraging words that were badly needed more than once. Anna, I cannot believe that we really made it and I thank you so much for caring for and standing by me in stressful times. Sybille, thanks for always checking in on me and cheering me up with your positive spirit. I thank you for the fruitful collaboration in the transactivation project as well as being such a great companion on our personal adventures. Lisa, I admire you for being such a passionate scientist and strong-minded woman and thank you for each scientific discussion, your help with programming and rephrasing my English sentences. Stanzi, I thank you for always offering your help and trying to take some work off my plate. Thank you, Marina, for your patience with me at the confocal. I thank René for performing the voltage-clamp measurements on 661W-pb cells.

Affidavit

Eidesstattliche Versicherung/ Affidavit

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation „GENOME EDITING AND TRANSCRIPTIONAL ACTIVATION USING CRISPR-CAS9 TECHNOLOGY - EXPLORING AND MANIPULATING EPIGENETIC MECHANISMS IN NEURONS“ selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation “GENOME EDITING AND TRANSCRIPTIONAL ACTIVATION USING CRISPR-CAS9 TECHNOLOGY - EXPLORING AND MANIPULATING EPIGENETIC MECHANISMS IN NEURONS” is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

München/Munich, 18.07.2019 Victoria Splith

Declaration of Author Contributions

TET3KO iNGN cell line

Quantification of cytosine modifications using gDNA from iNGN-derived neurons was performed by Franziska R. Traube (Carell group, LMU München). RNA-sequencing using RNA from iNGN-derived neurons was conducted by Dr. Gilles Gasparoni (Walter group, Universität des Saarlandes, Saarbrücken). Subsequent RNA-seq data analysis was done by Dr. Karl Nordström (Walter group, Universität des Saarlandes, Saarbrücken).

Transactivation

Inside-out voltage-clamp measurements of 661W-pb cells were performed by René Rötzer (Wahl-Schott group, LMU München). RT-PCRs using *Cnga1* and *Opn1mw* specific primers were conducted by Dr. Sybille Böhm (Biel group, LMU München). αCas9 WB of HEK293 cells transfected with E573 and V713 split Cas9 as well as quantification of the reconstitution efficiencies were performed by Dr. Sybille Böhm.

Munich, 18.07.2019

Supervisor (Prof. Dr. Stylianos Michalakis) Victoria Splith