Optimizing an AAV-CRISPR System for Retinal Gene Therapy

Marta Stevanovic A.B.
Department of Physiology, Anatomy, and Genetics
and
Merton College

This thesis was submitted as part of the requirement for the degree of Master of Science by Research in Anatomy, Physiology, and Genetics
University of Oxford, Michaelmas Term 2019

Supervisor: Professor Robert MacLaren

Declaration

I declare that this thesis is my own original work, unless otherwise indicated. This thesis contains no material that has previously been submitted for the award of any academic degree or qualification at the University of Oxford or any other institution.

Acknowledgements

My time in the MacLaren laboratory has been a formative and life-changing experience and would not have been possible without the support of many people.

I would first and foremost like to express my deepest gratitude to Robert E. MacLaren for his guidance, encouragement, and mentorship during my time in his laboratory. I am very grateful that I had the opportunity to study under his supervision. His skill as both a basic scientist and clinician and his dedication to his work is truly inspiring.

I would also like to express my sincere thanks to Michelle McClements for all of her support and guidance. From my very first day in the laboratory, Michelle taught me with infinite care, clarity, and patience and has shared her scientific wisdom on many occasions.

Of course, I am grateful for the wonderful members of “Team CRISPR:” Michelle McClements, Caroline Peddle, Lewis Fry, Elena Piotter, and Alessia Valteroni. It was always helpful to have such a supportive group with which I could share ideas, wallow in sorrow, and celebrate triumphs.

Thank you also to my fellow Mertonians, Laurel Chandler, Jasleen Jolly Kaur, and Imran Yusuf, for their friendship and support. I am grateful to many other people in the group for their guidance and camaraderie: Jasmina Kapetanovic, Alun Barnard, Cristina Martinez, Won Kyung Song, Moreno Menghini, Marco Bellini, Peter Charbel Issa, Federica Staurenghi, Iain Wilson, Tina Storm, Maria Patricio, Ahmed Salman, Harry Orlans, Ross Campbell, Arantxa Bolinches Amoros, and Celine-Lea Halioua-Haubold. I would like to

thank Kella Vangsness for helping me make some of the figures in this thesis. I also appreciate the support of James Cantley, my Department of Physiology, Anatomy, and Genetics supervisor, and my Merton College advisor, Kieran Clarke.

I would like to thank Mark Humayun and Dennis Clegg for their mentorship and teaching during my first year as an HHMI research fellow. I am very grateful to the Emory University School of Medicine, particularly Ira Schwartz, for being so supportive of my scientific pursuits. I would also like to thank my family for all of their encouragement and love.

Finally, I would like to acknowledge the Howard Hughes Medical Institute and the Foundation Fighting Blindness for funding me and Melanie Daub for making it possible for me to extend my research fellowship to go “across the pond.”

Abbreviations

AAV	Adeno-associated virus
ACTB	Beta-actin
AMD	Age-related macular degeneration
ANOVA	Analysis of Variance
ATP	Adenosine triphosphate
bp	Base pair
cDNA	Complementary DNA
CEP290	Centrosomal protein 290
CNV	Choroidal neovascularization
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
DNA	Deoxyribonucleic acid
DNMT3A	DNA methyltransferase 3 alpha
dNTP	Deoxynucleoside triphosphate
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced green fluorescent protein
EIAV	Equine infectious anemia virus
endA	Endonuclease A
ERG	Electroretinogram
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HDR	Homology-directed repair
HEK293	Human embryonic kidney 293 cells
Hif-1a	Hypoxia inducible factor 1 subunit alpha
His	Histidine
HITI	Homology-independent targeted integration
HP1	Heterochromatin protein
Indel	Insertion or deletion
IRD	Inherited Retinal Degeneration
KAP-1	KRAB-box-associated protein-1
KRAB	Krüppel-associated box
LB	Luria broth
LCA	Leber's Congenital Amaurosis
LHON	Leber's Hereditary Optic Neuropathy
Mertk	Proto-oncogene tyrosine-protein kinase MER
MHC	Major Histocompatibility Complex
mRNA	Messenger RNA
ND4	NADH dehydrogenase subunit 4
NHEJ	Non-homologous end joining

Nrl	Neural retinal leucine zipper
NTC	Non-targeting control
OCT	Optical coherence tomography
ONL	Outer nuclear layer
OS	Outer segment
PAM	Protospacer adjacent motif
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PND	Post-natal day
PR	Photoreceptors
PVDF	Polyvinylidene difluoride
qPCR	Quantitative PCR
rAAV	Recombinant AAV
RCS	Royal College of Surgeons
recA	Recombinase A
REP-1	Rab-escort protein 1
Rho	Rhodopsin
RNA	Ribonucleic acid
RNase	Ribonuclease
RNP	Ribonucleoprotein
RO	Reverse osmosis
RPE	Retinal pigment epithelium
RPGR	Retinitis pigmentosa GTPase regulator
SAM	Synergistic activation mediator
sFLT-1	Soluble fms-like tyrosine kinase-1
sgRNA	Single guide RNA
TALEN	Transcription activator-like effector nucleases
TE	Tris-EDTA
TIDE	Tracking of indels by Decomposition
tracrRNA	Trans-activating CRISPR RNA
UPR	Unfolded protein response
VEGF	Vascular endothelial growth factor
XYL	Xylose reductase
YFP	Yellow fluorescent protein
ZFN	Zinc finger nuclease

Abstract

Introduction: CRISPR/Cas9 is a bacterial defence system which evolved to cleave and deactivate bacteriophage DNA. The elements involved in the CRISPR/Cas9 system may be modified for gene therapy. CRISPR/Cas9 consists of two components, a single guide RNA (sgRNA) and a Cas9 protein. The sgRNA allows for sequence-specific binding. “Active” Cas9 creates a double-stranded break in the target region while the “deactivated” form can inhibit transcription. CRISPR/Cas9 can be packaged within adeno-associated viral vectors, which opens the possibility of using CRISPR/Cas9 to target mutations in the retina. We examine several optimizations to the AAV-CRISPR/Cas9 system *in vitro* in HEK293-*eGFP* cells.

Materials and Methods: We compare the effects of sgRNA orientation in both active SaCas9 constructs and deactivated SaCas9 constructs with the KRAB repressor (dSaCas9-KRAB). We also examine the cleavage ability of GeoCas9.

Results: There was no difference in Cas9 and sgRNA levels between the *in cis* and *in trans* SaCas9 constructs. We also found no difference in Cas9 levels, sgRNA levels, and target knock-down between the *cis* and *trans* orientations of the dSaCas9-KRAB constructs. For GeoCas9, we found that DNA cleavage levels were undetectable, despite confirmed expression of both sgRNA and GeoCas9 protein.

Conclusion: When delivering CRISPR/Cas9 in an AAV-based expression cassette, the efficacy appears unrelated to the orientation of the sgRNA (*in cis* or *in trans)* relative to the Cas9 gene. Furthermore, GeoCas9 may not be as efficacious as SaCas9, since it was unable to cleave DNA detectably in similar experiments.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 CRISPR/Cas

Discovery

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and the CRISPR-associated proteins (Cas) are part of a naturally-occurring bacterial and archaeal adaptive immune system designed to protect against bacteriophages. CRISPR/Cas was first characterized by Francisco Mojica in 1993, who noticed short repetitive palindromic sequences separated by DNA spacers in *Haloferax mediterranei*, an archaeal microbe.[1] Mojica soon discovered similar patterns in many other microbes, and termed them clustered regularly interspaced palindromic repeats (CRISPR).[2–4] CRISPR-associated (*cas*) genes were found to be located near the CRISPR sequences but the function of both *CRISPR* and the *cas* genes at that time was unknown.[3] Using bioinformatics analysis, Mojica determined that the DNA spacer sequences matched regions of viral genomes, specifically of bacteriophages.[5] This finding led him to postulate that CRISPR/Cas represented an adaptive immune system within bacteria designed to help fight bacteriophage infections.[5]

Experimental evidence of CRISPR's role in the bacterial immune system was first uncovered by Philippe Horvath. Horvath had been trying to find a way to protect dairy products against bacteriophage infections.[6] Since *Streptococcus thermophilus* is commonly-used to make yogurt and cheese, Horvath focused his experiments on this bacterium. Horvath and colleagues showed that the *Streptococcus thermophilus* CRISPR system integrates bacteriophage DNA into the CRISPR locus and that a higher number of copies of phage-derived sequences conferred greater immunity against infections.[7]

The mechanisms of the CRISPR/Cas system were slowly coming to light. It was found that in *Escherichia coli*, the spacer DNA regions are transcribed into RNA—also known as CRISPR RNAs (crRNAs).[8] Trans-activating RNA (tracrRNA) was subsequently discovered, which binds to the crRNA; together, the crRNA and tracrRNA help guide Cas9 to the target DNA region.[9] Garneau et al. 2010 determined that when Cas9 proteins bind to viral DNA, they create double-stranded breaks.[10]

Mechanisms of Type I, II, and III CRISPR/Cas Systems

The CRISPR/Cas9 system was initially organized into three types (I, II, and III) based on an understanding of the molecular mechanisms governing its interaction with viral DNA (Figure 1.1). The function of these CRISPR/Cas systems can be broken down into three steps: spacer adaptation, expression, and interference.[11]

Spacer adaptation is conserved amongst the three types of CRISPR/Cas and occurs during the first infection by a bacteriophage. In this step, the viral genetic material is digested and incorporated into the spacer DNA region by Cas1 and Cas2 enzymes.[12,13]

The next step is expression, which occurs after the CRISPR locus is transcribed into mRNA. In type I systems, the CRISPR mRNA sequence first forms a hairpin secondary structure and then Cas6e and Cas6f cut the spacer/CRISPR complexes.[11,14,15] These spacer/CRISPR complexes are the crRNA. In type II systems, tracrRNA is bound to the CRISPR repeats and the CRISPR/tracrRNA/spacer sequences are cut by Cas9 and RNaseIII.[9] In type III systems, a Cas6 homolog cleaves the mRNA.[16]

The third step is interference, which occurs after the CRISPR/Cas system binds to the target viral DNA. Type I and II systems require a protospacer adjacent motif (PAM) for binding, which is a 2-6 bp sequence within the viral genome located near the target binding region.[17–20] The PAM site increases targeting specificity and ensures that the Cas binds to only viral DNA and not endogenous bacterial DNA. In class I systems, the crRNA associates with a Cas protein. The crRNA/Cas complex then binds to the viral genetic material. After binding, a cascade of Cas enzymes is activated, which ultimately results in Cas3 cutting the viral genetic material.[8] In type II systems, the crRNA is associated with Cas9. Upon binding to the target region, the HNH and RuvC domains of Cas9 create a double stranded break.[18] Type III systems do not require a PAM sequence.[21] Much like type I systems, the type III systems rely on an activated complex of proteins to destroy the viral genetic material.[22,23] Several other types of CRISPR/Cas systems have subsequently been discovered after the initial description of type I, II, and III systems.[24]

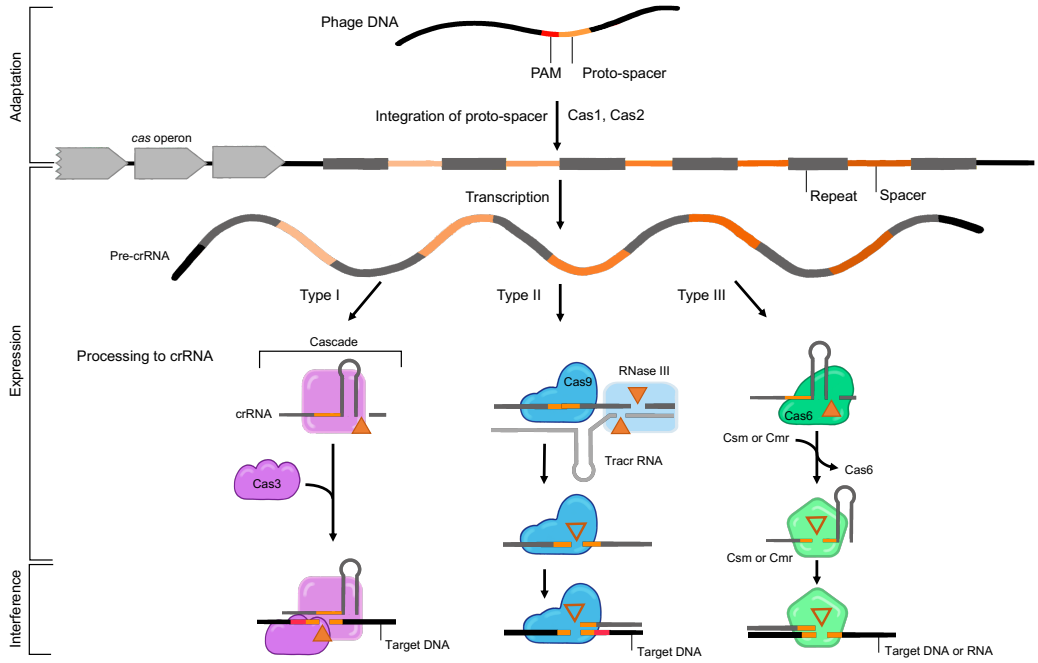

Figure 1.1 Mechanisms of Type I, II, and III CRISPR/Cas Systems

The figure above outlines the mechanisms of the three original CRIPSR/Cas systems.

Acquisition: After infection with a phage, Cas1 and Cas2 enzymes incorporate the phage proto-spacer DNA into the CRISPR locus as CRISPR spacers.

Expression: The CRISPR locus, along with the acquired spacers, is transcribed into the pre-crRNA. In type I systems, the CRISPR-associated complex for antiviral defence (Cascade) cleaves the pre-crRNA. In type II systems, the tracrRNA binds to the crRNA. Subsequently, RNaseII cleaves the crRNA-tracrRNA repeats in the presence of Cas9. In type III systems, Cas6 and either Csm or Cmr are responsible for pre-crRNA cleavage and processing.

Interference: In type I systems, the crRNA-Cascade complex binds to target DNA and the Cas3 is responsible for cleaving viral DNA. In type II systems, the Cas9 is guided to and cleaves viral DNA. Both type I and II systems require a PAM sequence in the viral DNA as well as crRNA complementarity for binding. In type III systems, either viral DNA or RNA is targeted by a crRNA bound to an enzyme. Type III systems do not rely on a PAM sequence for binding.

Image illustrated by Kella Vansgness.

Modifications to Type II CRISPR/Cas for Gene Editing

Once the basic elements of the CRISPR/Cas system had been determined, several adaptations were made before it was used for gene editing. Scientists

initially focused on modifying type II systems for use in synthetic biology. Gasiunas et al. 2012 determined that the crRNA, which is normally approximately 30 bp long, could be reduced to 20 bp and still produce efficient editing.[25] This group also determined that changing the crRNA could allow different target regions to be cut by the Cas9 protein.[25] Charpentier and Doudna simplified the CRISPR/Cas9 system by fusing the crRNA and tracrRNA into a single guide RNA (sgRNA).[18] In 2013, the Zhang laboratory was the first to show CRISPR/Cas9 editing in eukaryotic cells.[26] This group was also the first to package the CRISPR/Cas9 components into adeno-associated viral vectors for delivery into cells.[27,28]

Mechanism of Repair of Double-Stranded Breaks by CRISPR/Cas9

The mechanism of DNA repair is the same for both wild-type and engineered CRISPR systems. Once a double-stranded break is created by the Cas9 enzyme, there are two main types of repair mechanisms that may occur: non-homologous end joining (NHEJ) and homology-directed repair (HDR) (Figure 1.2). In NHEJ, the cell repairs the double-stranded break by ligating the two break ends together, a process that can randomly introduce insertions and deletions (indels).[29] These indels may be completely benign, and simply deactivate a certain gene. They may, however, create deleterious gain-of-function mutations. In HDR, the double-stranded break is repaired with the aid of a DNA template.[30] This template can either be the endogenous sister chromosome, which is present in the late S and G2 phase of the cell cycle, or it can be exogenous donor DNA.[31,32] While HDR is a relatively error-free process, and can thus be used to correct a mutated gene, its use is limited to the S and G2 phases of the cell cycle.

One innovation in CRISPR/Cas9 gene therapy is homology-independent targeted integration (HITI).[33] Mature differentiated cells that are not undergoing cell division, like photoreceptors, have difficulty performing homology-directed repair. These post-mitotic cells, however, have efficient NHEJ. In HITI, the target site is cut using CRISPR/Cas9 and a DNA template is supplied, in excess, in the hopes that this ectopic DNA will be directly inserted into the cut site during NHEJ repair.[33] HITI repair has been applied to correct the *Mertk* mutation in Royal College of Surgeons rats, which are animal models for retinal degeneration.[33] Treated rats had increased *Mertk* mRNA levels, better outer nuclear layer (ONL) preservation, and improved ERG b wave responses compared to the untreated controls.[33]

Off-Target Editing

CRISPR/Cas9 has the potential to induce off-target editing. Off-target effects broadly fall into two categories: those that occur in regions with great similarity to the target, which can be addressed by PCR of predicted off-target sites, and those that occur in regions that are dissimilar to the target, which can be determined by whole genome sequencing.[34]

There are several methods to assess off-target editing. Generally, *in silico* predictions are used first. There are many algorithms available to researchers, and they represent a cost-effective and efficient tool to evaluate CRISPR/Cas9 editing.[35–39] PCR can subsequently be performed on the predicted regions to determine if off-target editing has occurred. However, these *in silico* programs may yield very broad results, sometimes predicting hundreds or thousands of possible off-target sites. In the case of numerous predictions, it can be very difficult to perform PCR on all of the possible sites.

Another approach is to use *in vitro* genome-wide assays. In most of these assays, cell-free genomic DNA is cleaved by CRISPR/Cas9 and software is then used to detect double-stranded breaks. *In vitro* assays can be very sensitive, detecting sites with editing rates of less than 0.1%.[40] Several commonly-used *in vitro* assays include Digenome-seq, CIRCLE,-Seq, and SITE-Seq.[40–42]

Cell-based assays, in which CRISPR/Cas9-induced DNA cleavage is detected in cells, can also be helpful. Cell-based assays have the advantage of testing CRISPR/Cas9 off-target editing in chosen cell types and under specific experimental conditions. Examples of these assays include GUIDE-Seq, LAM-HTGTS, and BLISS.[43–45]

Although there is no gold standard for determining off-target effects, combining several methodologies is a robust way to uncover where off-target editing may occur. For example, Maeder et al. 2019 used three methods to assess possible off-target sites of a CRISPR/Cas9-based treatment for Leber Congenital Amaurosis 10 (LCA10): *in silico* prediction, Digenome-Seq, and GUIDE-Seq.[46] This CRISPR/Cas9 therapy for LCA10 is now being investigated in a clinical trial.

Deactivated CRISPR/Cas9 for Transcriptional Repression

There is concern about off-target cleavage by the CRISPR/Cas9 system due to non-specific binding of the sgRNAs. While “active,” cutting, versions of the CRISPR/Cas9 are certainly an important tool, an alternative option to cutting DNA may simply be to repress transcription of the target region.[47] Transcriptional repression can be achieved by using “deactivated” Cas9 proteins (dCas9).[47] While off-target effects may still occur when using the deactivated Cas9 proteins due to

non-specific binding of the sgRNA, the results are in theory less likely to be deleterious as there is no permanent change to the DNA sequence.

While dCas9 alone can affect transcription by interfering with transcription factor and RNA Polymerase II binding, this effect can be enhanced by fusing dCas9 to a repressor protein.[47] One of the most widely-used repressors is the Krüppel-associated box (KRAB).[48,49] KRAB modifies histones, thus reducing chromatin accessibility, and recruits additional co-repressor proteins like KRAB-box-associated protein-1 (KAP-1) and heterochromatin protein (HP1).[50–53] DCas9 can also be combined with DNA methyltransferases, like DNMT3A and MQ1, to reduce transcription.[54–56] To enhance gene transcription, dCas9 can be fused with activators like VP64-P65-Rta, synergistic activation mediator (SAM) complex, and SunTag.[57–62]

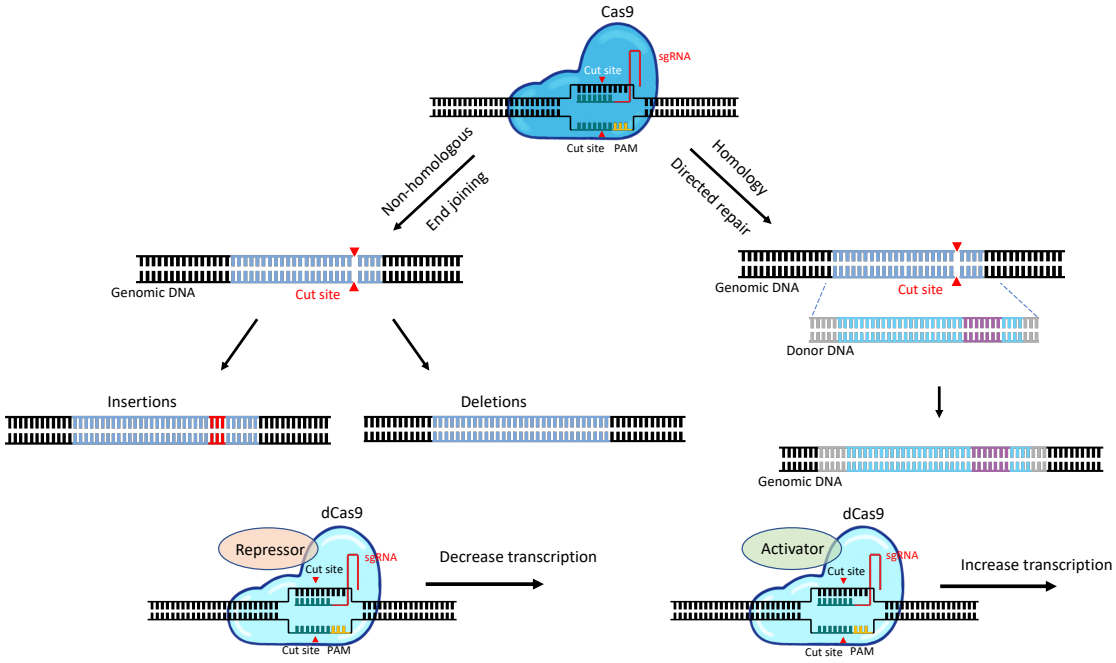

Figure 1.2. Schematic of CRISPR/Cas9 Editing

The sgRNA, which is comprised of the crRNA and tracrRNA components, is attached to the Cas9 protein (top). The sgRNA and protospacer adjacent motif (PAM) site allow for specific binding of the CRISPR/Cas9 system to the target DNA strand.

After binding, the Cas9 protein creates a double-stranded break in the DNA. In the absence of a template, non-homologous end joining will create indels (left side). If there is a repair template present, then homology-directed repair can occur and the double-stranded break will be repaired based on the sequence in the template (right side).

Deactivated CRISPR/Cas9 can be associated with a repressor (bottom left) to decrease transcription or an activator (bottom right) to increase transcription.

Image illustrated by Kella Vansgness.

1.2 Retinal Degeneration

A CRISPR/Cas9 system delivered to cells within an AAV may be helpful in treating retinal degenerations, most of which have a known genetic mechanism. Retinal degeneration occurs when parts of the retina dysfunction, largely due to genetic and/or environmental factors. There are many types of retinal degeneration, and two major ones are age-related macular degeneration and hereditary retinal dystrophies, which involved multiple and single genes, respectively.

Age-Related Macular Degeneration (AMD)

AMD is a leading cause of blindness in the developed world.[63,64] It is predicted that AMD will affect 196 million people worldwide by the year 2020 and 288 million by the year 2040.[65] At the moment, although abnormal blood vessel growth can be treated, there is no cure for the underlying disease process. Clinically, AMD may present with progressive distortion of central vision, decreasing visual acuity, and central scotomas. There are two main forms of the disease: “wet” and “dry.” The dry form of the disease, which accounts for approximately 80% of AMD, is characterized by drusen, retinal pigment epithelium abnormalities, hyperpigmentation, and atrophy.[66,67] The wet form of AMD is characterized by neovascularization and frequently develops as a complication of the dry form.[66,67] While AMD is a polygenic disease with some environmental influences, certain targets have been identified for therapy including the complement system and vascular endothelial growth factor.[68–73]

Inherited Retinal Degenerations

Inherited retinal degenerations (IRDs) comprise a group of inherited single gene disorders that cause blindness by affecting primarily the light-sensing photoreceptors in the eye and/or the retinal pigment epithelium (RPE) cells. Combined, inherited retinal degenerations have a prevalence of about one in 1490, affecting approximately two million people worldwide.[74,75] The vast majority of inherited retinal dystrophies are monogenic, meaning that they are caused by one mutation in one gene. Approximately 256 genes are known to contribute to retinal degeneration; within these genes, over 12,000 disease-causing mutations have been identified.[76–78] These diseases can have autosomal recessive, autosomal dominant, mitochondrial, or X-linked inheritance patterns. Because there are many possible genetic defects that can cause retinal degeneration, the range of phenotypic presentations of inherited retinal dystrophies is diverse. Current approaches to management include correcting refractive error with glasses, managing macular oedema, using low-vision aids, and performing cataract surgery, if necessary. While these treatments may help patients, there is still no cure for this group of diseases.

1.3 Retinal Gene Therapy

The eye has several advantages that make it suitable for receiving gene therapy. The blood-retina barrier, along with certain properties of the retinal cells, create an immune privileged environment; the immune response against foreign substrates, like gene therapy vectors, is thus relatively attenuated.[79] The blood-retinal barrier also prevents systemic distribution of substrates delivered to the eye, reducing the risk of systemic side effects.[80] Furthermore, because retinal cells do

not divide after birth, there are no new cells being created and a single treatment could thus be curative.[80] Tools like optical coherence tomography (OCT), fundus autofluorescence imaging, and microperimetry make it easy to assess and observe disease progression and efficacy of therapeutic intervention.[81] In diseases that affect both eyes equally, the eye without therapeutic intervention can serve as a control to evaluate the efficacy of new treatments.

To date, retinal gene therapy has largely involved gene augmentation. With this technique, a gene is packaged inside of a viral vector, usually an adeno-associated virus (AAV).[82] The virus transduces the target cells, and in this fashion delivers the gene. Numerous clinical trials have been initiated to treat retinal diseases with gene replacement therapy.

Leber Congenital Amaurousis Type 2 (LCA2)

In 2017, *Luxturna* (voretigene neparovovec/AAV2-hRPE65v2; Spark Therapeutics, Philadelphia, PA, USA) became the first gene therapy to gain approval from the Food and Drug Administration. Luxturna is indicated for biallelic *RPE65* mutation-associated retinal dystrophy, like Leber Congenital Amaurousis Type 2 (LCA2). *RPE65* encodes retinoid isomerase, an enzyme that is essential in the visual cycle for converting all-trans retinal into the 11-cis isomer in retinal pigment epithelium cells. Clinically, LCA2 usually manifests as severe vision loss in infants and young children.

The earliest clinical trials, in which several different vectors containing *RPE65* were administered, showed that the therapies were well-tolerated and resulted in some improvement in vision.[83–89] The most successful vector, which was the basis for the *Luxturna* therapy, was administered by Maguire et al. 2008.[90]

This vector contained *RPE65* in AAV2 combined with a surfactant to reduce vector loss during injection.[90] The three patients who initially received it showed improvement of pupillary light reflex and visual acuity six weeks after injection.[90] A follow-up of up to two years show that the 12 patients who were treated with this vector had sustained improvement in vision, including pupillary light response, and visual acuity.[91] Injections of the vector into the contralateral eye were well-tolerated and resulted in increased retinal function during the three year follow-up period.[92] A Phase III trial was initiated in 2017, in which bilateral subretinal injections of the vector were administered.[93] The success of this trial lead to FDA approval of *Luxturna* as a therapy for retinal degenerations caused by biallelic *RPE65* mutations.[93,94]

Choroideremia

Choroideremia is an X-linked retinal degeneration caused by mutations in Rab-escort protein 1 (*REP-1*), which is important for vesicle trafficking in retinal cells. Choroideremia has a prevalence of 1:50,000 and initially manifests as peripheral vision loss and night blindness, with eventual loss of visual acuity later in the disease course. Several clinical trials are investigating the therapeutic potential of delivering the Rab escort protein-1 gene in a type 2 adeno-associated viral vector. The University of Oxford and Nightstar Therapeutics (NCT01461213) have published results from a Phase I/II clinical trial, in which AAV2-*REP-1* was injected subretinally into patients with choroideremia.[95–97] The two year follow-up showed that all 14 patients in the cohort had improvement of visual acuity in the treated eye.[97] The median visual acuity gain was 4.5 letters in the treated eye while the untreated control eye had a 1.5 letter loss.[97] Six patients experienced

improvement of more than one line in the treated eye.[97] Similar results with this vector have been reported from clinical trials in Edmonton, Canada; Miami, Florida, USA; and Tubingen, Germany.[98–100] Currently, recruitment for Phase II and III clinical trials is underway (NCT03507686, NCT03496012, NCT02407678) and a long-term follow up of patients with gene therapy (NCT03584165) is in progress. Spark Therapeutics has developed another gene therapy vector for choroideremia, and a clinical trial is currently in progress.

Leber Hereditary Optic Neuropathy (LHON)

Leber Hereditary Optic Neuropathy is a mitochondrially-inherited disease that is characterized by painless vision loss. It is caused by mutations in NADH dehydrogenase subunit 4 complex I (*ND4*), which is a part of the electron transport chain. The mitochondria, however, cannot be transduced by adeno-associated viral vectors; to circumvent this issue, Wan et al. 2016 used a nuclear version of the *ND4* gene, which was then packaged within a rAAV2 for intravitreal delivery to nine patients with LHON.[101] The therapy was well-tolerated and six of the nine patients had improved visual acuity by at least 0.3 log MAR in the treated eye nine months after injection.[101] It is not known whether this improvement in vision is due to the gene therapy vector or the simply to the nature of LHON, which may at times spontaneously improve.[102,103]

Age-Related Macular Degeneration

While VEGF inhibitors have been a mainstay of wet AMD treatment for many years, the requirement for frequent injections can be burdensome. With gene

therapy, the goal is to continuously express a *VEGF* inhibitor, thus obviating the need for repeated therapeutic intervention.

Several clinical trials have demonstrated the efficacy of AAV-based gene therapy to treat AMD. Soluble fms-like tyrosine kinase-1 (*sFLT-1*) is a naturally occurring anti-angiogenic factor that can be delivered to the retina through gene therapy. Phase I and IIa studies of *sFLT-1* packaged in rAAV and delivered through subretinal injection have shown the safety and efficacy of this therapy.[104,105] 57% patients who received the viral vectors had maintenance or improvement of visual acuity while only 36% in the control group did during the 52 week follow-up period.[105] There were no major adverse events associated with the injections.[104,105]

Another approach to managing AMD has been through delivery of endostatin and angiostatin (Retino-Stat®) in a lentiviral Equine Infectious Anemia Virus (EIAV) vector. Endostatin and angiostatin have anti-angiogenic properties and are thus thought to reduce the neovascularization seen in AMD. A dose response study investigating lentiviral delivery of Retino-Stat® showed that the therapy was well-tolerated and that patients had long-term expression of endostatin and angiostatin (two patients for more than four years). [106]

While the use of anti-angiogenic therapies is beneficial for wet AMD, the dry form of the disease, which is more prevalent, requires another approach. Because overactivation of the complement system is seen in patients with dry AMD, Gyroscope Therapeutics use a vector modelled on the AAV.REP1 vector used in choroideremia trials to inhibit the complement system. A clinical trial using this vector was started in early 2019 (NCT03846193).

Retinitis Pigmentosa

The most prevalent inherited retinal dystrophy is retinitis pigmentosa, which affects approximately 1 in 4000 people.[74,75] Clinically, retinitis pigmentosa manifests with night blindness and progressive visual field loss. There are many mutations that can cause retinitis pigmentosa, and the common end result is a degeneration of first the rod and then the cone photoreceptor cells. Retinitis pigmentosa can be inherited in an autosomal recessive, autosomal dominant, multigenic, mitochondrial, or X-linked fashion.

The most common cause of X-linked retinitis pigmentosa is a mutation in the Retinitis pigmentosa GTPase regulator *(RPGR)* gene.[107] X-linked *RPGR* mutations cause a severe phenotype, characterized by early-onset vision loss which ultimately leads to legal blindness by middle age. Three phase I/II clinical trials testing X-linked *RPGR* treatment are currently underway. Biogen Inc. (formerly Nightstar Therapeutics) is delivering a human codon-optimized full-length *RPGF*^{ORF15} sequence driven by a human rhodopsin kinase promotor within an AAV8 vector.[108] Codon optimization is a strategy that can increase gene transcription and stabilize certain coding regions that are prone to mutation. Applied Genetic Technologies Corporation is also delivering a human codon-optimized *RPGR*^{ORF15} sequence driven by the human rhodopsin kinase promoter, but is instead using mutated forms of AAV2 as the delivery vector. A third clinical trial, sponsored by MeiraGTx, is using a non-codon optimized *RPGR* sequence delivered in a wild-type AAV2/5 capsid. The results of all three trials will be available soon.

Mutations in the Mer tyrosine kinase (*MERTK*), which causes autosomal recessive retinitis pigmentosa, have also been targeted by gene replacement therapy. *MERTK* is essential for phagocytosis of photoreceptor outer segments by the retinal pigment epithelium cells. A phase I clinical trial commenced in 2011 to determine the safety of delivering recombinant AAV2 expressing human *MERTK* to patients.

Other diseases

Clinical trials to treat Achromatopsia, X-linked retinoschisis, Stargardt disease, and Usher syndrome (the latter two using an equine lentivirus) are currently in progress, with results expected in the near future.

The Shortcomings of Gene Replacement Therapy

While gene augmentation is suitable for loss-of-function mutations, in which a normal copy of the gene is missing from the cells, it cannot be used to replace a large gene or treat dominant mutations that require knock-down of a pathogenic protein. Genes that exceed the viral vector coding capacity of ~4.7 kb are considered “large.” Dominant diseases exert an effect on the retina in two forms: either through a dominant negative effect or through gain-of-function. The dominant negative effect occurs when a mutated copy of a gene prevents the normal function of the wild-type copy. Gain-of-function mutations, on the other hand, give the gene a novel, deleterious, effect. A promising approach to treat diseases in which a large gene is affected or in which there is a dominant mutation is to use gene editing. This strategy can correct a pathogenic mutation at the DNA level.

Before the discovery of CRISPR/Cas9, zinc-finger nucleases (ZFN) and transcription activator-like effector nucleases (TALEN) represented the two main methods for editing DNA. ZFNs, however, have low efficiency while TALENs are difficult to deliver into cells.[82,109,110] Furthermore, because both ZFNs and TALENs rely on specific protein-DNA binding, modification of proteins is required for each target, which can be an arduous process. Using CRISPR/Cas9, which instead functions through specific RNA-DNA interactions, only requires changes to the sgRNA while the Cas9 protein remains unchanged. The relative simplicity of making sgRNA modifications, along with higher efficiency and ease of delivery, make CRISPR/Cas9 a promising tool to correct mutations.

The main delivery methods for CRISPR/Cas9 are ribonucleoproteins, electroporation, and AAVs. Ribonucleoproteins, however, have not been successfully delivered to the neuroretina and electroporation is best for use in mitotic cells, and not post-mitotic cells like those of the retina.[111,112] AAVs are an attractive vehicle for CRISPR/Cas9 because they have been used successfully and safely for gene therapy *in vivo* in humans. They also have low immunogenicity and can exhibit tropism for a variety of tissue types.[113–116] Because AAV has a small packaging capacity of ~4.7 kb, the commonly-used SpCas9 from *Streptococcus pyogenes*, which is 4.2 kb in size, cannot be incorporated into AAV along with AAV's essential regulatory elements.[28] Smaller Cas9 orthologs that can readily fit into an AAV, like the 3.2 kb SaCas9 from *Staphylococcus aureus*, could be used instead.[27]

1.4 *In Vivo* Applications of CRISPR/Cas9 to Treat Retinal Disease

The potential efficacy of CRISPR/Cas9 to treat retinal degenerations has been demonstrated in animal models for several types of retinal diseases. Proof-of-concept *in vivo* is important before approving clinical trials using CRISPR/Cas9 (Table 1).

Retinitis Pigmentosa

Out of the over 200 genes associated with retinitis pigmentosa, at least 20 of them cause an autosomal dominant form of the disease.[77] These are more prevalent as only one abnormal allele is needed for pathogenicity. Mutations in the rhodopsin (*RHO*) gene are the most common causes of autosomal dominant retinitis pigmentosa, and over 150 mutations in *RHO* have been identified.[77,117] Rhodopsin, which is expressed in the rod photoreceptors cells, is an essential component of the visual cycle and is present in large amounts in the retina; it makes up approximately 90% of the total protein in rod outer segments (OS).[118] The p.Pro23His (P23H) mutation in *RHO* is one of the most-studied causes of autosomal dominant RP.[117,119,120] This mutation leads to improper folding of rhodopsin. The misfolded protein is tagged for degradation by the ubiquitin proteosome.[121] The high number of ubiquitinated rhodopsin overwhelms the proteasome machinery, eventually causing a build-up of misfolded protein.[121] The misfolded protein is toxic and eventually leads to cell death via the unfolded protein response (UPR). This is an example of a toxic effect caused by an autosomal dominant mutation. Furthermore, wild-type rhodopsin created from the non-mutated allele can sometimes become entrapped in the mutated protein; because

the mutant is tagged for degradation, the attached wild-type protein is also broken down.[122,123] This is an example of a dominant negative effect.

CRISPR/Cas9 has been applied in animal models to treat P23H RP. Knockdown of exon 1 of the *rho* gene in transgenic mice carrying the P23H mutation was achieved using a CRISPR/Cas9 system with two sgRNAs, each targeting a different end of the exon.[124] One advantage of this approach is that it is not mutation specific, allowing it to target a wide range of mutations within exon 1.[125] The desired 24 bp deletion was present in 16% of analysed cells and a significant reduction in mutant P23H protein was observed.[124] Tsai et al. 2018 refined this protocol by using an “ablate-and-replace” system in which CRISPR/Cas9 with two sgRNAs targeting exon 1 was injected along with a wild-type copy of the *rho* gene.[125] Compared to eyes with *rho* supplementation alone, eyes that received both *rho* and the dual sgRNA CRISPR/Cas9 demonstrated a 17 to 36% increase in outer nuclear layer (ONL) thickness and improved preservation of a and b waves on electroretinography (ERG) three months after injection.[125]

Allele-specific editing of mutations in the rhodopsin gene has been achieved by several groups. Bakondi et al. 2016 targeted the *rho* gene in transgenic S334ter rats.[126] The S334ter mutation results in a novel PAM site not present on the wild-type allele, so a sgRNA that recognizes only the mutant PAM site was used.[126] Bakondi et al. 2016 had cleavage efficiencies of 33% and 36% in two rats, rescue of the photoreceptor phenotype, and a 53% higher visual acuity at post-natal day 39 compared to controls injected with non-targeting sgRNA.[126] Selective allele targeting was also achieved by Giannelli et al. 2018, who used the VQR variant of SpCas9, and a sgRNA that recognized a novel PAM site created by the P23H *rho* mutation.[127] Increased preservation of OS and ONL length was seen three months

after therapy in treated eyes compared to contralateral untreated eyes.[127] ERG readings at this time point showed rescue of a and b waves in treated eyes compared to non-treated controls.[127] Li et al. 2018 induced allele-specific editing in mice harbouring the P23H mutation by using a truncated 17 base pair crRNA along with a SpCas9 variant called VRQR.[128] The ONL thickness was significantly increased in retinas at five weeks post-therapy compared to untreated controls. MRNA transcripts of wild-type rhodopsin were increased and mutant P23H was decreased in treated heterozygous mice compared to the untreated controls.[128]

Another mutation that has been targeted by CRISPR/Cas9 is *PDE6B*, which is a cause of recessive retinitis pigmentosa. Homology-directed repair has been successfully used to treat *Pde6b* mutations. Wu et al. 2016 performed the first homology-directed repair in live mouse retinas by injecting rd1 mouse zygotes with a CRISPR/Cas9 system targeting the *Pde6b* gene and a single-stranded donor oligonucleotide to direct repair after DNA cleavage.[129] Mice with mutations in the *Pde6b* gene are preclinical models for retinitis pigmentosa and their mutation is analogous to the human *PDE6B* mutation. In this study, seven out of 11 mice had double stranded DNA cleavage, four out of 11 had homology-directed repair, and two mice had precise incorporation of the donor sequence in 35.7% and 18.8% of somatic cells, respectively.[129] At post-natal day 30, ERG responses from the two mice with donor template incorporation were comparable to those of wild-type mice.[129] Cai et al. 2019 targeted the *Pde6b* mutation by combining SpCas9/sgRNA with recombinase A (RecA), an enzyme derived from *Escherichia coli* that increases HDR efficiency.[130] This system relies on the wild-type allele to act as a template to repair the mutant allele. Although HDR is generally difficult to achieve in post-mitotic retinal cells, this CRISPR/RecA system was able to correct the

Pde6b mutation in rd1 mice, restore wild-type Pde6b protein expression, reduce photoreceptor degeneration, and improve visual function.[130]

Leber Congenital Amaurousis Type 10 (LCA10)

Another disease that has been targeted with CRISPR/Cas9 therapy is Leber Congenital Amaurousis Type 10 (LCA10), an autosomal recessive condition caused by mutations in the *CEP290* gene. *CEP290* encodes a protein that is important for normal structure and function of the photoreceptor cilia.[131,132] The most common cause of LCA10 is an adenine to guanine point mutation in intro 26 (IVS26, c.2991+1655A>G) of the *CEP290* gene.[133,134] This mutation creates a novel splice site, which causes the inclusion of an additional 128 base pair exon after RNA processing.[133] This new exon contains a premature stop codon, resulting in a truncated mRNA transcript that is prone to nonsense-mediated decay.[133]

Ruan et al. 2017 achieved deletion of mouse intro 25, which is homologous to intron 26 in humans, using SpCas9 and two guide RNAs, each flanking the target region.[135] The Cas9 and guide RNAs were delivered into wild-type mice using dual AAV5 vectors, one containing the guide RNAs and another containing the Cas9. Wild-type mice were used for this study because there are no suitable animal models for LCA10. Four weeks after therapy, next generation sequencing of four treated retinas revealed a successful editing rate of 7.5-26.4%.[135]

Maeder et al. 2019 successfully used a similar dual guide RNA approach but with the shorter SaCas9 instead of SpCas9.[46] They were able to deliver the CRISPR/Cas9 system in one AAV5 vector instead of two.[46] Maeder et al. 2019

achieved editing in humanized *CEP290* knock-in mice and in non-human primates (Figure 1.3). A productive editing rate of 60.8±30.2% was observed in mouse retinas and a 27.9±20.7% rate was seen in non-human primates.[46] Productive edits are those that remove the aberrant splice site. In mice, the observed edits persisted up to nine months after injection.[46] Recently, a clinical trial lead by EDitas Medicine has been approved to test this therapy, "EDIT-101," in human patients with LCA10. While CRISPR/Cas9 is already being used in clinical trials to treat sickle cell disease, beta thalassemia, and certain types of cancers, the EDitas clinical trials represent the first attempt to directly target mutations *in vivo* instead of using CRISPR/Cas9 to correct mutations *ex vivo*.

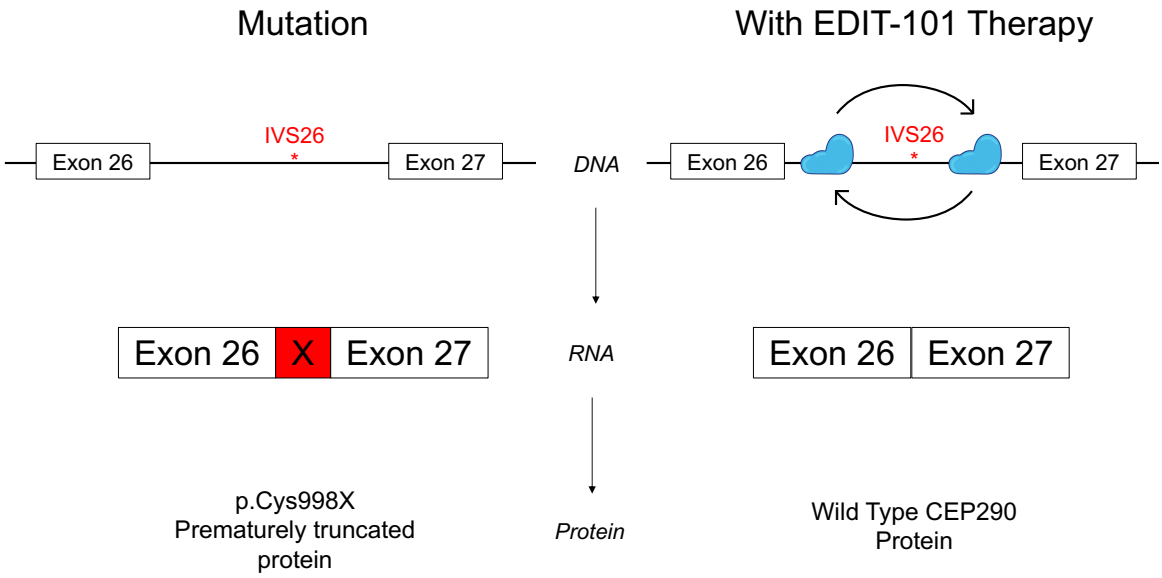

Figure 1.3. CEP290 Gene with IVS26 Mutation and CEP290 Gene Corrected Using EDIT-101 Gene Therapy (EDitas)

The left side of the figure shows the *CEP290* gene with the LCA10-causing IVS26 mutation. This mutation introduces a cryptic splice site, which results in the inclusion of a 128 base pair cryptic exon ("X") in the processed mRNA. This mRNA is translated into a non-functional protein. The EDIT-101 CRISPR/Cas9 gene therapy, shown on the right, removes the IVS26 mutation and allows for proper splicing to occur.

Image illustrated by Kella Vansgness.

Age-related Macular Degeneration (AMD)

Age-related macular degeneration (AMD) has also been a target for CRISPR/Cas9 therapies. The neovascularization observed in wet AMD is associated with overexpression of *VEGFA* and *HIF-1a*, a transcription factor that increases *VEGFA* expression.[136–138] CRISPR/Cas9-based therapies with several different Cas proteins have been used to target both *VEGFA* and *HIF-1a* in order to reduce angiogenesis.

Kim E et al. 2017 used a smaller Cas9 ortholog, cjCas9, packaged in AAV9 to target either *Vegfa* or *Hif1a*, which are associated with choroidal neovascularization in AMD.[139] Indels at both targets were observed in the retina and the retinal pigmented epithelium (20±5% and 22±3% for *Vegfa* and 58±12% and 31±2% for *Hif1a*, respectively).[139] The *Vegfa*- and *Hif1a*- targeting CRISPR/Cas9 systems reduced laser-induced choroidal neovascularization by 24±4 and 20±4 % respectively.[139] The *Hifa*-targeting system did not cause any retinal toxicity 14 months after injection based on histological and electroretinographic findings.[140]

Several studies have described SpCas9-based treatments for AMD. Kim K et al. 2017 showed that ribonucleoproteins of SpCas9 with a guide RNA targeting *Vegfa* were able to induce indels in 25±3% of retinal pigment epithelium cells three days after injection and significantly reduce choroidal neovascularization in the treated area.[112] Holmgaard et al. 2017 achieved an *in vivo* editing efficiency of up to 84% in C57BL/6J mice with subretinal delivery of lentivirus coding for SpCas9 and sgRNA targeting *Vegfa* in retinal pigment epithelium cells.[141] Huang et al. 2017 delivered CRISPR/Cas9 using recombinant AAV1, which was found preferentially to transduce pathological endothelial cells.[142] They used a sgRNA that targets *Vegfr2*.[142] In mouse models of oxygen-induced retinopathy, there was a 2% rate of indels and a 30% reduction in Vegfr2 protein compared to controls.[142] The CRISPR/Cas9 system decreased choroidal neovascularization in mice with laser-induced injury compared to controls.[142]

Koo et al. 2018 used the LbCpf1 ribonuclease delivered intravitreally in AAV to target *Vegfa* and *Hif1a* genes in mice.[143] LpCpf1 does not require a tracrRNA and relies simply on the crRNA and PAM site for targeting.[143] Mice with laser-

induced choroidal neovascularization six weeks after injection of the CRISPR system had reduced choroidal neovascularization area (42±4 % for *Vegfa*-targeting and 34±5% for *Hif1a*-targeting) compared to controls injected with aflibercept, the standard of therapy.[143] These results show the long-lasting effects of LbCpf1 to inhibit neovascularization unlike aflibercept, which has a shorter half-life.[143] Furthermore, VEGFA protein levels in the RPE were reduced compared to controls. No significant difference in scotopic and photopic responses on ERG was observed at this time point.[143]

Mutation-Independent Approaches

Because of the heterogeneity of inherited retinal disease, one approach is to target common pathways of degeneration in order to restore retinal function. Knocking down the neural retinal leucine zipper (*Nrl*), a transcription factor that is important for normal rod photoreceptor function, causes mature rods to exhibit characteristics of cones and ultimately improves retinal function. This may, for example, be a viable option for many mutations that cause retinitis pigmentosa.[144,145] Yu et al. 2017 used a dual AAV8 vector approach to deliver SpCas9 and sgRNA targeting *Nrl* into the subretinal space of Crxp-Nrl mice, which have only rod-like photoreceptors.[144] After therapeutic intervention, some photoreceptors developed a cone-like morphology.[144] Injection of this CRISPR/Cas9 system into three different mouse models of retinal disease slowed rod degeneration and improved cone function.[144] Similar findings of improved rod and cone preservation were seen after using CRISPR/Cas9 to knock down *Nrl* or *Nr2e3*.[145]

The successful use of CRISPR/Cas9 for gene editing *in vivo* shows that it is a promising tool to treat patients with a variety of retinal degenerative diseases. Optimization of CRISPR/Cas9 is an important step before it is applied in the treatment of retinal disease. The aims of the thesis are therefore the investigate further factors that my influence the efficacy of CRISPR/Cas9 for gene therapy.

Table 1. Summary of *In Vivo* Cas9 Applications to Treat Retinal Disease

Study	Disease	Gene Target	Approach	Result
Suzuki et al. 2016[33]	Retinitis Pigmentosa	*Mertk*	Insertion of exon 2 using HITI repair. Dual AAV delivery of SpCas9 to RCS rats.	Increased *Mertk* mRNA levels; ONL preservation; improved ERG b wave responses.
Latella et al. 2016[124]	Retinitis Pigmentosa	Exon 1 of *Rho* gene carrying P23H mutation	Electroporation of SpCas9 and two sgRNAs plasmids targeting different ends of exon into P23H *Rho* transgenic mice.	Desired 24 bp deletion was present in 16% of analysed cells; reduction in mutated protein.
Tsai et al. 2018[125]	Retinitis Pigmentosa	Exon 1 of *Rho* gene	Dual AAV delivery of two sgRNAs targeting different ends of exon 1, wild-type copy of *Rho* gene, and human codon-optimised SpCas9 to mice with human *Rho* mutation knock-in.	17-36% increase in ONL thickness; improved preservation of a and b waves 3 months after injection.
Bakondi et al. 2016[126]	Retinitis Pigmentosa	S334ter *Rho* mutation	Electroporation of SpCas9/sgRNA plasmids in S334ter-3 rats for allele-specific editing.	33-36% cleavage efficiency in two rats; rescue of photoreceptor phenotype; 53% higher visual acuity at PND 29.
Giannelli et al. 2018[127]	Retinitis Pigmentosa	P23H *Rho* mutation	Electroporation of plasmids or dual AAV delivery of SpCas9 VQR and sgRNA for allele-specific editing of P23H knock-in mice.	Increased OS preservation and ONL length at 3 months; rescue of a and b waves on ERG.
Li et al. 2018[128]	Retinitis Pigmentosa	P23H *Rho* mutation	Electroporation of SpCas9/sgRNA plasmids for allele-specific editing in *Rho*-P23H mice.	Significant increase in ONL thickness 5 weeks after therapy; increased mRNA transcript in treated eyes.
Wu et al. 2016[129]	Retinitis Pigmentosa	*Pde6b*	Injection of sgRNA plasmid, SpCas9 protein, and single-stranded oligonucleotide to induce homology-directed repair in rd1 mice.	7/11 mice had dsDNA cleavage; 4/11 had HDR repair; 2/11 had precise incorporation of donor sequence in 35.7% and 18.8% of somatic cells and normal ERG responses at PND 30.
Cai et al. 2019[130]	Retinitis Pigmentosa	*Pde6b*	Electroporation of SpCas9, sgRNA, ssDNA template, and RecA to induce homology-directed repair in mice with the *Pde6b rd1* mutation.	Correction of *Pde6b* mutation; restoration of Pde6b protein expression; reduction in PR degeneration; improvement in visual function.

Ruan et al. 2017[135]	LCA10	*CEP290*	Dual AAV delivery of SpCas9 and two sgRNAs flanking the target region to delete intron 25 in wild-type mice.	Successful editing rate of 7.5-26.4%.
Maeder et al. 2019[46]	LCA10	*CEP290*	Single AAV delivery of SaCas9 and two sgRNAs flanking target region to delete a mutated region in humanized *CEP290* knock-in mice and non-human primates.	Productive editing rate of 60.8±30.2% in mouse retinas and 27.9±20.7% in non-human primates.
Kim E et al. 2017 and Jo et al. 2019[139,140]	AMD	*Vegfa* and *Hif1a*	Single AAV delivery of CjCas9 and sgRNA to C57BL/6J mice.	Indels in retina and RPE (20±5% and 22±3% for *Vegfa* and 58±12% and 31±2% for *Hif1a*, respectively); reduction in CNV; no retinal toxicity 14 months after injection from *Hif1a* -targeting system.
Kim K et al. 2017[112]	AMD	*Vegfa*	Subretinal injections of SpCas9 RNP in C57BL/6J mice.	Indels in 25±3%RPE cells 3 days after injection; reduction in CNV.
Holmgaard et al. 2017[141]	AMD	*Vegfa*	Lentiviral delivery of SpCas9 in C57BL/6J mice.	*In vivo* editing efficiency up to 84% in RPE cells.
Huang et al. 2017[142]	AMD	*Vegfr2*	Dual AAV delivery of SpCas9/sgRNA to C57BL/6J mice.	2% indel rate and 30% reduction in Vegfr2 protein in mice with oxygen-induced retinopathy; decrease in CNV in mice with laser-induced injury.
Koo et al. 2018[143]	AMD	*Vegfa* and *Hif1a*	LbCpf1 delivered in AAV to C57BL/6J mice.	Reduction in CNV (42±4% for *Vegfa* target and 34±5% for *Hif1a* target); reduction in VEGFA protein levels in RPE.
Yu et al. 2017[144]	Retinitis Pigmentosa (mutation-independent)	*Nrl*	Dual AAV8 delivery of SpCas9 and sgRNA to mice.	Development of cone-like morphology in PR; slowing of rod degeneration; improved cone function.
Zhu et al. 2017[145]	Retinitis Pigmentosa (mutation-independent)	*Nrl* or *Nr2e3*	Dual AAV delivery of SpCas9/sgRNA to mice.	Development of cone-like morphology in PR; restored visual function.

1.5 Aims

1. Including the sgRNA and the Cas9 on the same AAV construct requires two independent sets of transcriptional riboprotein assemblies in very close proximity to one another to read DNA. This raises the question as to whether having the Cas9 and sgRNA on separate plus and minus DNA strands (forward/*in cis* and reverse/*in trans*) might improve editing rates. The first aim of this thesis is to determine the effects of orientation of the sgRNA and its associated promoter in SaCas9 and dSaCas-KRAB constructs in a plasmid *in vitro*.
2. Although SaCas9 is effective *in vivo*, its use in gene editing is restricted by its specific PAM site. Exploring other Cas9 proteins that have different PAM requirements would expand editing possibilities. The second aim of this thesis is to determine the feasibility of using a novel Cas9 protein, GeoCas9, for editing HEK293-*eGFP* cells *in vitro*.

2 Materials and Methods

2.1 Polymerase Chain Reaction (PCR)

PCR is used to make many copies of a target DNA region. Two primers (Sigma-Aldrich) flanking the target DNA region were designed *in silico*, following the guidelines below:

- Primers should be 18-25 bp long, so they are short enough to efficiently bind to the DNA template but long enough to ensure specificity of binding to the target region.
- The melting temperature (Tm)=40-65°C. Tm indicates the stability of primer binding to the DNA template. Both the forward and reverse primer should have a similar melting temperature.
- The GC content should be ~50%
- The last five bases at the 3' end of the primer should have G and C bases. Having more G and C bases at the 3’ end helps increase the strength and specificity of primer binding. G and C bases form three hydrogen bonds as opposed to A and T, which form two hydrogen bonds.
- Ideally, primers should not form secondary structures or dimers. If they do, then the Tm of the secondary structures should be ≤35°C.
- Long “runs,” which are repeats of the same base pair (e.g. GGGG), and “dinucleotide repeats” (e.g. ATATA) should be avoided. Both of these types of sequences make it difficult for the primer to bind to the correct region.
- If the amplicon will subsequently be digested with restriction enzymes, there should be a leader sequence present at the 5’ end of the primer. A leader sequence consists 2-6 random base pairs, which are incorporated into the final PCR product and increase restriction digest efficiency.

To reduce contamination by aerosolized PCR products from previous experiments, the bench surface and pipettes were treated with 10% bleach for 10 minutes and then cleaned with 70% IMS before setting up the PCR reaction. Aliquots of PCR grade water were UV treated for 10 minutes. KOD Hot Start DNA polymerase, 10X Buffer for KOD Hot Start DNA polymerase, 25 mM MgSO4, and dNTPs (2 mM each) were used for the PCR reaction, per the manufacturer's protocol (Novagen, EMD Chemicals, Merck, UK). KOD Hot Start DNA Polymerase is a high fidelity KOD DNA Polymerase with 3' to 5' exonuclease activity. Briefly, each reaction had a final volume of 20 µL or 50 µL and contained:

Reagent	Amount in µL
10X Buffer for KOD Hot Start DNA Polymerase	5
25 mM MgSO4	3
dNTPs (2 mM each)	5
PCR grade water	X
Sense (5') Primer (10 µM)	1.5
Anti-Sense (3') Primer (10 µM)	1.5
Template DNA	Y
KOD Hot Start DNA polymerase	1

The thermocycler conditions were:

Step	Amplicon <500 bp	Amplicon 500-1000 bp	Amplicon 1000-3000 bp	Amplicon >3000 bp
1. **Polymerase Activation**	95 °C for 2 min	95 °C for 2 min	95 °C for 2 min	95 °C for 2 min
2. **Denaturation**	95 °C for 20 s	95 °C for 20 s	95 °C for 20 s	95 °C for 20 s
3. **Annealing**	Lowest primer Tm for 10 s	Lowest primer Tm for 10 s	Lowest primer Tm for 10 s	Lowest primer Tm for 10 s
4. **Extension**	70 °C for 10 s/kb	70 °C for 15 s/kb	70 °C for 20 s/kb	70 °C for 25 s/kb

Repeat steps 2-4 for 20-40 cycles.

2.2 Restriction Digest

Restriction endonuclease digestion of PCR amplicons and plasmid vectors was performed for sticky end cloning. Reactions contained 1 µL of each restriction enzyme, 5 µL of either CutSmart buffer or 3:1 buffer, 500-1000 ng of DNA, and PCR grade water in a total reaction volume of 50 µL (New England Biolabs, UK). Samples were incubated in a water bath at the appropriate working temperature for the restriction enzymes for a period of 15 min to 1 hour. The restriction enzymes were subsequently heat inactivated when possible, per the manufacturer's instructions. When necessary, digested plasmid vectors were dephosphorylated with 1 µL of shrimp alkaline phosphatase (New England Biolabs, UK).

For samples in which two or more fragments >100 bp in length each were created, the full sample volume was loaded onto an agarose gel for gel electrophoresis separation of the DNA bands. The desired band was identified and then excised and purified using the QIAquick gel extraction kit, per the manufacturer's instructions (Qiagen, UK). Samples with only one product >100 bp in length were purified using the QIAquick PCR purification kit, per the manufacturer's protocol (Qiagen, UK).

2.3 Gel Electrophoresis

PCR products and DNA fragments digested with restriction enzymes were separated and visualized on agarose gels. Each gel consisted of 200 mL tris/acetate/EDTA (TAE) buffer (Thermo Fisher, UK), 2-3 g of agarose, and 10 µL ethidium bromide. Samples were mixed with 6x Gel Loading Dye (New England Biolabs) and loaded onto the gel. Appropriate DNA ladders were also loaded onto

the gel in order to help determine the length of the samples 1kb DNA ladder (UltraRanger, Norgen, UK) for samples >1 kb and 100 bp DNA ladder (PCRRanger, Norgen, UK) for samples <1 kb. Each gel was submerged in TAE buffer in a gel electrophoresis tank. A voltage of 100-150 V was applied for a period of 30 minutes to 2 hours. The gel was then removed from the tank and DNA bands were visualized on an ultraviolet (UV) transilluminator (U:Genious, Syngene, UK).

2.4 Annealing and Phosphorylation of Oligonucleotides

Complementary oligonucleotides were annealed and phosphorylated for subsequent ligation into digested vectors. 25 pmol of each oligonucleotide was mixed with 2 µL of 10x reaction buffer A for T4 PNK (Thermo Fisher Scientific), 2 µL of 10 mM ATP (New England Biolabs), and 1 µL of T4 PNK (Thermo Fisher Scientific) in a 20 µL reaction. The samples were placed on the thermocycler under the following conditions.

Step	Temperature	Time	Note
1.	37°C	20 minutes	
2.	75°C	10 minutes	
3.	95°C	4 minutes	
4.	95°C	1 minute	
5.	-1.5°C	1 minute	47 cycles to reach 25°C
6.	4°C	Storage	

Samples were then PCR purified using a QIAquick PCR purification kit, per the manufacturer's protocol (Qiagen, UK). The concentration and purity of each

sample was determined using the NanoDrop 1000 Spectrophotometer (Thermo Scientific, Northumberland, UK).

2.5 Ligation

Ligation of DNA fragments to create the desired plasmids was achieved using reactions with 1 µL of T4 DNA Ligase (New England Biolabs), 1 µL of 10x T4 DNA Ligase buffer (New England Biolabs), and PCR grade water in a total reaction volume of 10 µL. Different insert:vector ratios were used, depending on the size of each component. Insert mass was calculated using the NEBioCalculator (https://nebiocalculator.neb.com/#!/ligation). Samples were incubated at room temperature overnight. These ligated plasmids were subsequently transformed into XL-10 Ultracompetent *E. coli* cells (Strategene, USA). Controls for the ligation reaction included:

1. A no insert control (N), in which the vector plasmid is incubated with ligase but without insert. Positive bacterial colony growth after transformation indicates re-ligation of the vector on itself or contamination by undigested vector.
2. A ligase control (L+), in which a linearized plasmid is incubated with ligase. Presence of bacterial colonies after transformation indicates that the ligase enzyme is successfully able to catalyse the ligation reaction.
3. A no ligase control (L-), in which digested vector is not treated with ligase. Bacterial colony growth after transformation indicates presence of incompletely digested vector plasmids.

2.6 Transformation

In order to create plasmid DNA libraries, plasmids were transformed into XL-10 Ultracompetent *E. coli* cells (Strategene, USA). These cells have the Hte phenotype, which increases their transformation efficiency. They also do not contain any known restriction enzymes [*Δ(mcrA)183 Δ (mcrCB-hsdSMR-mrr)173*]. They lack endonuclease (*endA*), which enhances the quality of plasmid DNA, and are deficient in recombination (*recA*), which improves the stability of plasmids. Transformation was carried out per the manufacturer's protocol, with some modification, as follows:

1. 15 mL falcon tubes were pre-chilled on ice.
2. 50 µL of XL-10 Ultracompetent *E. coli* cells were added into each tube.
3. 3 µL of β-Mercaptoethanol was added to each tube. β-Mercaptoethanol can improve transformation efficiency.
4. Cells were gently mixed with a pipette every 2 minutes for 10 minutes total.
5. ~1 ng of DNA was added (~4 uL of ligated plasmid). 1 µL of the positive control pUC18 plasmid included in the kit was always added to one tube in order to confirm the success of the transformation protocol.
6. Cells were incubated on ice for 30 minutes.
7. Each tube was heat pulsed in a water bath at 42°C and then chilled on ice for 2 minutes.
8. 450 µL of SOC broth was added to each tube.
9. The tubes were incubated at 37°C and 250 rpm for 1 hour.
10. 200 µL of the transformation mixture was placed onto LB agar plates containing ampicillin. Because all transformed plasmids have an ampicillin

resistance gene, only the bacteria with the desired plasmid will be able to grow on these plates.

11. Plates were incubated at 37°C for at least 17 hours. Plates were then stored at 4°C.
12. Success of the transformation reaction was determined by observing colony growth on the positive control (pUC18) plate.

2.7 Plasmid Isolation

Plasmids were isolated using the Zyppy Plasmid Miniprep Kit (Zymo Research Corp) and the EndoFree Plasmid Maxi Kit (Qiagen, UK).

Miniprep

After transformation of plasmid DNA, bacterial colonies were picked from LB agar plates using an autoclaved cocktail stick or a sterile filtered 1000 µL pipette tip, and placed into a sterile falcon tube containing 3 mL of autoclaved luria broth and 100 µg/mL of ampicillin. These picked colonies were incubated in a shaker set to a temperature of 37°C and 250 rpm for 16-18 hours. Plasmids were isolated from the incubated culture using a Zyppy Plasmid Miniprep Kit (Zymo Research Corp), per the manufacturer's instructions, and eluted into 30 µL of Zyppy Elution Buffer. Plasmid concentration and quality was determined using a NanoDrop 1000 Spectrophotometer (Thermo Scientific, Northumberland, UK). Plasmids were then stored at -20°C. To confirm that the plasmids had been successfully ligated, miniprepped plasmids were tested with restriction digest and/or commercial Sanger sequencing of the ligated region (Eurofins Genomics, EU or Source Bioscience, UK).

Maxiprep

Once the correct DNA sequence was confirmed, colonies were either transformed again or were picked directly from the luria broth (LB) agar plates used for the miniprep preparation. Picked colonies were placed into a sterile falcon tube containing 3 mL of autoclaved luria broth and 100 µg/mL of ampicillin to make a starter culture. These cultures were incubated at 37°C or, for plasmids containing the KRAB sequence, at 30°C in a shaker set to 250 rpm for 8 hours. 400 µL of the starter culture was then added to 200 mL of autoclaved luria broth with 400 µL of 100 µg/mL ampicillin in a 1 L beaker and incubated for 16-18 hours (overnight) at either 37°C or 30°C in a shaker set to 250 rpm. If necessary, 250 µL of this overnight culture was combined with 250 µL of 50% glycerol and stored at -80°C for future use. These glycerol stocks could be used to make a starter culture for subsequent plasmid preparations. Plasmids were then isolated from the remaining overnight culture using an endotoxin-free maxiprep kit (EndoFree Plasmid Maxi Kit, Qiagen, UK) and then precipitated using PureLink HiPure Precipitator Modules, per the manufacturers' protocol (Thermo Fisher Scientific). Plasmids were eluted in 750 µL of TE Buffer (10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA) (Qiagen, UK) and the correct DNA sequence was confirmed using commercial Sanger sequencing (Eurofins Genomics, EU or Source Bioscience, UK).

2.8 Cell Culture

Cell culture was conducted in Class II sterile positive pressure fume hoods using aseptic technique. All culture media [Dulbecco's Modified Eagle Medium supplemented with 10% foetal bovine serum (FBS), 2 nM L-glutamine, 100 U/mL

penicillin, 1000 ug/mL streptomycin, 1% non-essential amino acids, and phenol red pH indicator (Gibco, Thermo Fisher Scientific, UK)] and PBS were pre-warmed in a 37°C water bath. *In vitro* experiments were conducted on the immortal Human Embryonic Kidney 293 cell line that stably expresses enhanced green fluorescent protein (HEK293-*eGFP*). Cells were maintained in a tissue culture incubator at 37°C with 5% CO2.

Resuscitation of Cells

HEK293-*eGFP* cells in cryovials were stored in liquid nitrogen. The vials were thawed in a 37°C water bath for approximately 1-2 minutes. A p1000 pipette was used to mix and transfer the entire contents of the vial to a 15 mL falcon tube with 3 mL of pre-warmed media. This cell suspension was then spun for 5 minutes at 300xg. The supernatant was removed and 2 mL of pre-warmed media was used to resuspend the pellet. The total cell suspension was then seeded into a T25 flask with 4 mL of pre-warmed media. Cells were passaged at least three times before being used in experiments.

Passage

Cells were passaged into T75 flasks two times per week when they were 70-90% confluent. For this procedure, old tissue culture medium was removed and discarded. Cells were then washed with 5 mL of pre-warmed PBS. 3 mL of TrypLE^{TM} Express solution (Gibco, Thermo Fisher Scientific, UK) was applied and the flask was incubated at 37°C in the tissue culture incubator. The flasks were then removed from the incubator and gently tapped to help dissociate the cells. 7

mL of pre-warmed media was added to each flask; this mixture was pipetted up and down to ensure the formation of a single cell suspension. 1 mL of this cell suspension was added into a new T75 flask with 10 mL of pre-warmed media. The new T75 flasks were then stored in the tissue culture incubator. The remaining cell suspension was either discarded or used for seeding.

Seeding

The cell suspension created during passaging was seeded onto tissue culture plates. First, the concentration of cells in the suspension was determined either through counting cells on a haemocytometer or by using a TC20 Automated Cell Counter (Bio-Rad). For counting on the haemocytometer, 10 µL of cell suspension was diluted 1:2 in 10% trypan blue. 10 µL of this mixture was loaded onto a haemocytometer and the live cell count was used to estimate the concentration. For the TC20 Automated Cell Counter (Bio-Rad), 10 µL of cell suspension was diluted 1:2 in trypan blue (neat). 10 µL of this mixture was loaded onto the slide and inserted into the cell counter.

Based on the determined concentration, the appropriate amount of pre-warmed media was added to dilute the cell suspension to $5x10^5$ cells/mL. 24-well or 12-well tissue culture plates were seeded with 0.5 mL and 1 mL of this suspension, respectively. The plates were then stored in the tissue culture incubator.

2.9 Transfection

When HEK293-*eGFP* cells were ~70-90% confluent, which usually occurred 24 hours after seeding, they were transfected with plasmid DNA. *Trans*IT® LT-1 reagent (Mirus Bio, USA) and Opti-MEM™ serum-free culture media (Thermo Fisher Scientific, UK) were brought to room temperature prior to transfection. *Trans*IT® LT-1 was briefly vortexed. A master mix of *Trans*IT® LT-1 and Opti-MEM™ serum-free culture media was created, per the manufacturer's instructions (3 µL *Trans*IT® LT-1/1 ug DNA; 50 µL serum-free culture media/1.9 cm^2 well surface area), and incubated at room temperature for 5 minutes. This master mix was then added to the plasmid DNA and incubated at room temperature for 15-20 minutes. The *Trans*IT® LT-1: serum-free culture media: DNA complexes were then added to each well in a drop-wise fashion and the tissue culture plate was rocked back and forth to ensure even distribution of plasmid DNA. The plates were returned to the tissue culture incubator.

2.10 Harvesting

Cells were harvested on ice 48 hours after transfection. For 24 well plates, wells were washed with 500 µL of chilled PBS. 500 µL of chilled PBS was then added to each well and pipetted up and down to create a cell suspension. This suspension was transferred to a 1.5 mL Eppendorf tube. 500 µL of chilled PBS was used to wash each well again and was added to the appropriate 1.5 mL Eppendorf tube. The tubes were spun at 500xg for 10 minutes at 4°C. The supernatant was removed and the pellet was resuspended in 500 µL of chilled PBS. The tubes were spun at 500 g for 10 minutes at 4°C. The supernatant was removed and the pellets were stored at -80°C.

For cells harvested from 12-well plates, the procedure described above was used. The only modification is that a volume of 1000 µL of PBS was used instead of 500 µL for all steps except the second wash with PBS, in which a volume of 500 µL was used.

2.11 DNA and RNA Extraction

DNA was extracted from harvested cells using the QIAamp DNA Minikit and RNA was extracted from harvested cells using the Qiagen miRNeasy Mini Kit, per the manufacturer's protocols (Qiagen, UK). Extracted DNA was eluted in Buffer AE and RNA in RNase-free water. DNA and RNA concentration and purity were determined using a NanoDrop 1000 Spectrophotometer (Thermo Scientific, Northumberland, UK). DNA samples were stored at -20°C while RNA was stored at -80°C.

2.12 Complementary DNA (cDNA) Synthesis

cDNA synthesis from RNA is necessary because DNA is much more stable than RNA and can be used for further experiments such as quantitative real time PCR. 1 µg of RNA extracted from cells was used for cDNA synthesis with the SuperScript III kit (Invitrogen, UK), per the manufacturer's protocol. Oligo DTs, which anneal to poly(A) tails of eukaryotic mRNA, were used to copy Cas9 RNA (SaCas9, dSaCas9, and GeoCas9). Because the sgRNA component does not contain a poly(A) tail, gene specific primers were used to amplify the sgRNAs. Each primer was designed to bind to the sgRNA scaffold.

2.13 Real time quantitative polymerase chain reaction (qPCR)

qPCR is used to quantify gene expression levels from cDNA. The TaqMan qPCR assay was used for all experiments to quantify cDNA levels (Thermo Fisher Scientific). TaqMan probes for Cas9 and sgRNA detection were designed using the custom TaqMan Assay Design Tool (https://www.thermofisher.com/order/custom-genomic-products/tools/gene-expression/). TaqMan probes for quantification of reference “housekeeping” genes ACTB and GAPDH were purchased (Thermo Fisher Scientific).

To reduce contamination by aerosolized plasmid DNA from previous experiments, the bench surface and pipettes were treated with 10% bleach for 10 minutes and then cleaned with 70% industrial methylated spirit before setting up the PCR reaction. Aliquots of PCR grade water were UV treated for 10 minutes. The reagents used were TaqMan™ Fast Universal Master Mix (2X) (Thermo Fisher Scientific), Betaine (5M) (Sigma Aldrich), and the specific TaqMan probes (Thermo Fisher Scientific). Experimental cDNA was diluted 1:10. For the standard curves, serial ten-fold dilutions of the cDNA template or plasmid DNA were created. All samples were set up in triplicate. When appropriate, “no template” controls, in which PCR grade water was used in place of cDNA, were included to show the level of background nucleic acid contamination. A “no reverse transcriptase” control was included to determine the amount of DNA contamination in the original RNA sample. The assays were set up on ice in 96-well hard-shell qPCR plates, with the following components in each reaction.

Component	Amount in µL
TaqMan Master Mix	10
Taqman Assay (20x)	1
Betaine 5M (only for sgRNA quantification)	4
PCR grade water (ACTB, GFP, and Cas9 reactions)	7
PCR grade water (sgRNA reactions)	3
cDNA	2

qPCR plates were loaded onto the CFX Connect™ real time PCR detection thermal cycler (Bio-Rad, UK) and the following conditions were used:

Step	Temp (°C)	Time	Cycles
Heat activation	95	20s	
Denaturation	95	3s	40x
Annealing and Extension	60	30s	

Analysis of qPCR results was conducted as follows[146]:

1. The arithmetic mean of each triplicate was calculated to determine the Ct(sample) and Ct(housekeeping).
2. All of the values determined in step 1 were transformed to a non-log scale using 2^{-Ct}. This is the R_0 value.
3. In the cases where two "housekeeping" genes were used, $R_{0(HKMean)}$ was determined where $R_{0(HKMean)}$= Geometric mean of $R_{0(HK1)}$ and $R_{0(HK2)}$.
4. $R_{0(sample)}/R_{0(HK)}$ was calculated for the experiments in which one housekeeping gene was used. $R_{0(sample)}/R_{0(HKMEAN)}$ was calculated for the experiments in which two housekeeping genes were used.
5. Analysis was performed on the values determined in step 4.

6. For the graphs of the qPCR data:
    a. A reference sample was chosen and the $R_{0(reference)}$ was calculated. The $R_{0(reference)}$ is the arithmetic mean of the $R_{0(sample)}$ values from the reference.
    b. The fold change was determined by dividing the $R_{0(sample)}$ values determined in step 2 by the $R_{0(reference)}$ determined in step 6a.
    c. The qPCR graph contains the values determined in step 6b.

3 Comparison of a CRISPR/Cas9 System Arranged *in cis* or *in trans* on an AAV Transgene

3.1 Introduction

Intracellular Cas9 and sgRNA levels

Optimizing Cas9 and sgRNA levels in cells is a crucial step to improving gene editing and reducing off-target effects of CRISPR/Cas9 systems. It has been shown that editing efficiency improves and specificity decreases when higher doses of sgRNA and SpCas9 plasmid DNA at a 1:1 molar ratio are used in HEK293FT cells.[37] While low concentrations of Cas9 and sgRNA may not be effective, it appears that high concentrations should also be avoided because they can increase the likelihood of off-target editing.[147,148]

Varying the ratios of sgRNA and Cas9 can further improve efficiency. In Drosophila, maximum editing occurs with mid-range sgRNA doses while altering Cas9 levels does not greatly change on-target editing efficiency.[149] One study found that changing the ratio of AAV-delivered sgRNA:Cas9 from 1:1 to 10:1 resulted in a 10-fold increase in editing efficiency *in vivo* in mice with a mutation in the Duchenne muscular dystrophy gene.[150] While low sgRNA doses may not produce efficient editing, it is important to note that delivering too much sgRNA may be deleterious since elevated levels of exogenous RNA are known to be fatal in mice.[151]

Although we do not yet know the ideal sgRNA and Cas9 doses for the treatment of retinal degenerative disease, it is still essential to investigate factors that influence sgRNA and Cas9 expression levels in the cell. One important element that may influence transcription of the CRISPR/Cas9 elements is promoter orientation, mainly through the mechanisms of transcriptional interference, DNA topology-induced transcriptional regulation, and antisense RNA interference.

Mechanisms that Affect Transcription from Multiple Promoters

3.1.1.1 Transcriptional Interference

RNA polymerases are responsible for transcribing DNA into RNA; they initiate transcription by binding to DNA regions known as promoters. Transcriptional interference occurs when one transcribing RNA polymerase directly impedes another transcribing RNA polymerase.[152] Transcriptional interference can occur between two promoters that are *in cis*, *in trans,* or divergent (facing away from each other) (Figure 3.1). The mechanisms of transcriptional interference are listed below:

- *Sitting Duck:* In the sitting duck model, there is a "strong" promoter located upstream and convergent to a "weak" promoter.[152] RNA polymerases rapidly transcribe DNA from the strong promoter but are slow to initiate transcription from the weak promoter. When the RNA polymerases are bound to the weak promoter but have not yet started transcription, they are considered to be "sitting duck." These sitting duck RNA polymerases can easily be dislodged by the RNA polymerases arriving from the strong upstream promoter.[153]
- *Collision:* Collision between two RNA polymerases was first described by Prescott and Proudfoot, who observed this phenomenon in *Saccharomyces cerevisiae*.[154–157] Sometimes the collided RNA polymerases bypass each other and in other cases, they bind and form stable complexes.[158,159]
- *Occlusion:* In the occlusion model, one transcribing RNA polymerase unit prevents binding of a second RNA polymerase to a downstream promoter as it is passing over this downstream promoter.[160]

- *Promoter competition:* Promoter competition occurs when one RNA polymerase unit has the ability to bind to two or more promoters.[152] These promoters are thus in competition for the same RNA polymerase. Promoter competition occurs when promoters overlap, are located close to each other, or share the same enhancer site.[152]
- *Roadblock:* The roadblock mechanism occurs when a protein binds to a promoter and prevents RNA polymerase from initiating transcription.[152]

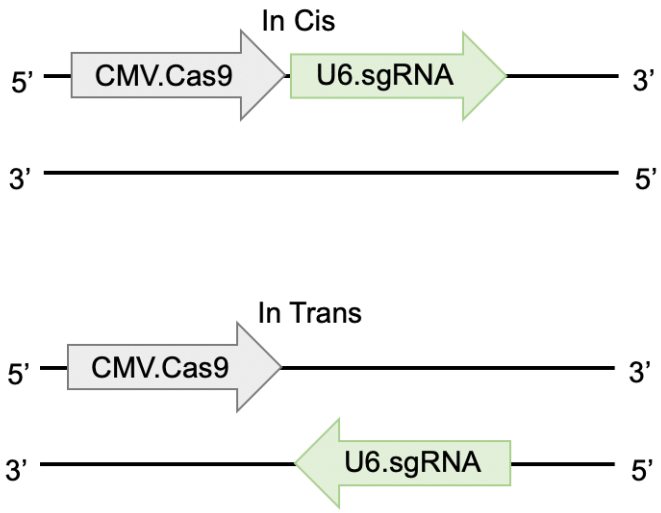

Figure 3.1 Examples of *Cis* (forward) and *Trans* (reverse) Promoter Orientations

Top: The Cas9 driven by the CMV promoter is oriented *in cis* to the sgRNA driven by the U6 promoter.

Bottom: The Cas9 driven by the CMV promoter is oriented *in trans* to the sgRNA driven by the U6 promoter.

Image illustrated by Kella Vansgness.

In theory, transcriptional interference may influence the efficacy of CRISPR/Cas9 gene therapy. The CRISPR/Cas9 system can be delivered into

target cells as a transgene, which is a single-stranded DNA molecule contained within an adeno-associated virus (AAV) vector. This transgene contains a polymerase III promoter that drives the transcription of the sgRNA and a polymerase II promoter that drives Cas9 expression. Transcriptional interference between these two promoters can alter the total amount of sgRNA and Cas9 that is produced from the transgene, which may in turn affect the efficiency of the CRISPR/Cas9 gene editing system.

3.1.1.2 DNA Topology

DNA topology may influence transcription from different promoters. RNA polymerases induce torque on the DNA strand during transcription; this torsional force can be propagated along the DNA both up- and down-stream of the RNA polymerase machinery.[161–163] The forces generated by RNA polymerases can change the three-dimensional conformation of DNA, which can in turn influence the initiation and rate of transcription from adjacent promoters.[161–163]

Because the Cas9 and sgRNA components are transcribed by different RNA polymerases simultaneously, it is possible that one transcribing RNA polymerase may influence transcription by the other RNA polymerase by affecting the topology of the adjacent DNA.

3.1.1.3 Antisense RNA Interference

For constructs arranged *in trans*, in addition to the effects of transcriptional interference, antisense RNA interference may occur; in this case, an antisense RNA transcript binds to its complementary RNA sequence, thus promoting sense RNA decay or preventing the complementary sequence from being translated into

protein.[154,164] For sgRNA and Cas9 genes oriented *in trans*, antisense RNA interference may affect total levels of sgRNA, Cas9, or both.

Effects of Promoter Orientation on Transcription

The effects of promoter orientation on transcription are variable, and seem to differ based on the specific system that is being assessed. Under certain conditions, promoters oriented *in cis* reduce transcription the most. Kholod and Mustelin 2001 tested artificial constructs containing GFP and luciferase, each driven by a promoter of similar strength, in plant protoplasts.[165] When genes were oriented *in cis*, expression of the downstream gene was reduced by 80%.[165] For constructs *in trans*, the expression of the upstream gene was reduced 53%.[165] Divergent promoters produced no change in gene expression nor was any interference seen on the upstream gene for orientations *in cis*.

Another study found that promoters oriented in *trans* have the largest effect on expression. Eszterhas et al. 2002 tested different orientations of two reporter genes, GFP and YFP, in two different genomic regions of mouse erythroleukemia cells.[166] Both reporters were driven by CMV promoters and contained strong polyadenylation transcriptional termination signals. Eszterhas et al. 2002 found that constructs *in trans* produced the greatest reduction in transcription.[166]

In some systems, it appears that promoter orientation has no effect on gene transcription. Bae et al. 2008 tested the effects of *in cis, in trans*, and divergent orientations of XYL1 and XYL2 genes in plasmids expressed in *Saccharomyces cerevisiae* but did not find any significant differences in gene expression between the three orientations.[167]

3.2 Aims

The aim of this chapter is to assess the effects of U6.sgRNA orientation on sgRNA and Cas9 levels in *Staphylococcus aureus* Cas9 constructs and to test the effects of U6.sgRNA orientation on sgRNA, Cas9, and *eGFP* expression in deactivated *Staphylococcus aureus* Cas9-KRAB constructs (Figure 3.2).

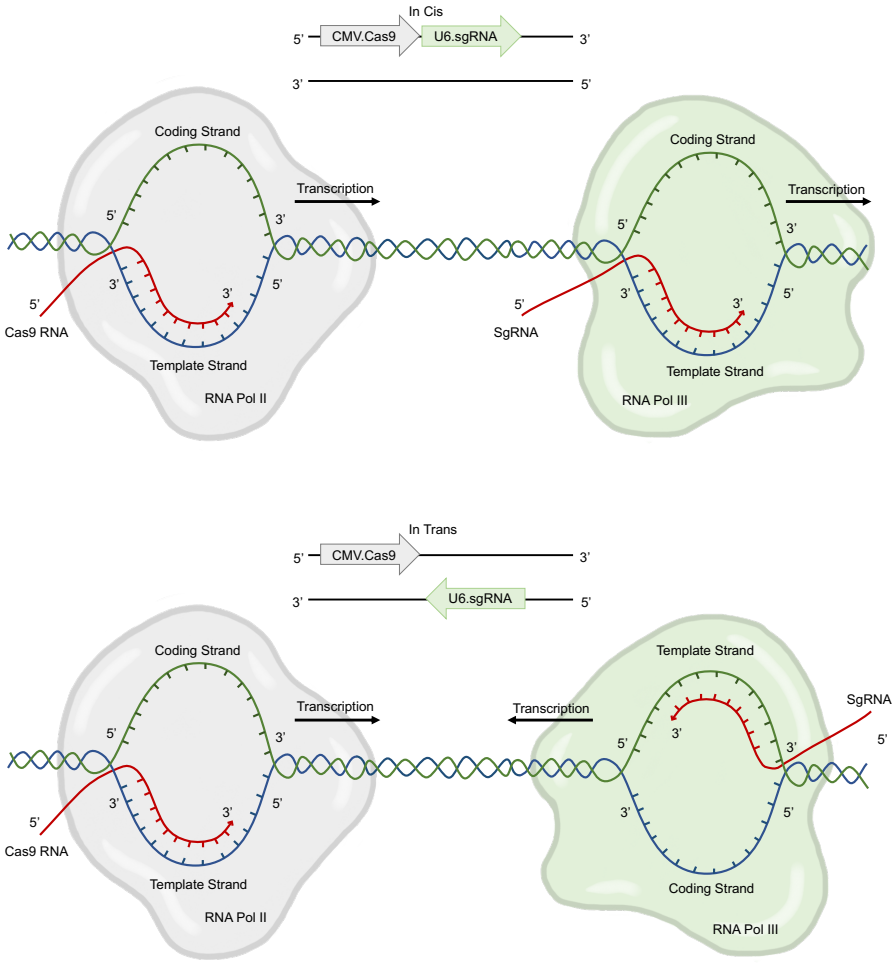

Figure 3.2. Schematic of of Forward/*Cis* (top) and Reverse/*Trans* (bottom) Orientation of U6.sgRNA in CRISPR/Cas9 Plasmids

The CMV.Cas9 is transcribed by RNA Polymerase II (grey) and the U6.sgRNA is transcribed by RNA Polymerase III (green).
Image illustrated by Kella Vansgness.

3.3 Materials and Methods

All plasmids used in this study are described in Table 2, and plasmid maps can be found in Figure 3.4 and Figure 3.5.

Table 2. List of Plasmids Used in U6.sgRNA Orientation Study

Plasmid Code	Cas9	Repressor	Puromycin Resistance	sgRNA Target	sgRNA Orientation
RHO-F	SaCas9	-	No	Rhodopsin	Forward (in *cis*)
RHO-R	SaCas9	-	No	Rhodopsin	Reverse (in *trans*)
VEGFA-F	SaCas9	-	No	*VEGFA*	Forward (in *cis*)
VEGFA-R	SaCas9	-	No	*VEGFA*	Reverse (in *trans*)
NTC-F	dSaCas9	KRAB	Yes	Non-targeting control	Forward (in *cis*)
NTC-R	dSaCas9	KRAB	Yes	Non-targeting control	Reverse (in *trans*)
13-F	dSaCas9	KRAB	Yes	*eGFP*	Forward (in *cis*)
13-R	dSaCas9	KRAB	Yes	*eGFP*	Reverse (in *trans*)
15-F	dSaCas9	KRAB	Yes	*eGFP*	Forward (in *cis*)
15-R	dSaCas9	KRAB	Yes	*eGFP*	Reverse (in *trans*)

Cloning dSaCas9-KRAB with U6.sgRNA in Trans (NTC-R, 13-R, 15-R)

Plasmids NTC-F, 13-F, and 15-F, which were previously designed, were used as templates for PCR amplification of the U6.sgRNA region with primers NotI_hU6_F2 and KpnI_hU6_R1 (Table 5, Appendix). Sequences 13 and 15 target enhanced green fluorescent protein (*eGFP*) in HEK293-*eGFP* cells and NTC is a non-targeting control (Table 4, Appendix). All three sgRNA variants are preceded by a U6 promoter sequence. A plasmid containing dSaCas9, a KRAB repressor, puromycin resistance, and ampicillin resistance genes was used as a vector for insertion of these three sgRNA variants. The cloning strategy is outlined in Figure 3.3.

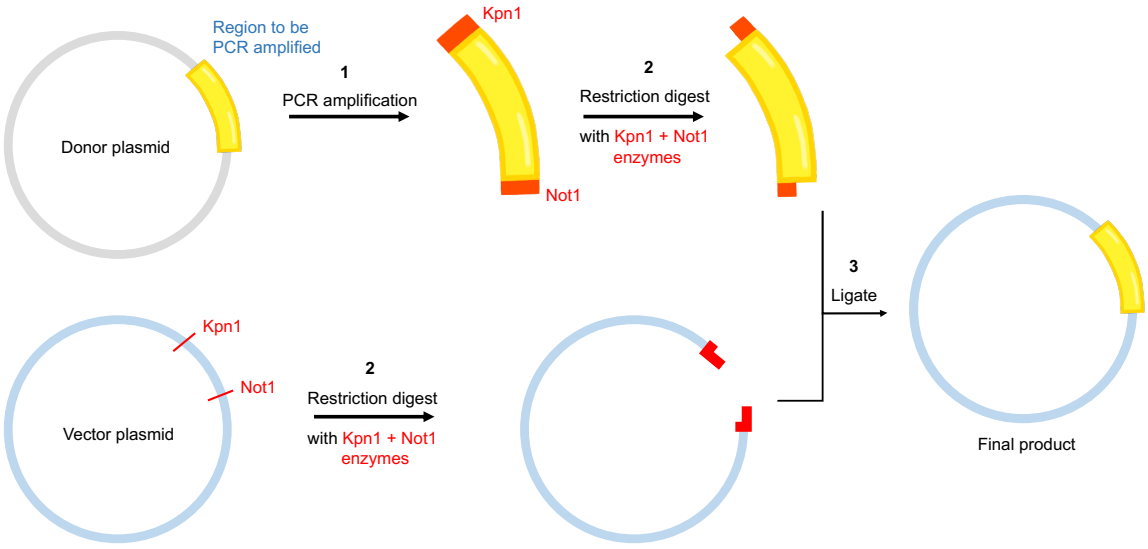

Figure 3.3. Schematic of Plasmid Cloning Strategy

1. The U6.sgRNA (yellow region in donor plasmid) was PCR amplified with primers that added KpnI and NotI restriction sites (red areas) onto the ends of the PCR product.
2. Both the vector plasmid and the PCR amplified U6.sgRNA were digested with KpnI and NotI restriction enzymes.
3. The digested vector and U6.sgRNA were ligated to produce the final product.
   Image illustrated by Kella Vansgness.

After PCR amplification of the U6.sgRNA (Figure 7.1, Appendix), a small aliquot of the PCR product was tested for appropriate size and amplification using gel electrophoresis. Once the correct size of the amplicons was confirmed, the remaining product was PCR purified, as outlined in chapter 2.

Subsequently, both the PCR purified U6.sgRNA and the vector plasmid were digested with KpnI and NotI restriction enzymes (Figure 7.2, Appendix). The digested U6.sgRNA was PCR purified again while the digested vector plasmid was loaded onto a 1% agarose gel for electrophoresis and the appropriately-sized fragment was gel extracted. The digested U6.sgRNA and plasmid vector were

ligated using 30 ng of the vector plasmid and a 10:1 ratio of U6.sgRNA insert:vector following the protocol outlined in chapter 2 (Figures 7.3 and 7.4, Appendix).

All plasmids (NTC-F, NTC-R, 13-F, 13-R, 15-F, and 15-R) were transformed into *E.coli* cells and plated onto LB agar plates with ampicillin. Colonies were subsequently selected for miniprep. Miniprepped plasmids were sequenced to confirm correct orientation of the U6.sgRNA region. These miniprepped colonies were either re-transformed or colonies were selected from the original LB agar ampicillin plates for maxiprep (Figure 7.5, Appendix). Restriction digests using the KpnI and NotI enzymes were performed prior to sequencing to confirm the presence of the restriction sites. Plasmids were also subjected to restriction digest with the XmaI enzyme in order to confirm the presence of inverted terminal repeats (ITRs) (Figure 7.6, Appendix). ITRs are difficult to sequence as they form strong secondary structures. Maxiprepped plasmids were then fully sequenced prior to being used *in vitro*.

In Vitro Testing of Plasmids for the DSaCas9-KRAB Study

HEK293-*eGFP* cells were seeded onto 24 well plates (Corning) (Figure 7.7, Appendix). Plasmids were transfected 24 hours after seeding, as described in chapter 2. Twenty-four hours after transfection, the cells were treated with puromycin in order to select for cells that had successfully been transfected with the plasmid. Cells were harvested 48 hours after puromycin treatment, as described in chapter 2.

RNA Extraction, CDNA Synthesis, and QPCR

Harvested cell pellets were used for RNA extraction and subsequent cDNA synthesis with primer sgRNA_R1 (Table 7, Appendix). For qPCR, a single reference gene, beta-actin (ACTB), was used for normalization for the dSaCas9-KRAB study. Two reference genes, ACTB and GAPDH, were used in the SaCas9 study. The sequences of the TaqMan Probes for the sgRNA and the SaCas9 can be found in Table 9 of the Appendix. Two-way analysis of variance (ANOVA) with Sidak's post-test for multiple comparisons was performed. The independent variables were: *orientation of U6.sgRNA* (forward v reverse) and *sgRNA identity* (NTC, 13, 15, RHO, or VEGFA). A p-value ≤ 0.05 was deemed to be significant.

For the SaCas9 study, a total of six biological replicates were included. Two replicates were completely removed from analysis: one of the VEGFA-F-transfected replicates due to failure of the RNA extraction and one of the VEGFA-R-transfected replicates due to failure of the cDNA synthesis reaction. The dSaCas9-KRAB study had a total of eight biological replicates. One replicate was completely removed from the analysis because of highly variable ACTB values (standard deviation >0.3), which precluded analysis of *eGFP*, sgRNA, and Cas9 levels. Another was removed from only the sgRNA analysis due to highly variable sgRNA levels (standard deviation >0.3).

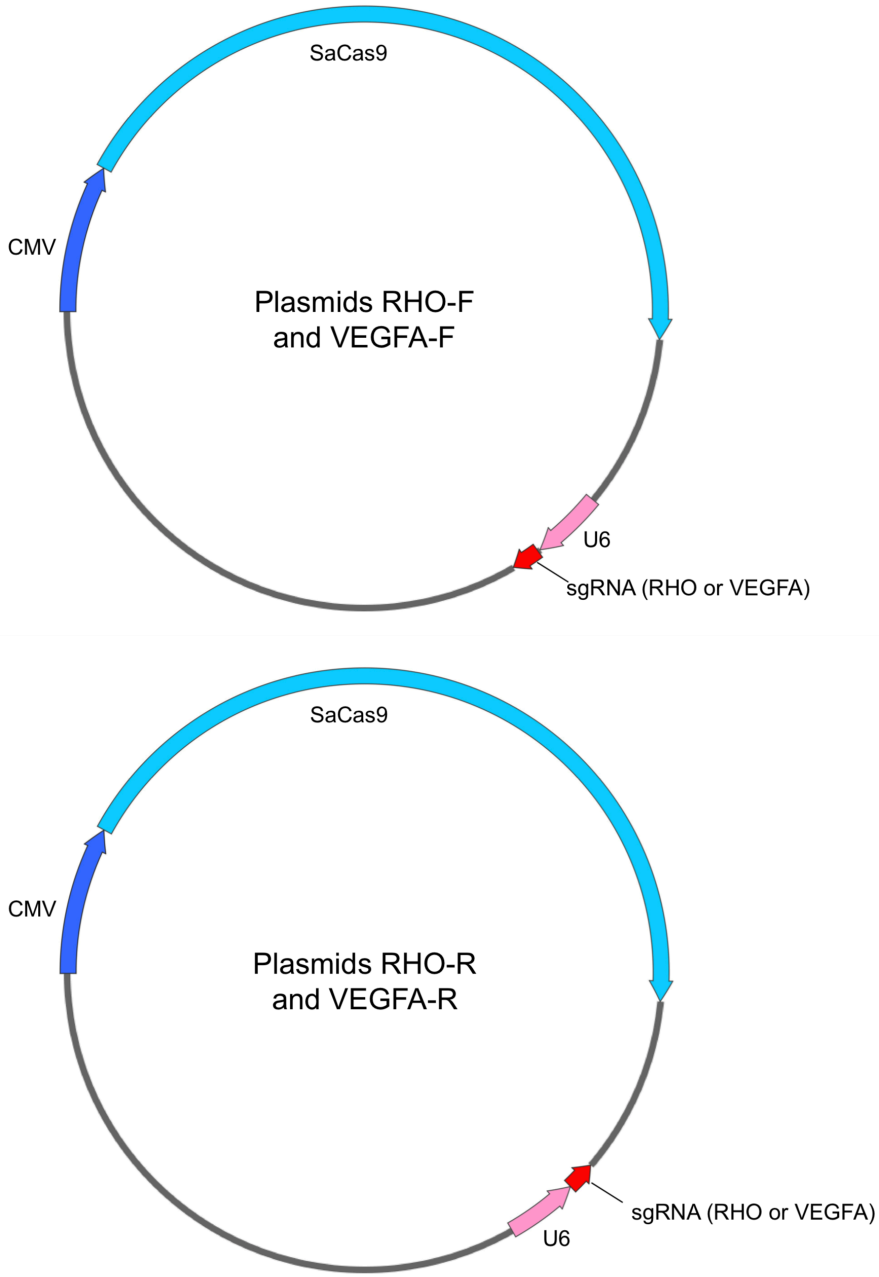

Figure 3.4 Plasmid Maps for the SaCas9 Study

The plasmid map on the top depicts the structure of plasmids RHO-F and VEGFA-F. The plasmid map on the bottom depicts the structure of plasmids RHO-R and VEGFA-R.

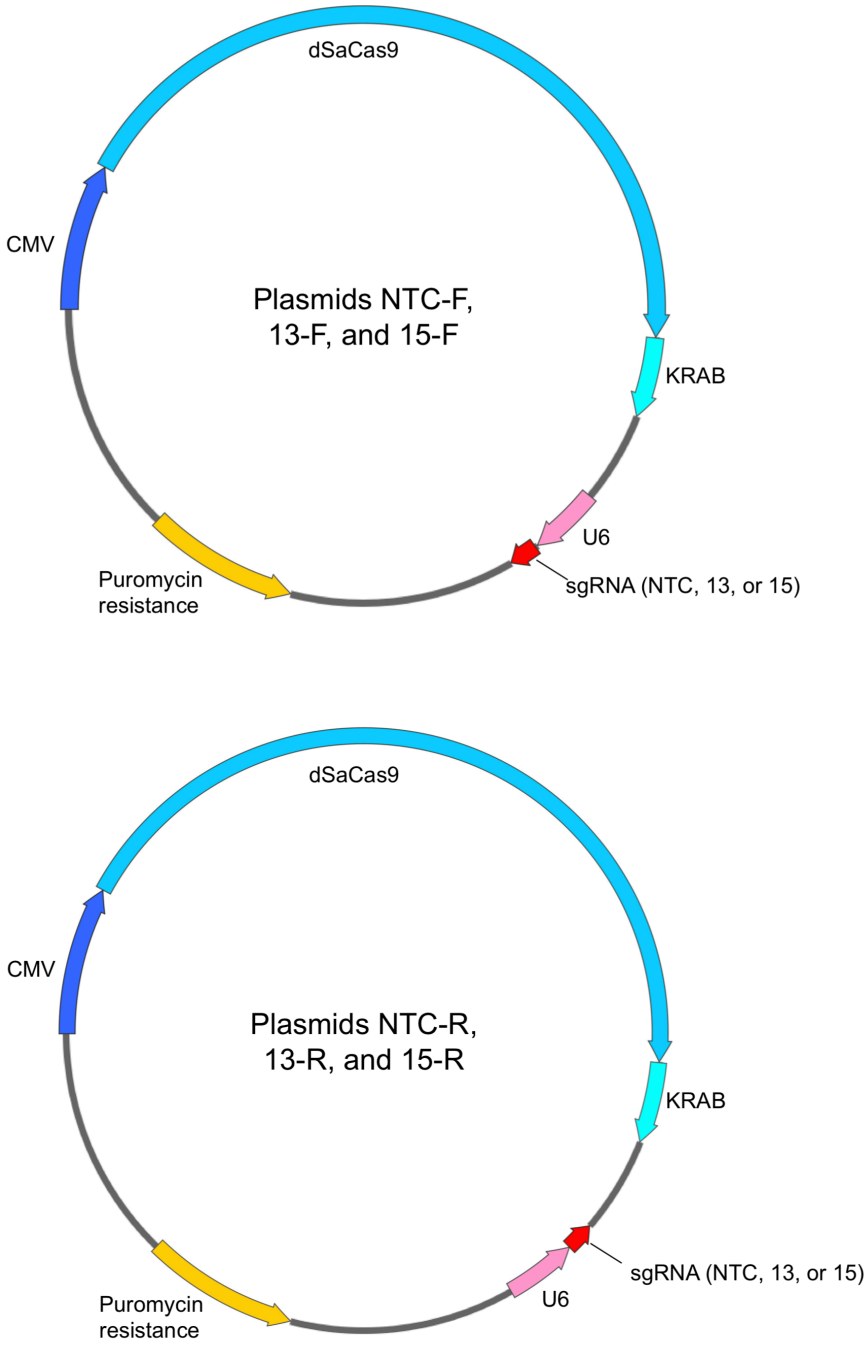

Figure 3.5 Plasmid Maps for the DSaCas9-KRAB Study

The plasmid map on the top depicts the structure of plasmids NTC-F, 13-F, and 15-F. The plasmid map on the bottom depicts the structure of plasmids NTC-R, 13-R and 15-R.

3.4 Results

Standard Curves for SgRNA and SaCas9 Taqman Probes in HEK293-eGFP Cells Transfected with RHO-R and VEGFA-R Plasmids

Taqman qPCR probe efficiency was tested using the standard curve method. The RHO-R and VEGFA-R plasmids were used for these experiments. For the RHO-R plasmid, the efficiency of the sgRNA Taqman probe was 78.4% and the R^2 was 0.983 (Figure 3.6). The efficiency of the SaCas9 Taqman probe in cells transfected with this plasmid was 79.0% and the R^2 was 0.999. For the VEGFA-R plasmid, the efficiency of the sgRNA Taqman probe was 86.1% and the R^2 was 0.997. The efficiency of the SaCas9 Taqman probe was 91.1% and the R^2 was 0.998 (Figure 3.6).

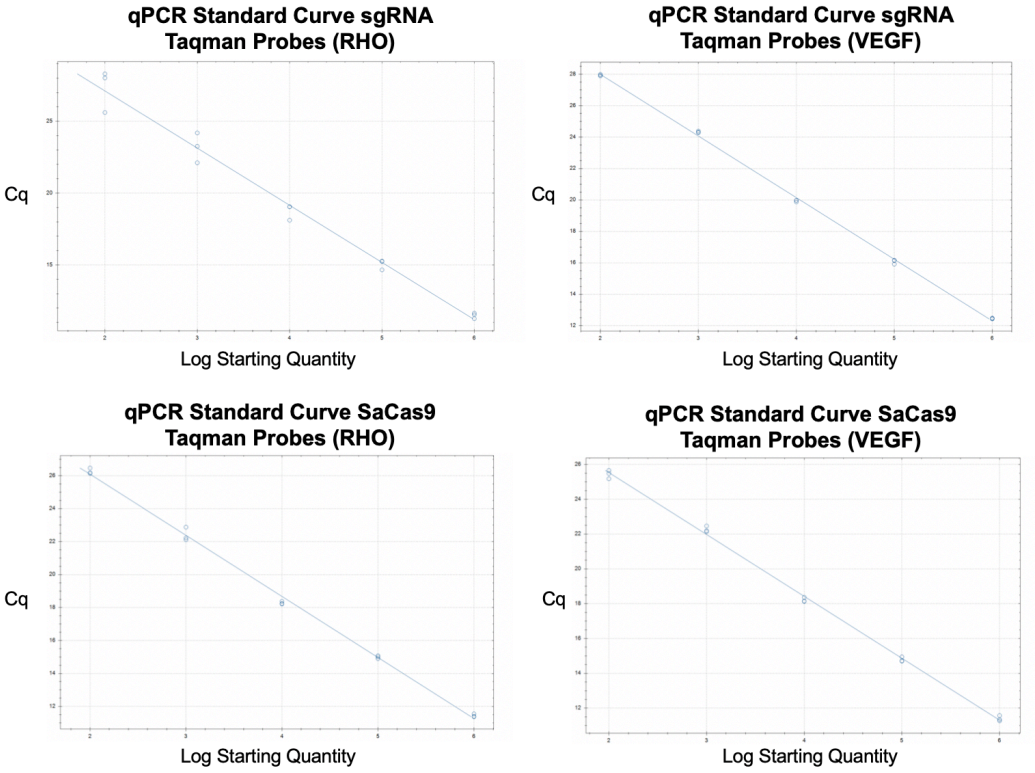

Figure 3.6. Standard Curves for SgRNA and SaCas9 Taqman Probes in HEK293-*eGFP* Cells Transfected with RHO-R and VEGFA-R Plasmids

Top left: Standard Curve for sgRNA probes in cells transfected with RHO-R plasmids. Efficiency=78.4%. R^2=0.983.

Top right: Standard Curve for sgRNA probes in cells transfected with VEGFA-R plasmids. Efficiency=79.0%. R^2=0.999.

Bottom left: Standard Curve for SaCas9 probes in cells transfected with RHO-R RHO plasmids. Efficiency=86.1%. R^2=0.997.

Bottom right: Standard Curve for SaCas9 probes in cells transfected with VEGFA-R plasmids. Efficiency=91.1%. R^2=0.998.

U6.sgRNA Orientation Does Not Affect Cas9 and SgRNA Levels in HEK293-eGFP Cells Transfected with SaCas9 Plasmids

The gene expression levels of SaCas9 and sgRNA were determined for the RHO-F, RHO-R, VEGFA-F, and VEGFA-R plasmids. The sgRNA identity did not significantly affect Cas9 expression levels ($F_{Identity}$ (1,18) = 0.03987, p = 0.8440). Sidak-adjusted comparisons revealed no significant effect of U6.sgRNA orientation on Cas9 expression ($F_{Orientation}$ (1,18) = 6.816, p = 0.0177; Sidak adjusted p_{RHO}=0.1338; Sidak-adjusted p_{VEGFA}=0.1781). There was also no significant interaction effect between sgRNA identity and orientation on Cas9 expression ($F_{Interaction}$ (1,18) = 3.510e^{-005}, p = 0.99531) (Figure 3.7).

The sgRNA identity and orientation did not significantly affect sgRNA expression levels ($F_{Identity}$ (1,18) = 1.220, p = 0.2839; $F_{Orientation}$ (1,18)=0.8046, p=0.3816). There was also no significant interaction between sgRNA identity and orientation on Cas9 expression effect ($F_{Interaction}$= (1,18) = 0.04430, p = 0.8357) (Figure 3.8).

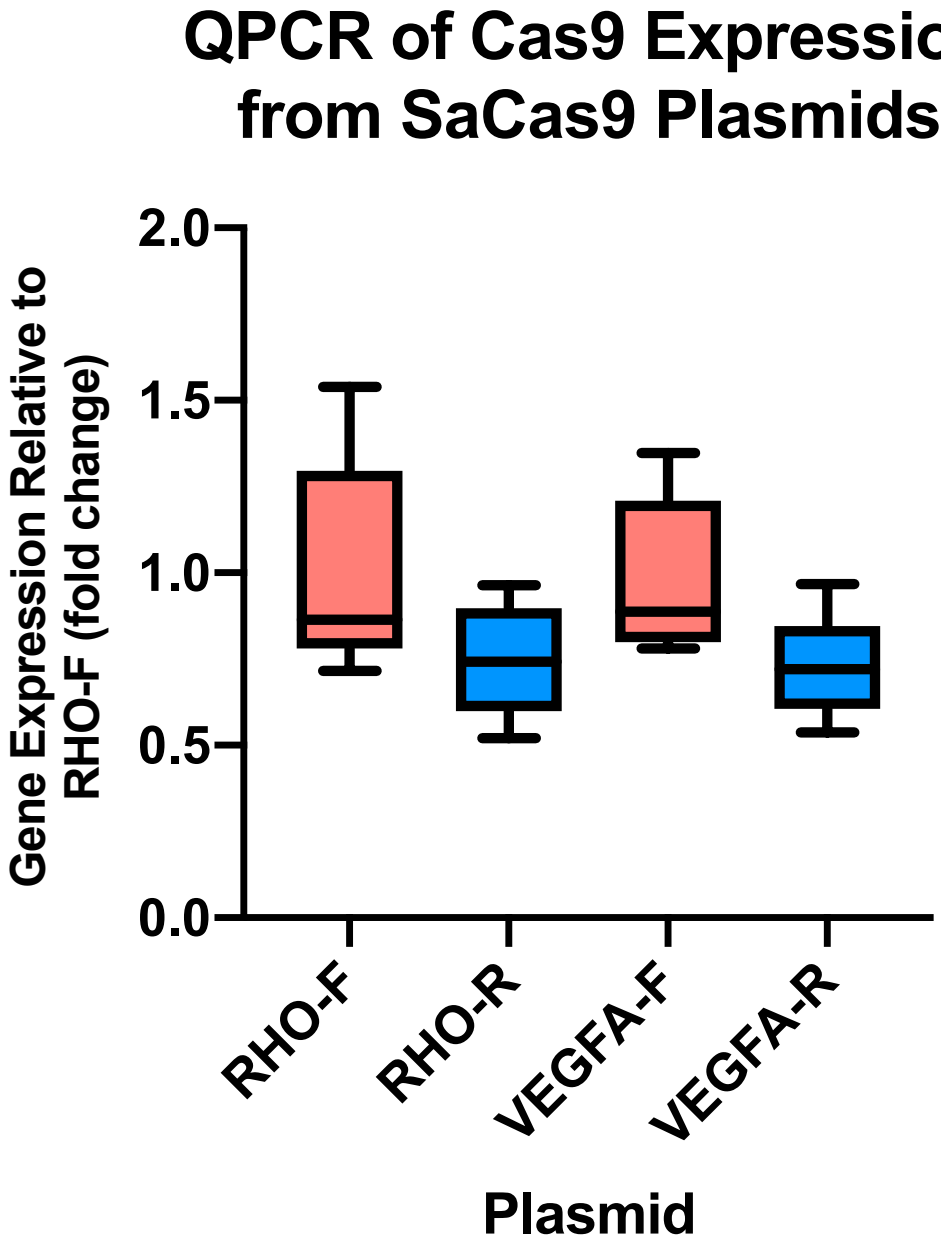

Figure 3.7. U6.sgRNA Orientation Does Not Affect Cas9 Expression in HEK293-*eGFP* Cells Transfected with SaCas9 Plasmids

The x-axis denotes the transfected plasmid. The y-axis shows SaCas9 gene expression relative to the average SaCas9 expression of cells transfected with RHO-F. The edges of each box extend from the 25th to 75th percentiles. The line within each box denotes the median. Error bars extend from the 5th to 95th percentiles. There is no significant difference in SaCas9 expression between any of the transfected samples. n=6 for RHO and n=5 for VEGFA.

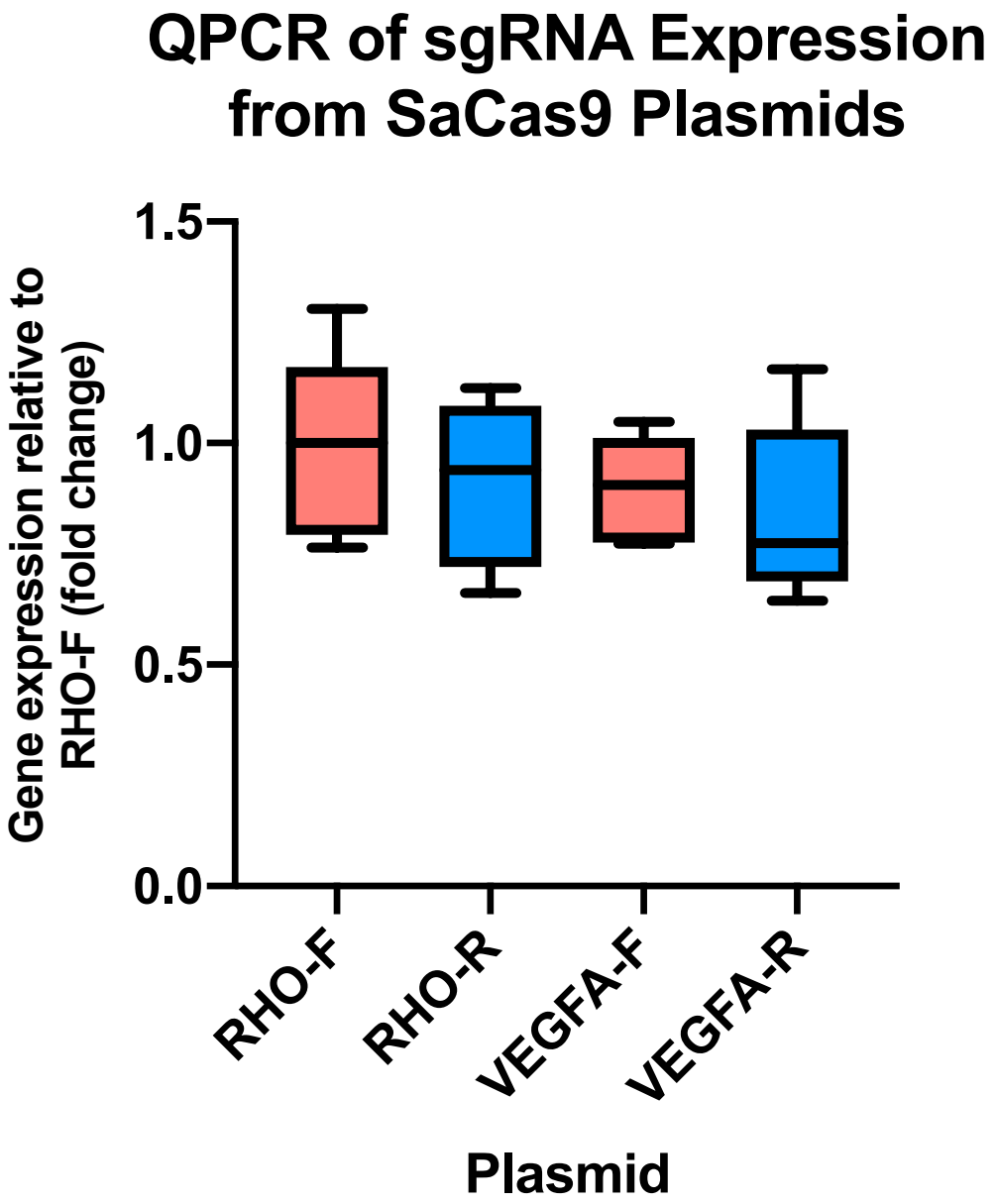

Figure 3.8. U6.sgRNA Orientation Does Not Affect sgRNA Expression in HEK293-*eGFP* Cells Transfected with SaCas9 Plasmids

The x-axis denotes the transfected plasmid. The y-axis shows sgRNA gene expression relative to the average sgRNA expression of cells transfected with RHO-F. The edges of each box extend from the 25th to 75th percentiles. The line within each box denotes the median. Error bars extend from the 5th to 95th percentiles. There is no significant difference in sgRNA expression between any of the transfected samples. n=6 for RHO and n=5 for VEGFA.

U6.sgRNA Orientation Has No Effect on Cas9, sgRNA, and EGFP Levels in HEK293-eGFP Cells Transfected with EGFP-targeting DSaCas9-KRAB Plasmids

The gene expression levels of dSaCas9, sgRNA, and *eGFP* were determined for the non-targeting (NTC-F and NTC-R) and *eGFP*-targeting plasmids (13-F, 13-R, 15-F, and 15-R).

dSaCas9 expression was not significantly affected by sgRNA identity or U6.sgRNA orientation ($F_{Identity}$ (2,36) = 1.073, p =0.3528; $F_{Orientation}$ (1,36) = 3.713, p = 0.0619). There was also no significant interaction effect of sgRNA identity and U6.sgRNA orientation on dSaCas9 expression ($F_{Interaction}$ (2,36) = 0.01433, p =0.9858).

We found no significant effect of sgRNA identity or orientation on sgRNA expression levels, and no significant interaction effect was observed ($F_{Identity}$ (2,30)=3.068, p=0.0613; $F_{Orientation}$ (1,30)=0.1086 , p=0.7440; $F_{Interaction}$ (2,30)=0.5832, p=0.5643).

Furthermore, we did not observe any effect of sgRNA identity or orientation on *eGFP* expression ($F_{Identity}$ (2,36)=0.3197, p=0.7284; $F_{Orientation}$ (1,36)=0.5038, p=0.4824). There was also no interaction effect of sgRNA identity and orientation on *eGFP* levels $F_{Interaction}$ (2,36)=0.05462, p=0.9469) (Figures 3.9, 3.10, and 3.11).

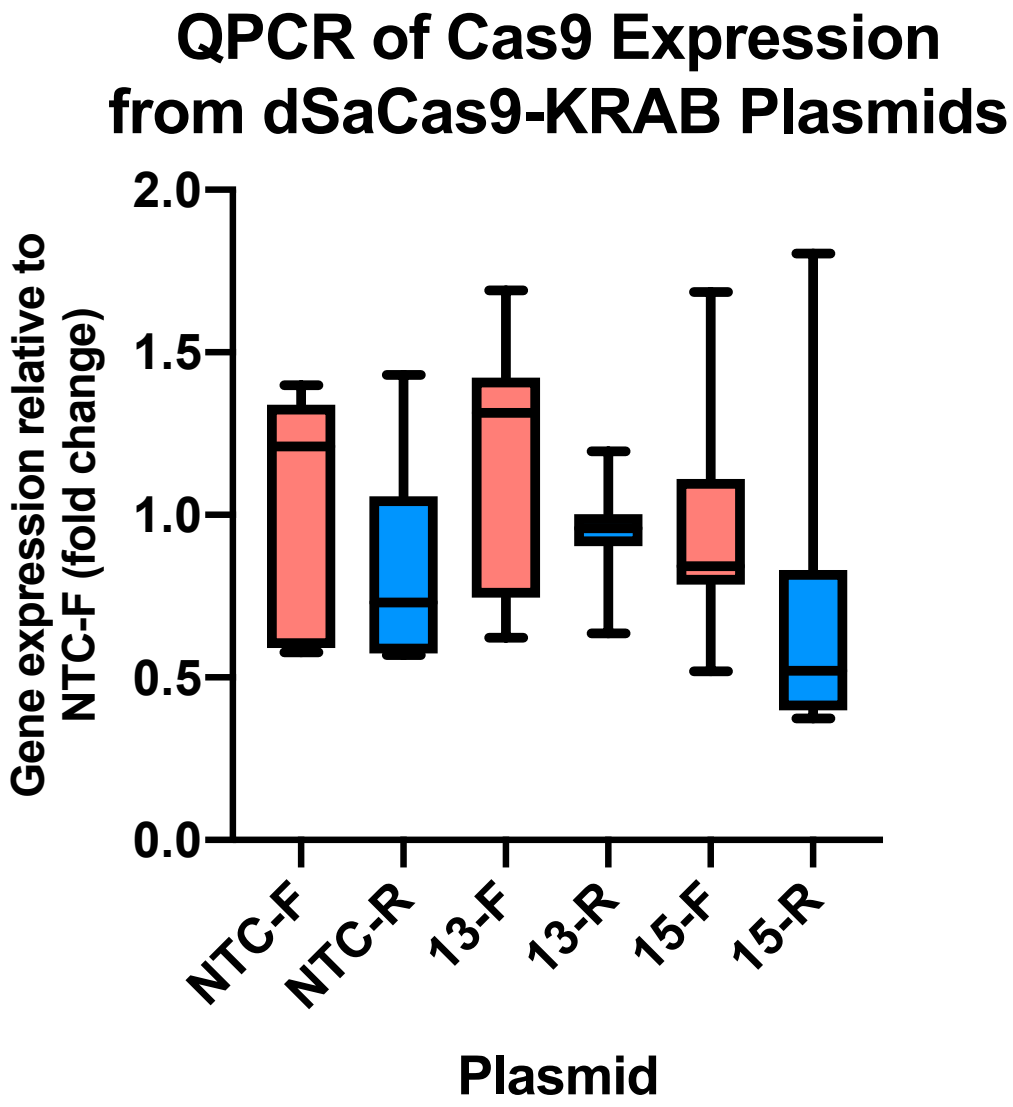

Figure 3.9. U6.sgRNA Orientation Does Not Affect Cas9 Expression in HEK293-*eGFP* Cells Transfected with DSaCas9-KRAB Plasmids

The x-axis denotes the transfected plasmid. The y-axis shows dSaCas9 gene expression relative to the average dSaCas9 expression of cells transfected with NTC-F. The edges of each box extend from the 25th to 75th percentiles. The line within each box denotes the median. Error bars extend from the 5th to 95th percentiles. There is no significant difference in dSaCas9 expression between any of the transfected samples. n=7 for all samples.

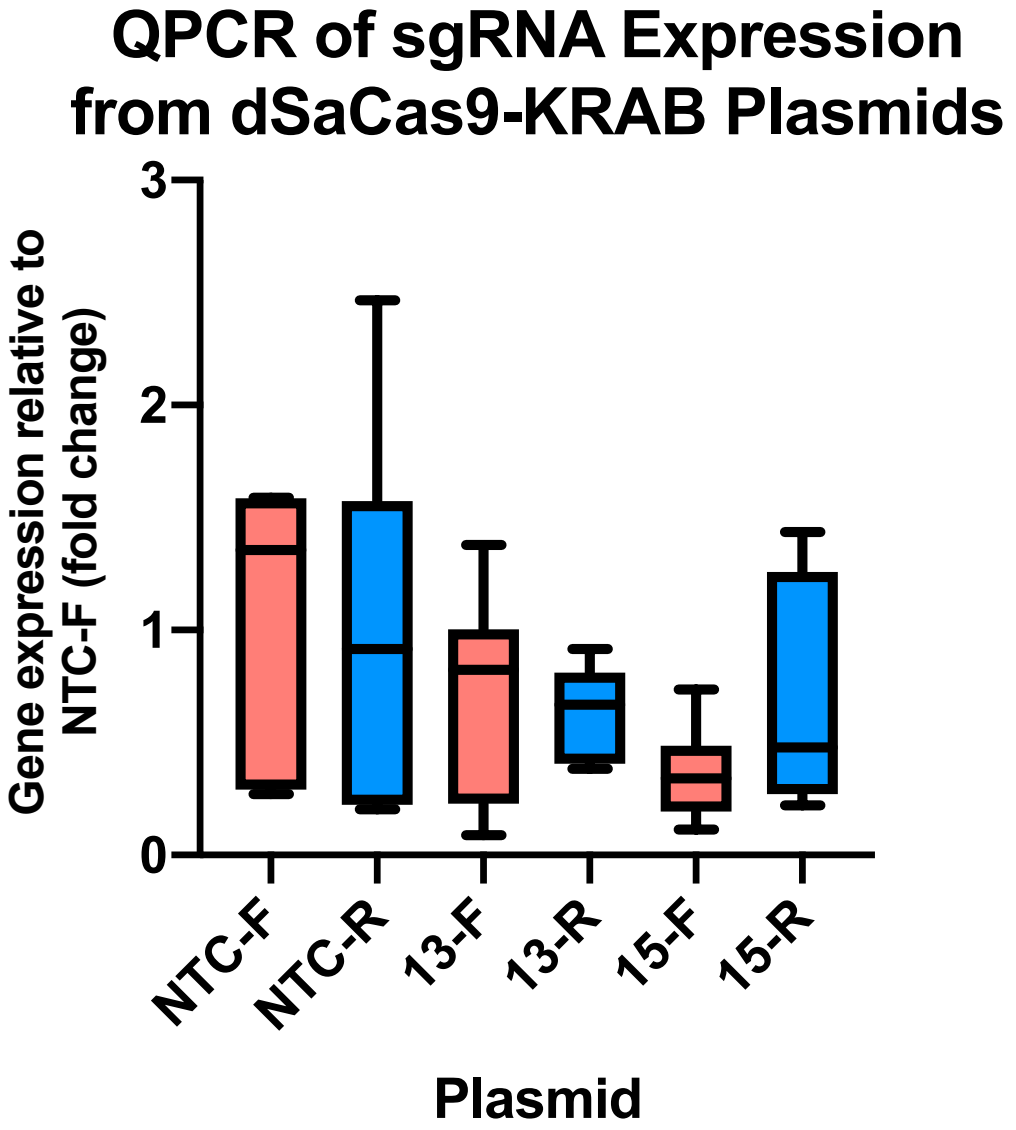

Figure 3.10. U6.sgRNA Orientation Does Not Affect SgRNA Expression in HEK293-e*GFP* Cells Transfected with DSaCas9-KRAB Plasmids

The x-axis denotes the transfected plasmid. The y-axis shows sgRNA gene expression relative to the average sgRNA expression of cells transfected with NTC-F. The edges of each box extend from the 25th to 75th percentiles. The line within each box denotes the median. Error bars extend from the 5th to 95th percentiles. There is no significant difference in sgRNA expression between any of the transfected samples. n=6 for all samples.

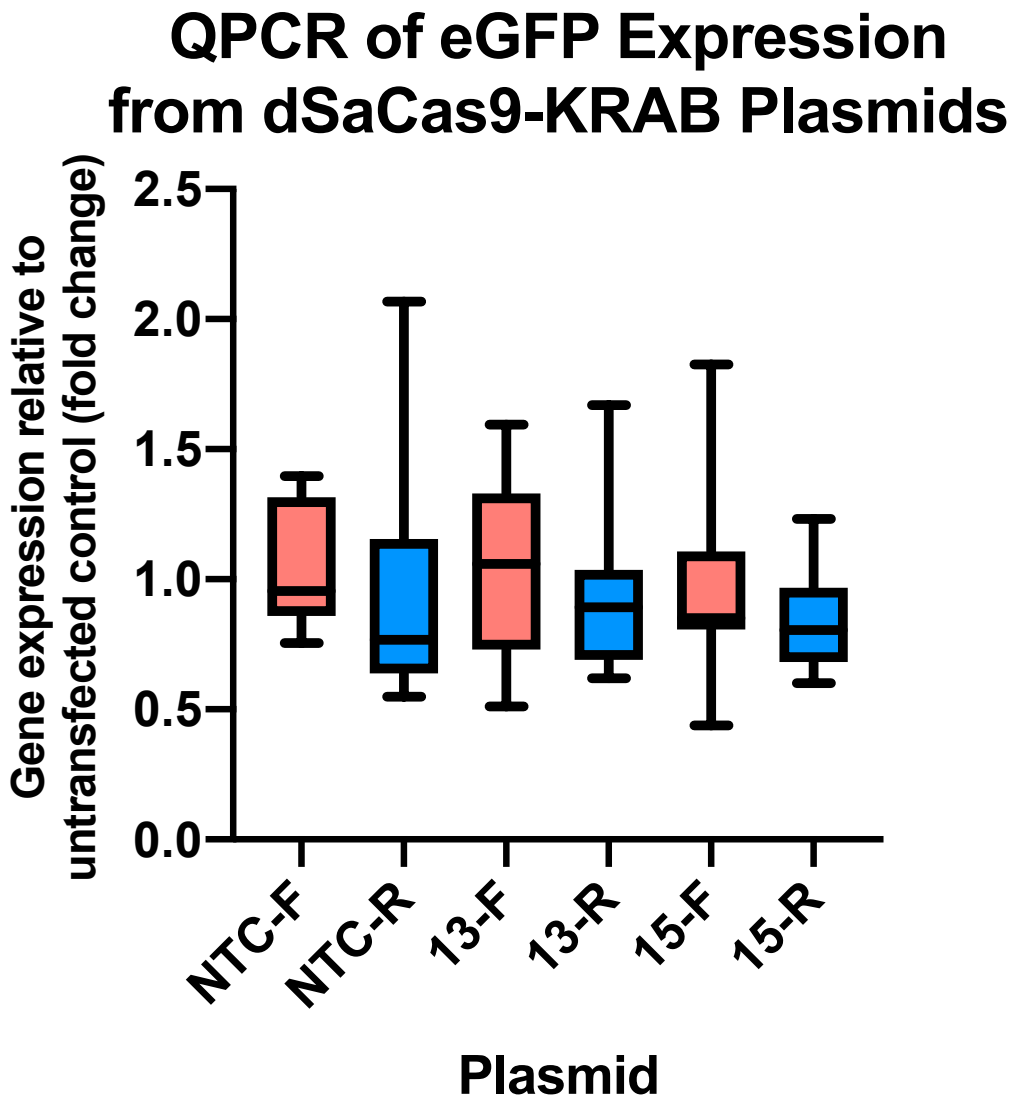

Figure 3.11. U6.sgRNA Orientation Does Not Affect *EGFP* Knock-Down in HEK293-*eGFP* Cells Transfected with DSaCas9-KRAB Plasmids

The x-axis denotes the transfected plasmid. The y-axis shows *eGFP* gene expression relative to the average *eGFP* expression of the untransfected control. The edges of each box extend from the 25th to 75th percentiles. The line within each box denotes the median. Error bars extend from the 5th to 95th percentiles. There is no significant difference in *eGFP* expression between any of the transfected samples. n=7 for all samples.

3.5 Discussion

The data presented here show that U6.sgRNA promoter orientation has no significant effect on Cas9 and sgRNA transcription from both active *Staphylococcus aureus* Cas9 and deactivated *Staphylococcus aureus* Cas9.KRAB CRISPR constructs. Furthermore, U6.sgRNA orientation does not affect *eGFP* transcription from deactivated *Staphylococcus aureus* Cas9-KRAB CRISPR constructs.

QPCR Normalization

Reference gene expression levels can be variable, and use of one reference gene may introduce bias.[168,169] It is possible that by including only one reference gene for the dSaCas9-KRAB qPCR analysis, any differences in gene expression levels between the two U6.sgRNA promoter orientations were masked. However, this outcome is unlikely because the reference gene levels were stable across samples, indicating that the different conditions tested did not affect reference gene expression; therefore, it is appropriate to use the single reference gene for analysis. Furthermore, it is important to note that even with the use of one reference gene, large differences in expression levels would be apparent. We are investigating factors that make a clinically-relevant difference in CRISPR/Cas9 efficiency; we do not believe that any small changes in gene expression level would translate to a clinically meaningful effect *in vivo*.

EGFP Expression

The lack of *eGFP* knock-down by *eGFP*-targeting constructs compared to the non-targeting controls (NTC-F and NTC-R) in our dSaCas9-KRAB plasmids is

not surprising given previous findings in our laboratory. It is likely that the HEK293-*eGFP* reporter cell line used has many copies of the *eGFP* gene, so very high amounts of gene disruption would be required to see any difference on qPCR. Our lack of significant *eGFP* reduction is partially why we chose to focus on the rhodopsin and *VEGFA* genes for the active SaCas9 study. These two genes are also clinically relevant and a validated guide RNA sequence for *VEGFA* was available.[170]

Factors Influencing Transcription

3.5.1.1 Charge of RNA Polymerases

One factor that may affect transcription is the charge of the RNA Polymerases. RNA polymerases are highly negatively charged and may repel each other if they are too close.[171,172] It is possible that in our constructs, the RNA Polymerase III and RNA Polymerase II are located too far apart to significantly repel each other and cause transcriptional interference.[171,172] The main difference between the SaCas9 and dSaCas9-KRAB constructs is that the dSaCas9-KRAB constructs contain an additional 312 bp KRAB sequence. Despite being 312 bp shorter than the dSaCas9-KRAB plasmids, the SaCas9 constructs did not experience electrostatic RNA polymerase repulsion.

3.5.1.2 Three-Dimensional Structure of Plasmids

The specific sequence of the plasmid is known to affect it's three-dimensional conformation, which may in turn influence transcription.

DNA normally exists in the B-conformation, which is the double helix first proposed by Watson and Crick. Plasmid DNA, however, may adopt several non-B DNA conformations, which include left-handed Z-DNA, intramolecular triplex H-

DNA, cruciforms, and R-loops.[173] These conformations are in part influenced by the DNA sequence and can affect the process of transcription.[166,173,174]

Furthermore, transcription by RNA polymerases influences the topology of the surrounding DNA by causing torque.[161–163] Thus, it is theoretically possible that the interplay between U6.sgRNA orientation and plasmid sequence, specifically presence and absence of a 312 bp KRAB element, may influence transcription by affecting the three-dimensional structure of the plasmid. This does not appear to be the case in our study, as we did not observe any difference in Cas9 and sgRNA levels between two different U6.sgRNA orientations in plasmids that do (dSaCas9-KRAB) and do not (SaCas9) contain KRAB.

3.5.1.3 Termination of Transcription

One mechanism that may contribute to transcriptional interference is ineffective termination of transcription from one promoter, which allows for read-through of the RNA Polymerase. This read-through causes transcriptional interference by collision or occlusion at the second promoter. Either RNA Polymerase II or III can experience weak termination of transcription.

RNA Polymerase II, which initiates transcription from the CMV promoter, relies on certain motifs to promote polyadenylation and transcriptional termination. Some of these motifs, like the SV40 late polyA and the rbGlob polyA, are more efficient because they contain enhancer elements while others, like the BGH contained in our plasmids, may be weaker.[175,176] It appears that in our constructs, the BGH signal is sufficient because we did not observe any reduction in transcription from the U6 promoter, which would be expected for *in cis* constructs if read-through were to occur from the CMV promoter.

Weak termination of transcription can also occur from RNA Polymerase III, which normally transcribes the U6.sgRNA and usually relies on poly-T residues on the non-template strand for termination. Human RNA Polymerase III requires at least four sequential thymidine residues, and termination efficiency increases with longer poly-T residues.[177,178] The strength of the poly-T sequence is not only dependent on its length but also on the surrounding sequences and on the ability of the transcribed RNA to from a hairpin secondary structure.[179–181] In our constructs, there is a five nucleotide thymidine sequence immediately following the sgRNA to facilitate termination. If this sequence is not strong enough to terminate RNA Polymerase III, read-through will occur. This effect would possibly be more pronounced in the SaCas9 constructs than in the dSaCas9-KRAB constructs because the 312 bp KRAB sequence contains an additional five nucleotide poly-T sequence, which may further act as a transcriptional terminator. We did not observe any evidence of read-through from the U6 promoter in either the SaCas9 or the dSaCas9-KRAB plasmids, as we did not see differences in transcription of Cas9 levels between forward and reverse U6.sgRNA constructs.

Antisense RNA Interference

Another consideration, one that applies specifically to constructs oriented *in trans*, is the effect of antisense RNA interference. RNA Polymerase III read-through can cause antisense RNA transcripts of Cas9 to interfere with Cas9 transcripts produced by RNA Polymerase II, and thus promote degradation of Cas9 RNA.[182,183] We did not see any evidence of antisense RNA interference in our study in either the SaCas9 or the dSaCas9-KRAB constructs.

Conclusion

Based on our results, it appears that the orientation of the U6.sgRNA does not cause transcriptional interference in the constructs was we studied. Our observations are consistent with previous reports on promoter configuration, in which the amount of transcriptional interference varies depending on the genes being studied.

4 Investigation of GeoCas9 as an Alternative CRISPR/Cas9 System for an AAV Vector

4.1 Introduction

PAM Site Variation

One limitation of SaCas9 is the PAM sequence (5'-NNGRRT-3' or 5'-NNGRR-3'), which restricts CRISPR/Cas9 editing to certain regions of the DNA.[27] Exploring the potential of other Cas9 proteins with different PAM sequences would be beneficial in creating an arsenal of CRISPR/Cas9 systems that can target a wide range of DNA sequences. The GeoCas9, with a PAM sequence is that 5'-NNNNCRAA-3', is one such alternative, where N represents any base pair and R represents adenine or guanine. GeoCas9 is derived from *Geobacillus stearothermophilus*, a bacterium that inhabits ocean sediments and hot springs. GeoCas9 is 3.2 kb in size, so it is small enough to fit within the ~4.7 kb coding capacity of AAV for delivery into the retina (Figure 4.1).[184] Furthermore, GeoCas9 RNP has increased stability in blood plasma, a feature that may be beneficial if CRISPR/Cas9 systems were ever to be delivered into the eye in this form.[184]

Immunology of CRISPR/Cas9

Another potential benefit of GeoCas9 is that it may reduce the risk of an immune response. Currently, the potential immunological risks of CRISPR/Cas9 therapy are largely unknown. There have been conflicting reports of CRISPR/Cas9's ability to activate the immune system. Several experiments in mice have shown that CRISPR/Cas9 can induce an immune response. It is not certain, however, exactly which component of the therapy is responsible for the observed effect; possibilities include the delivery vector and the Cas9 protein.

There is evidence that CRISPR/Cas9 therapies delivered in adenoviral vectors, lentiviral vectors, or through electroporation can activate the immune

system.[185–188] These immunologic changes are not observed when SpCas9 is delivered to mice as a naked DNA construct, implicating the vectors and not the CRISPR/Cas9 as the cause of the immune response.[189] CRISPR/Cas9 systems delivered in adeno-associated vectors (AAV), which are much less immunogenic, appear to be well-tolerated. In the eye, particularly, AAV-SpCas9 did not induce any immune response for up to four weeks after delivery, although the SpCas9 had self-limiting expression.[135] AAV-SaCas9 and AAV-CjCas9 have not caused any adverse immunological events in the eye for up to nine months after injection.[139,144] Other *in vivo* studies have found no evidence of morphological and physiological abnormalities in cells expressing Cas9 from AAV.[144,190]

However, even with an AAV delivery system, there is still some concern that the Cas9 proteins may induce an immune response. Cas9-specific T cells and anti-Cas9 antibodies have been identified in mice injected with AAV-SpCas9-sgRNA constructs. [188] These mice also developed lymph node enlargement and an increase in myeloid cells.[188] Despite evidence of immune system activation, the AAV-Cas9-sgRNA constructs did not cause extensive tissue damage and the proliferated T cells appeared to be immature and inactive.[188]

Another reason for caution with CRISPR/Cas9 therapies is that humans may have a pre-existing immune response to the commonly-used Cas9 proteins, SpCas9 and SaCas9. SpCas9 and SaCas9 are derived from *Streptococcus pyogenes* and *Staphylococcus aureus*, two bacteria that have frequently caused disease during human evolution. Varying levels of pre-existing immunity to Cas9 proteins have recently been identified. Simhadri et al. 2018 observed that out of 200 individual human serum samples, 10% had anti-SpCas9 antibodies and 2.5% had anti-SaCas9 antibodies.[191] Charlesworth et al. 2019 found a prevalence of

78% and 58% for SpCas9 and SaCas9 antibodies, respectively, in a cohort of 125 patients.[192] Furthermore, Charlesworth et al. 2019 tested 18 donors for T cells and observed that 78% and 68% of subjects had specific reactive T cells against SpCas9 and SaCas9, respectively. Wagner et al. 2018 examined the peripheral blood of 48 subjects and found that 46 of them had specific effector T cell activation by SpCas9 components.[193] Some of the effector T cells were of the T regulatory subtype, which may act to temper the immune response to SpCas9.[193] Despite the evidence of both reactive and regulatory T cells in the peripheral blood of these patients, the immunological impact of this T cell repertoire *in vivo* is unknown.[193]

Using a Cas9 protein derived from bacteria that do not cause disease in humans is a potential strategy to reduce the risk of an immune response. Unlike *Streptococcus pyogenes* and *Staphylococcus aureus*, *Geobacillus stearothermophilus*, from which GeoCas9 is derived, is not known to cause disease in humans. The risk of pre-existing antibodies and T cells to GeoCas9 is in theory quite low. Furthermore, a deactivated GeoCas9 protein could possibly be immunologically safer than dSpCas9 and dSaCas9. Deactivated Cas9 proteins must be continuously expressed for a sustained therapeutic effect. This prolonged Cas9 expression, however, increases the risk of an immune response; therefore, using a less immunogenic Cas9 protein would be advantageous. In this study, we investigate the editing potential of GeoCas9, which may be a suitable alternative to SpCas9 and SaCas9 for gene therapy.

Figure 4.1 Image of *Geobacillus Stearothermophilus* Bacteria

Geobacillus Stearothermophilus is a gram positive spore-forming rod-shaped bacterium that inhabits hot springs and ocean sediments and can grow at temperatures between 35°C and 75°C.[194]

Figure adopted from Novik et al 2018.

4.2 Aims

The aim of this chapter is to assess whether the 3.2 kb GeoCas9, whose PAM site differs from that of SaCas9, can cleave DNA when delivered to HEK293-*eGFP* cells through plasmid transfection. Furthermore, the potential immunogenicity of GeoCas9 will be assessed using *in silico* modelling.

4.3 Materials and Methods

A dual plasmid approach was used to deliver the GeoCas9 components to HEK293-*eGFP* cells. One plasmid contained a CMV promoter driving GeoCas9 expression while the other contained a U6 promoter driving sgRNA expression. A summary of the plasmids used in this study can be found in Table 3 and plasmid maps are in Figures 4.2 and 4.3.

Table 3. List of Plasmids Used in GeoCas9 Study

Plasmid Code	Cas9	Ampicillin Resistance	6x Histidine Tag Present	sgRNA Target
Geo	GeoCas9	Yes	-	-
Geo-6x His	GeoCas9	Yes	Yes	-
NTC	-	Yes	-	Non-targeting control
G1	-	Yes	-	*eGFP*
G2	-	Yes	-	*eGFP*
G3	-	Yes	-	*eGFP*
G4	-	Yes	-	*eGFP*
RHO	-	Yes	-	Rhodopsin
VEGFA	-	Yes	-	*VEGFA*
SaNTC	SaCas9	Yes	-	Non-targeting control (NTC)
SaVEGFA	SaCas9	Yes	-	*VEGFA*
ABCA4-6xHis			Yes	

Cloning of GeoCas9 Plasmids

The PX404 *Campylobacter jejuni* Cas9 plasmid (#68338, Addgene) was used as a vector for insertion of the GeoCas9 coding sequence from the pET-MBP-NLS-Geo_st plasmid (#87703, Addgene).

First, the unique restriction sites KpnI and HindIII were inserted into plasmid #68338 to allow for subsequent sticky-end cloning of the GeoCas9. Two

complementary oligonucleotides coding for four consecutive restriction sites (AgeI-KpnI-HindIII-EcorI FW and RV) were annealed and phosphorylated (Table 6 Appendix). The PX404 *Campylobacter jejuni* Cas9 plasmid was then digested with AgeI and EcorI restriction enzymes. The digested vector was loaded onto a 1% agarose gel for electrophoresis and subsequent gel extraction of the appropriately-sized band. The oligonucleotide sequences were PCR purified and diluted 1/25 in molecular grade water. 1 µL of this dilution was used for a ligation reaction with 10 ng of the linearized PX404 *Campylobacter jejuni* Cas9 plasmid vector.

Subsequently, the GeoCas9 coding region and nuclear localization signal from plasmid #87703 were PCR amplified using primers GeoCas9 KpnI FW6 and GeoCas9 HindIII RV5 to add KpnI and HindIII restriction sites to the 5' and 3' ends, respectively (Table 5 and Figure 7.8, Appendix). The products were PCR purified. Both the PCR products and the vector plasmid were digested with the KpnI and HindIII restriction enzymes and PCR purified. Ligations of the two sequences were performed using a 2:1 ratio of insert:vector and 10 ng vector in a 10 µL reaction (Figure 7.9, Appendix).

Cloning of GeoCas9-6xHis Plasmids

Two complementary oligonucleotides coding for a 6x Histidine tag (6xHis) with KpnI overhangs were annealed and phosphorylated (Table 5, Appendix). The GeoCas plasmid was subjected to digestion with the KpnI restriction enzyme. The 6xHis-coding oligonucleotide was diluted 1/25 in molecular grade water. 1 µL of this dilution was used for a ligation reaction with 10 ng of the linearized GeoCas9 plasmid vector.

Cloning of GeoCas9 SgRNA Plasmids

The pU6-Cj-sgRNA plasmid (#89753, Addgene) was used as a vector for insertion of the "98 nt tracrRNA sgRNA" (GeoCas9 sgRNA) from Harrington et al. 2017.[184] A double-stranded oligonucleotide coding for the GeoCas9 sgRNA with SalII and NdeI restriction sites on the 5' and 3' ends, respectively, was ordered (GeneArt™, Thermo Fisher; GeoCas9 sgRNA scaffold, Table 5, Appendix). Both the oligonucleotide and the pU6-Cj-sgRNA plasmids were digested with SalI and NdeI. The digested vector was loaded onto a 1% agarose gel for electrophoresis and subsequent gel extraction of the appropriately-sized band. Ligations were performed using a 3:1 ratio of insert:vector and 30 ng vector in a 10 µL reaction (Figure 7.10, Appendix). The resulting plasmid was termed "NTC."

To insert the crRNA sequences, complementary oligonucleotides coding for the appropriate crRNA with BsmbI restriction site overhangs were phosphorylated, annealed, and diluted 1/25 in molecular grade water. The "NTC" plasmid was then digested with the BsmbI restriction enzyme. The digested vector was loaded onto a 1% agarose gel for electrophoresis and subsequent gel extraction of the appropriately-sized band. Ligations were performed with 1 µL of the diluted oligonucleotides and 10 ng of the digested vector (Table 4).

Transfection of Plasmids

HEK293-*eGFP* cells were seeded at a density of $5x10^5$ cells/mL onto 24 well plates for qPCR analysis or 12-well plates for western blot (Corning). Cells were transfected with either 500 ng of DNA for qPCR or 750 ng for western blot (1:3 ratio of GeoCas9:sgRNA) 24 hours after seeding, as described in chapter 2. Cells were harvested 48 hours after transfection, as described in chapter 2.

RNA Extraction, CDNA Synthesis, and QPCR

Harvested cell pellets were used for RNA extraction and subsequent cDNA synthesis using primer GeosgRNA2 for qPCR, as described in chapter 2 (Table 7, Appendix). Three biological replicates were used for each sample. For qPCR, a single reference gene, beta-actin (ACTB), was used for normalization. One-way analysis of variance (ANOVA) was performed. The Taqman probe sequences targeting GeoCas9 and GeoCas9 sgRNA are listed in Table 9 of the Appendix.

Tracking of Indels by DEcomposition (TIDE) Analysis

TIDE is a bioinformatics tool that quantifies CRISPR/Cas9 editing efficiency by aligning an “uncut” reference Sanger sequence with a Sanger sequence from edited DNA.

PCR was performed on genomic DNA samples isolated from harvested cells using the QIAamp DNA Minikit (Qiagen, UK). Primers pairs were designed to amplify a ~700 base pair region, with one primer ~200 bp away from the predicted CRISPR cut site and the other ~500 bp away from the predicted CRISPR cut site, as per the TIDE analysis protocol (Figure 7.11, Appendix). All primers used for TIDE analysis are listed in Table 7 of the Appendix. Amplified DNA was PCR purified using a QIAquick PCR purification kit, per the manufacturer’s protocol (Qiagen, UK) and sent for Sanger sequencing (Eurofins Genomics, EU or Source Bioscience, UK) using the primer 200 bp away from the cut site. Sequences were uploaded to the TIDE analysis website (https://tide.deskgen.com/) and editing efficiency was determined (example in Figure 7.12, Appendix). The SaCas9 was selected from the drop-down menu for the TIDE analysis settings because it was the only nuclease that allowed input of a 22 base pair guide RNA sequence.

Three biological replicates were used for TIDE analysis, except for the samples transfected with Geo+G3, which had four replicates. A one-way ANOVA with multiple comparisons was conducted. When necessary, the Benjamini-Krieger-Yekutieli procedure was used to correct for the false discovery rate.

Western Blot

Western blot was performed to detect the presence of 6xHis tagged GeoCas9 proteins in harvested cells. Frozen cell pellets were thawed on ice. 100 µL of lysis buffer [1 mL 10x RIPA (Merck Millipore), 8 mL PBS, protease inhibitor tablet (Roche), and 1 mL glycerol] was added to each pellet and cells were lysed with a hand-held homogenizer. Samples were then incubated on ice for 30 minutes and centrifuged at 10,000xg at 4ºC for 20 minutes. The supernatant was removed and mixed with 5 µL of 5x protein loading buffer (GeneFlow, UK) in a total volume of 25 µL. Samples were then incubated at room temperature for 20 minutes and subsequently loaded into wells of a 4-20% Criterion™ Tris-Glycine extended (TGX™) Precast Gel (Bio-Rad, UK) alongside 5 µL of BLUeye pre-stained protein ladder (Geneflow, UK). The gel was placed in a tank containing running buffer (100 mL Tris-glycine SDS [GeneFlow, UK] and 900 mL RO water) for sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), which separates the denatured proteins by size. A 100 V electric current was applied to the gel for a period of 1 hour and 30 minutes. For subsequent immunofluorescence staining of the proteins, the gel was removed from the tank and the proteins were transferred to a polyvinylidene difluoride (PVDF) membrane using the Trans-Blot Turbo system (Bio,Rad, UK), per the manufacturer's protocol. The PVDF membrane was incubated in blocking buffer (Odyssey, TBS, LI-COR,

NE, USA) for one hour with gentle rotation. The membrane was then incubated for one hour under gentle rotation with 20 mL primary antibody solution containing 1% blocking buffer TBS-T [0.1% Tween 20 (Merck) in 1 L 1x TBS], 1:2000 mouse anti-GAPDH (TA802519, Origene), and 1:1000 rabbit anti-6x His tag (ab9108, Abcam). The membrane was then washed three times with TBS-T before being incubated for 30 minutes under gentle rotation with 20 mL secondary antibody solution containing 1:10,000 donkey anti-rabbit 800CW (926-32212, LICOR) and 1:10,000 donkey anti-mouse 680RD (926-68072, LICOR). After addition of the fluorescent secondary antibodies, the membrane was kept in the dark to prevent photobleaching. The membrane was subsequently washed three times, dried, and imaged on the LICOR Odyssey Imager.

In Silico Prediction of Binding to MHCI and MHCII

The protein sequences for SaCas9 and GeoCas9 were entered into the IEDB 3.0 server (www.iedb.org). Default settings were used to predict the ability of the two Cas9 proteins to bind MHCI and MHCII.

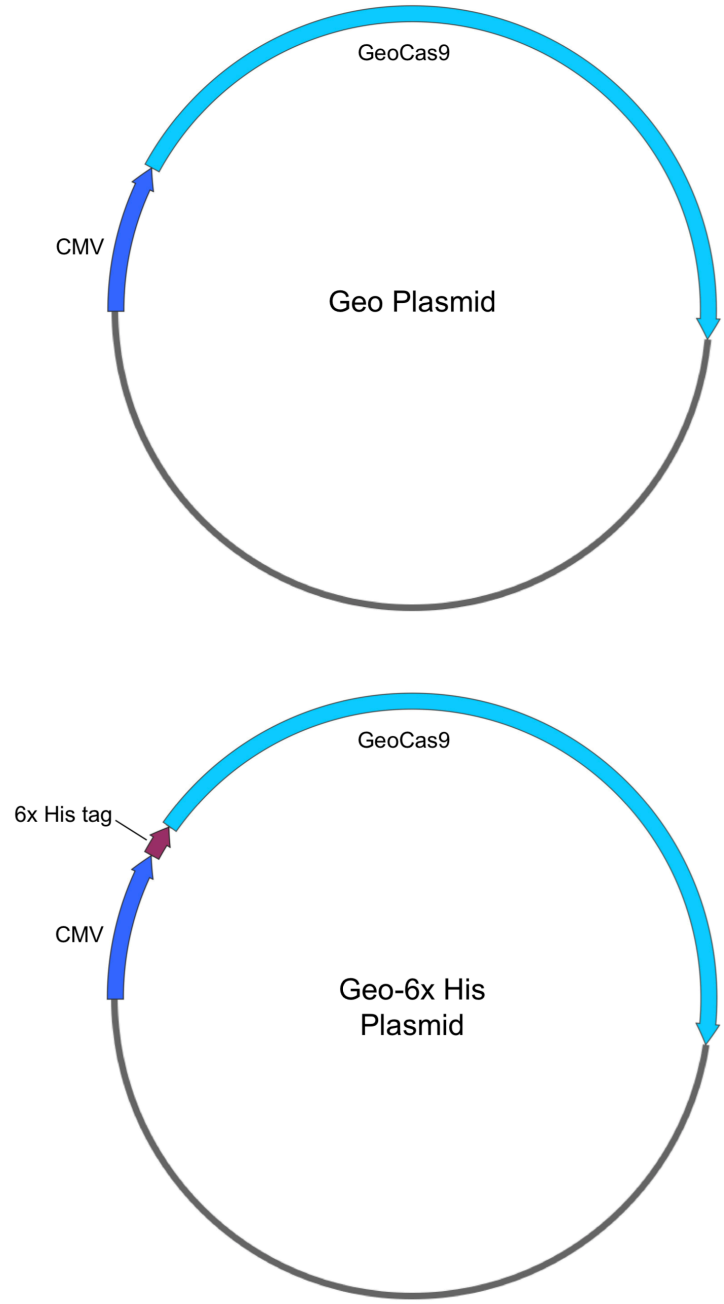

Figure 4.2 Plasmid Maps for Geo and Geo-6x His

The plasmid map on the top depicts the structure of the Geo plasmid and the map on the bottom depicts the structure of the Geo-6x His plasmid.

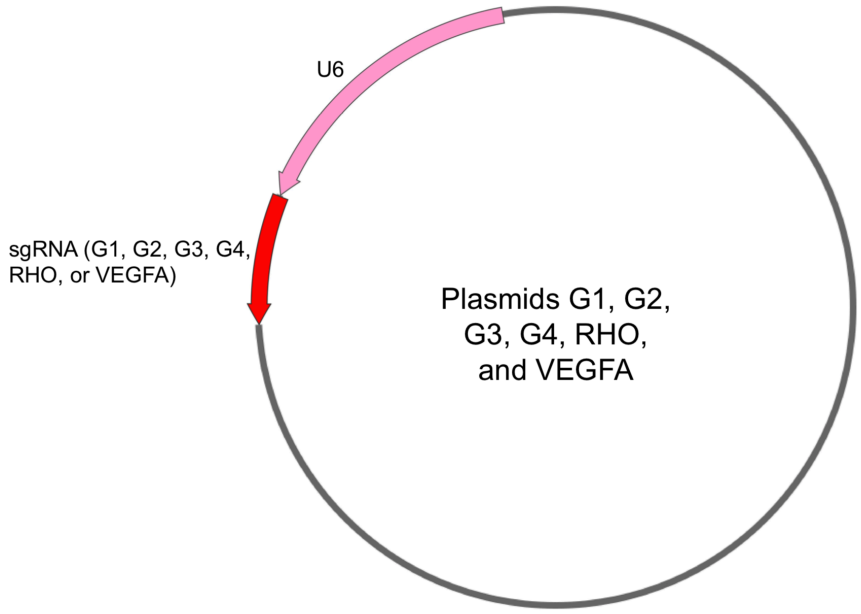

Figure 4.3 Plasmid Map for G1, G2, G3, G4, RHO, and VEGFA

The plasmid map depicts the structure of the G1, G2, G3, G4, RHO, and VEGFA plasmids.

4.4 Results

Standard Curve for SgRNA and GeoCas9 Taqman Probes in Cells Transfected with GeoCas9 and SgRNA G3 Plasmids

Taqman qPCR probe efficiency was tested with the standard curve method using cDNA derived from HEK293-*eGFP* cells co-transfected with the G3 and Geo plasmids. The efficiency of the sgRNA Taqman probe was 73.3% and the R^2 was 0.998 while the efficiency of the GeoCas9 probe was 82.2% and the R^2 value was 0.999 (Figure 4.4).

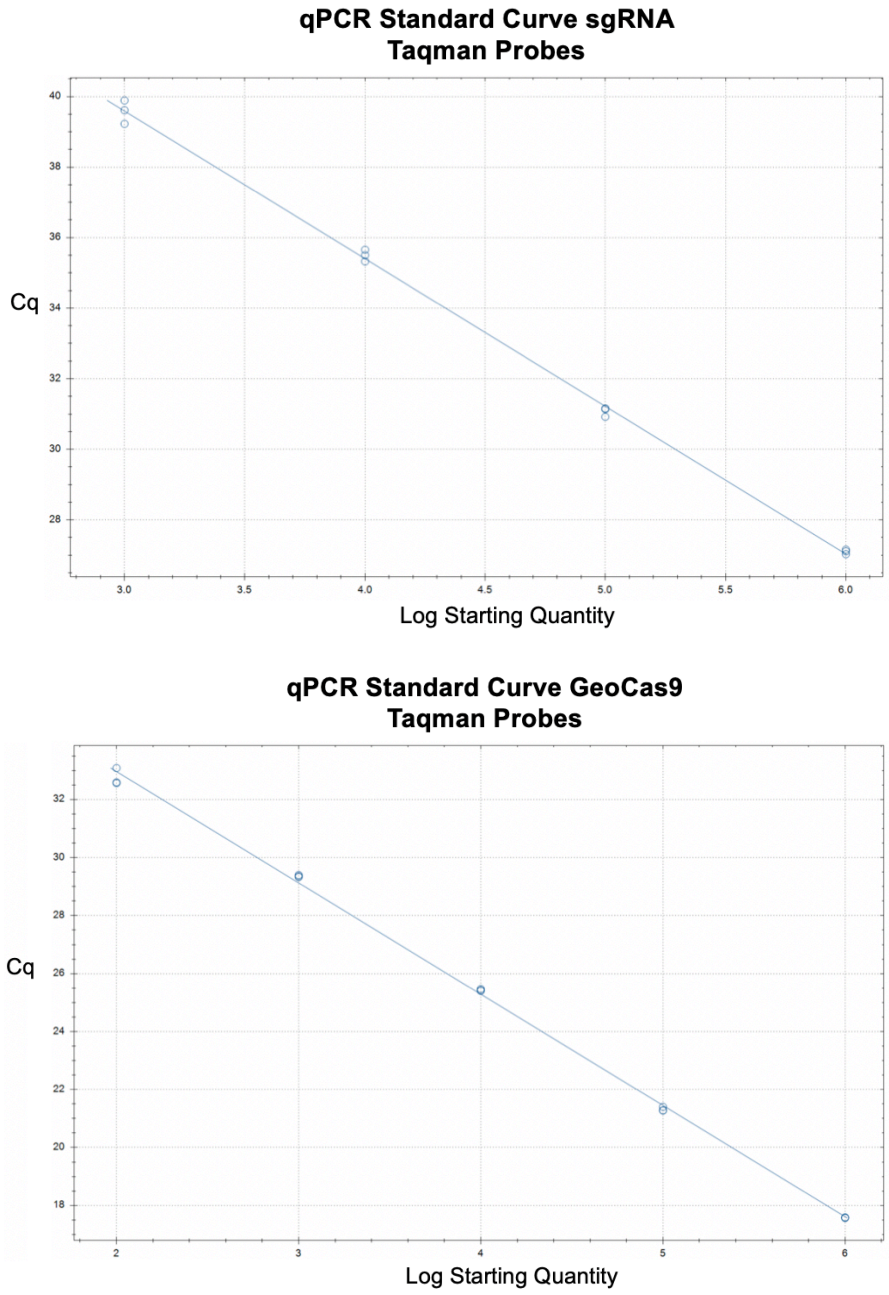

Figure 4.4. Standard Curve for SgRNA and GeoCas9 Taqman Probes in Cells Transfected with GeoCas9 and SgRNA G3 Plasmids

Top: Standard Curve for sgRNA probes in cells transfected with GeoCas9 and sgRNA G3 plasmids. Efficiency=73.3%. R^2=0.998.

Bottom: Standard Curve for GeoCas9 probes in cells transfected with GeoCas9 and sgRNA G3 plasmids. Efficiency=82.2%. R^2=0.999.

TIDE Analysis Shows No DNA Cleavage by GeoCas9 with EGFP-Targeting SgRNAs

Cleavage rates were determined using TIDE analysis. The *eGFP* cleavage rates from the dual-transfected wells (Geo plasmid plus G1, G2, G3, or G4 plasmids) were compared to those from wells transfected with only *eGFP*-targeting sgRNA (G1, G2, G3, or G4), Geo, Geo with non-targeting sgRNA plasmid (NTC), untransfected wells, and non-targeting sgRNA (NTC) (Figure 4.5).

There was no significant disruption of the *eGFP* locus when the *eGFP*-targeting sgRNA plasmids (G1, G2, G3, and G4) were dual-transfected with the GeoCas9 plasmid [F_{Geo+G1} (5,12) = 6.679, p = 0.0034; F_{Geo+G2}(5,12) = 0.6746, p = 0.6507; F_{Geo+G3}(5,17) = 3.082, p = 0.0368; F_{Geo+G4}(5,12) = 2.819, p = 0.0656]. For the Geo+G1 plasmid transfections, no significant difference in cutting efficiency was found after using the Benjamini-Krieger-Yekutieli procedure to correct for the false discovery rate. For the Geo+G3 plasmid transfections, multiple comparisons revealed no significant difference in cutting efficiency between any of the plasmid transfections ($p>0.05$).

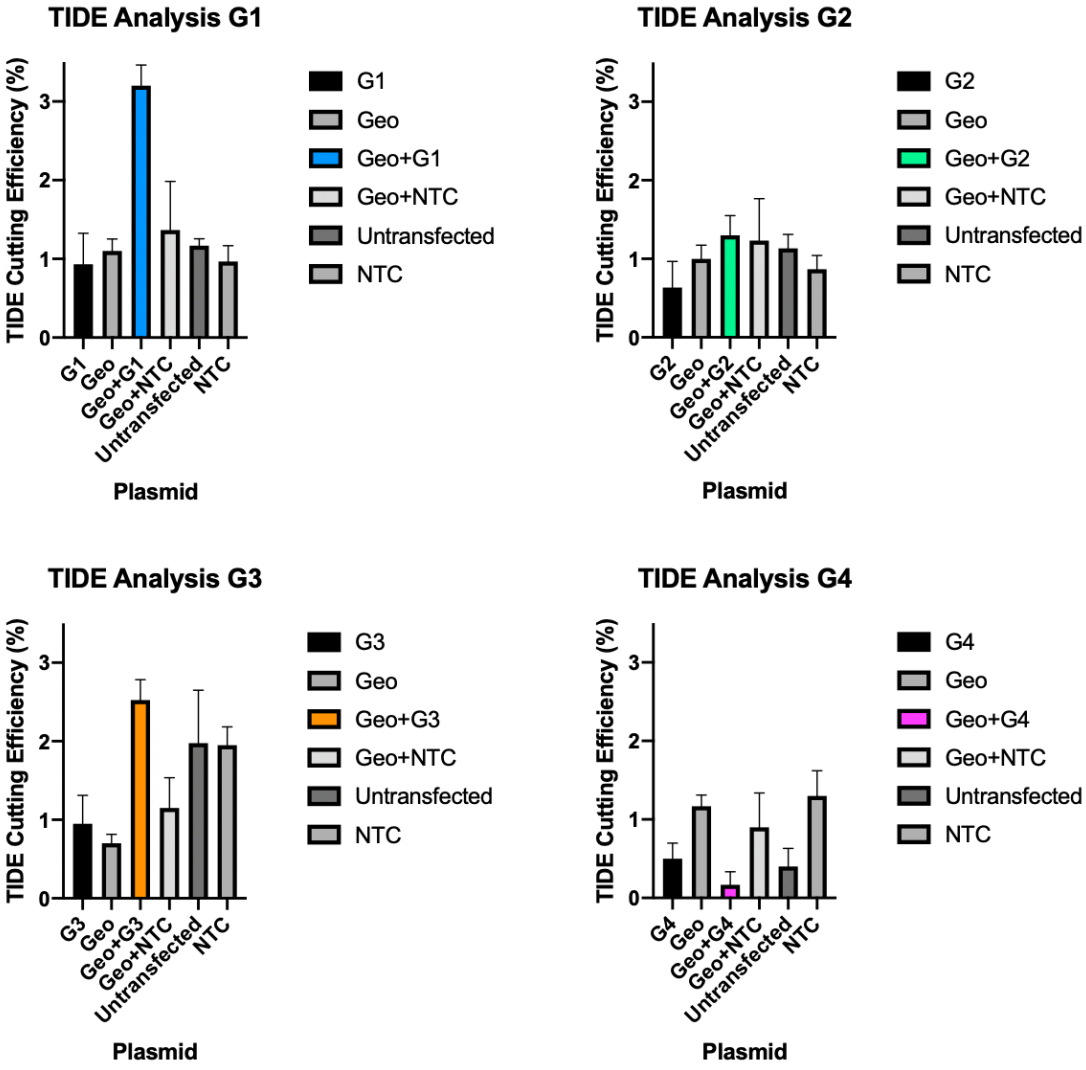

Figure 4.5. TIDE Analysis Shows No DNA Cleavage by GeoCas9 with *EGFP*-Targeting SgRNA

The x-axis shows the identity of the transfected plasmid. The y-axis depicts the cutting efficiency as predicted by TIDE analysis. The error bars represent the standard error of the mean. Cells were transfected with a 1:3 ratio of GeoCas9:sgRNA and harvested 48 hours after transfection. sgRNA G3 was previously validated by Harrington et al. 2017.[184] Control samples are in greyscale and results from cells transfected with both the GeoCas9 and *eGFP*-targeting sgRNA plasmids are shown in colour. There is no significant disruption of the *eGFP* locus by any of the tested sgRNAs. n=3 for all samples except Geo+G3, for which n=4.

TIDE Analysis Shows No DNA Cleavage by GeoCas9 with VEGFA- and RHO-Targeting SgRNA

Cleavage rates were determined using TIDE analysis. The *VEGFA* cleavage rates from the dual-transfected wells (Geo plasmid with VEGFA sgRNA plasmid) were compared to those from wells transfected with *VEGFA*-targeting SaCas9 plasmids (SaVEGF), non-targeting SaCas9 plasmids (SaNTC), VEGFA, Geo, Geo and non-targeting sgRNA plasmids (NTC), untransfected wells, and non-targeting sgRNA (NTC). The *VEGFA* targeting efficiency of SaCas9 (62.7±4.7%) was significantly higher than that of GeoCas9 (0.33±0.24%) ($p<0.0001$). There was no significant difference in *VEGFA* targeting between Geo+VEGFA transfected wells and those transfected with control plasmids ($p>0.05$). $F_{Geo+VEGFA}(7,16)=168.8$, $p<0.0001$.

For *RHO*-targeting, the cleavage rates from the dual-transfected wells (Geo plasmid with RHO sgRNA plasmid) were compared to those from wells transfected with RHO, Geo, Geo and non-targeting sgRNA plasmids (NTC), untransfected wells, and non-targeting sgRNA (NTC). There was no significant difference in rhodopsin disruption between any of the experimental samples, $F_{Geo+RHO}(5,12) = 0.7098$, $p = 0.6276$ (Figure 4.6).

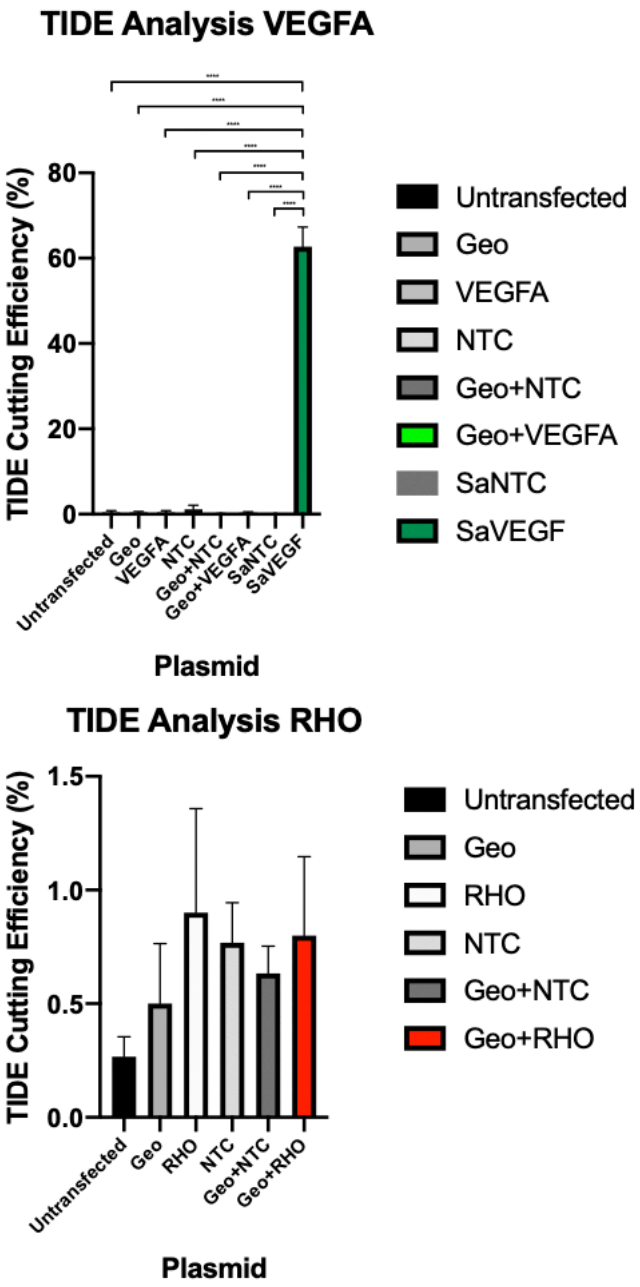

Figure 4.6. TIDE Analysis Shows No DNA Cleavage by GeoCas9 with *VEGFA*- and *RHO*-Targeting SgRNA

The x-axis shows the identity of the transfected plasmid. The y-axis depicts the cutting efficiency as predicted by TIDE analysis. The error bars represent the standard error of the mean. Cells were transfected with a 1:3 ratio of GeoCas9:sgRNA or SaCas9 and harvested 48 hours after transfection. Control samples are in greyscale and results from cells transfected with Cas9 and *RHO/VEGFA*-targeting sgRNA plasmids are shown in colour. There is no significant disruption of the *RHO* and *VEGFA* loci by GeoCas9. The SaCas9 with the *VEGFA*-targeting sgRNA serves as a positive control and has a targeting efficiency of 62.7±4.7%. n=3 for all samples.

QPCR Shows SgRNA Expression in Plasmid-Transfected HEK293-eGFP Cells

In order to confirm that the lack of cutting activity found with GeoCas9 was not due to failure of expression of the relevant molecular components, sgRNA (guide variant G3) expression after transfection with plasmids was determined using qPCR (Figure 4.7). There was a significant difference in sgRNA expression between the untransfected cells and those transfected with either G3 or Geo and G3 plasmids ($p<0.0001$). There was also a significant difference in sgRNA expression between cells transfected with only the Geo plasmid and those transfected with G3 or Geo and G3 plasmids ($p<0.0001$). $F_{sgRNA}(3,8)=127.9$, $p<0.0001$. Hence sgRNA deficiency was not the cause of the reduced cutting efficacy of GeoCas9.

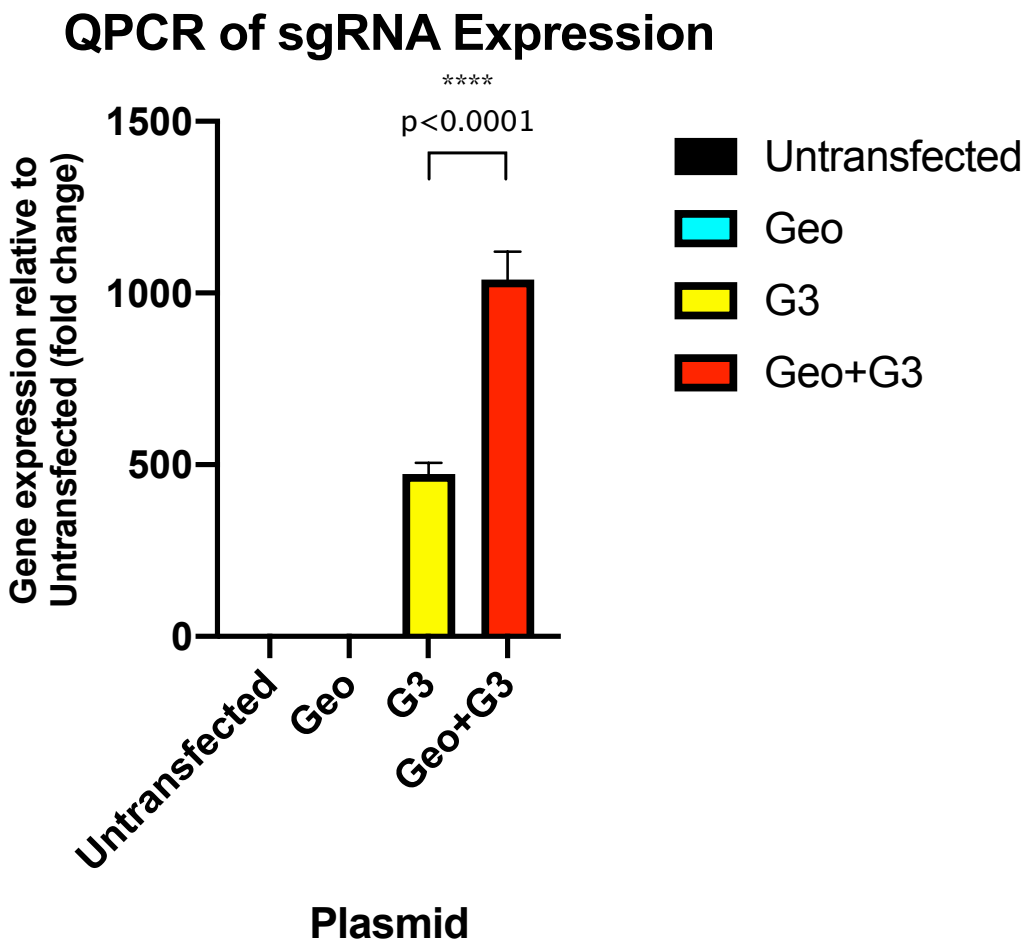

Figure 4.7 QPCR Shows SgRNA Expression in Plasmid-Transfected HEK293-*eGFP* Cells

The x-axis shows the identity of the transfected plasmid. The y-axis depicts the sgRNA gene expression relative to the untransfected control. The error bars represent the standard error of the mean. Cells were transfected with a 1:3 ratio of GeoCas9:sgRNA and harvested 48 hours after transfection. There is expression of sgRNA in the cells transfected with sgRNA alone and GeoCas9 with sgRNA. n=3 for all samples.

QPCR and Western Blot Shows GeoCas9 RNA and Protein Expression in Plasmid-Transfected HEK293-eGFP Cells

In order to determine whether the lack of detectable DNA cleavage was due to an absence of GeoCas9, GeoCas9 expression in transfected cells was determined using qPCR and Western Blot (Figures 4.8 and 4.9). For the qPCR analysis, there was a significant difference in GeoCas9 expression between the untransfected cells and those transfected with either Geo or Geo and G3 plasmids ($p<0.0001$). There was also a significant difference in GeoCas9 expression between cells transfected with only the G3 plasmid and those transfected with Geo or Geo and G3 plasmids ($p<0.0001$). $F_{GeoCas9}(3,8)=788.6$, $p<0.0001$.

Similarly, the level of GeoCas9 protein expression was determined by western blot. An antibody against the 6xHis tag was used to visualize the GeoCas9-6xHis protein. A ~128 kDa protein containing the 6xHis tag was present in cells transfected with the Geo6xHis plasmids and in cells dual transfected with the Geo6xHis and G3 plasmids. No fluorescence signal was detected from untransfected controls or from cells transfected with untagged GeoCas9 plasmids (Geo). Hence, both the GeoCas9 and the relevant sgRNA (G3) were confirmed to be expressed efficiency in the same HEK293-eGFP cell line in which SaCas9 showed high cutting activity.

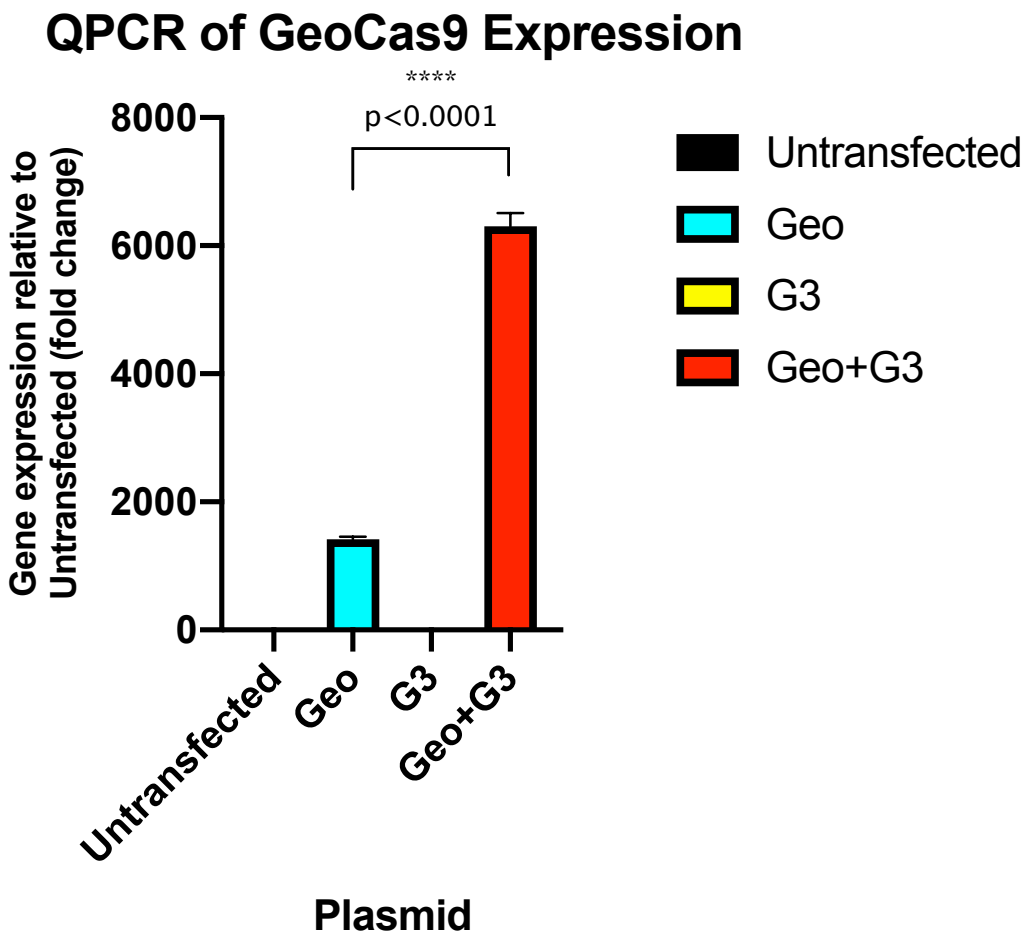

Figure 4.8. QPCR Shows GeoCas9 Expression in Plasmid-Transfected HEK293-*eGFP* Cells

The x-axis shows the identity of the transfected plasmid. The y-axis depicts the GeoCas9 gene expression relative to the untransfected control. The error bars represent the standard error of the mean. Cells were transfected with a 1:3 ratio of GeoCas9:sgRNA and harvested 48 hours after transfection. There is expression of GeoCas9 in the cells transfected with GeoCas9 alone and GeoCas9 with sgRNA. Hence the lack of cutting efficiency with GeoCas9 cannot be explained by lack of Cas9 mRNA or sgRNA expression. n=3 for all samples.

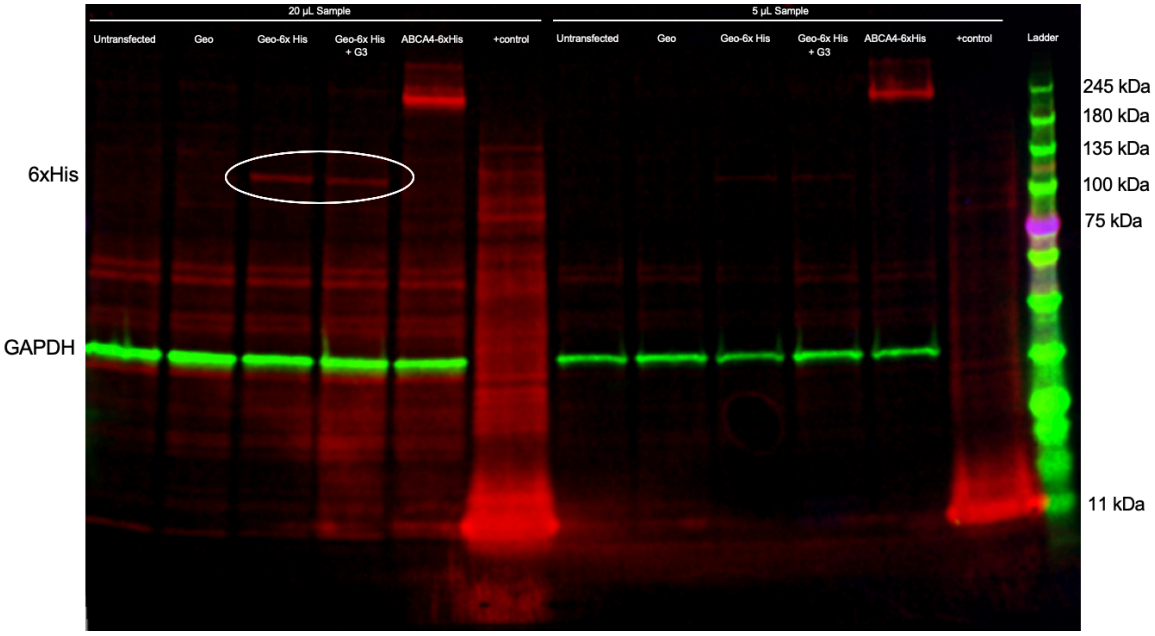

Figure 4.9. Western Blot Shows GeoCas9 Protein Expression in Plasmid-Transfected HEK293-*eGFP* Cells

In order to confirm that the GeoCas9 protein was being translated from the mRNA transcripts, cells were transfected and harvested 48 hours after transfection and a western blot was performed on the protein lysates. Antibodies targeting a 6xHis tag were used for the immunofluorescence staining (red). The GeoCas9-6xHis protein is 128 kDa in size. There is expression of GeoCas9-6xHis in the wells transfected with GeoCas9-6xHis and GeoCas9-6xHis + sgRNA G3 (denoted by the white oval). His tagged recombinant ABCA4 protein was used as a positive control for the His antibody on the Western blot.

In Silico Prediction Reveals that GeoCas9 and SaCas9 Have Similar Ability to Bind to MHCI and MHCII

The first stage in determining whether or not GeoCas9 might have immunological advantages over SaCas9 in terms of avoiding the innate immune response in human cells is to compare the three-dimensional protein structures in terms of their major histocompatibility complex (MHC) Class I and II binding ability. This was assessed in various small peptide regions from the SaCas9 and GeoCas9 proteins using the IEDB 3.0 server (www.iedb.org). SaCas9 and GeoCas9 yielded similar patterns of predicted MHCI and MHCII binding for the various small peptide sequences (Figure 4.10).

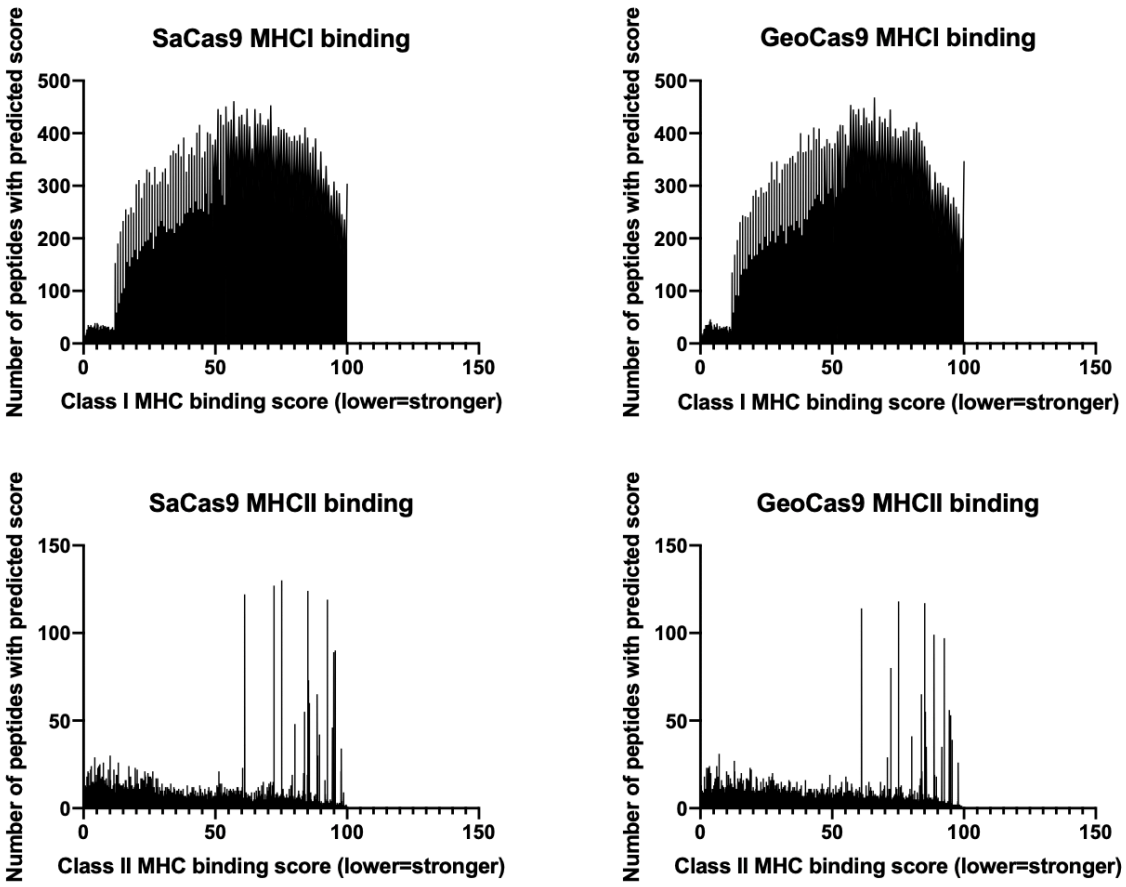

Figure 4.10. *In Silico* Prediction Reveals that GeoCas9 and SaCas9 Have Similar Ability to Bind to MHCI and MHCII

The x-axis in each graph indicates the predicted "binding score" of small peptide fragments from the Cas9 proteins. A lower score denotes a higher binding ability. The y-axis represents the number of small peptide fragments with the predicted binding score. MHC=Major Histocompatibility Complex.

4.5 Discussion

The potential advantages of using GeoCas9 are that it is small enough to fit into an AAV vector and, compared to SaCas9, it has a different PAM site. It also has increased stability in blood plasma compared to SpCas9 and lower potential to induce an immune response. We have demonstrated that plasmids encoding a GeoCas9 CRISPR/Cas9 system cannot edit DNA to detectable levels in HEK293-*eGFP* cells. We confirmed expression of both sgRNA and GeoCas9 RNA using qPCR after single and dual plasmid transfection; the higher levels of RNA after dual plasmid delivery are likely due to the synergistic effects of co-transfection.[195,196] We also confirmed GeoCas9 protein expression using western blot. The lack of DNA editing may possibly be attributed to inaccessibility of the target DNA, defects in the sgRNA component, and inability of the GeoCas9 protein to cleave DNA.

An additional *in silico* comparison of the two Cas9 protein structures predicted comparable binding of both SaCas9 and GeoCas9 in relation to Class I and II MHC receptors. This suggests that the immunological responses to either of the two proteins may be similar.

Accessibility of Target DNA

Heterochromatin, which consists of DNA that is tightly bound to nucleosomes, may be difficult to access by CRISPR editing machinery.[197,198] We know from successful SaCas9 editing of *VEGFA* that the *VEGFA* gene in the HEK293-*eGFP* cells is accessible to CRISPR/Cas9 systems. We therefore chose to compare SaCas9 and GeoCas9 disruption of the *VEGFA* gene as a positive control. Because we were examining the editing ability of two different Cas9

proteins with different PAM requirements, we could not use the exact same sgRNA. We did, however, choose sgRNAs that bind to adjacent regions of the target DNA. The SaCas9 constructs produced high levels of *VEGFA* disruption (69.3±0.2%) while GeoCas9 did not induce *VEGFA* cleavage. It therefore seems unlikely that our observation is due to inaccessibility of the target DNA region.

SgRNA

The binding of the CRISPR/Cas9 system to the target DNA could be affected by sgRNA sequence. Specific crRNA sequences have preferences for different PAM sites.[199] It is possible that the crRNA sequences tested may have a stronger affinity for a PAM site other than 5'-NNNNCGAA-3', which was used in our study.

Furthermore, the presence of specific motifs near the crRNA sequence may influence editing efficiency. For our sgRNA constructs, we chose to insert two extra guanine base pairs at the 5' end, before the crRNA, in order to replicate the sgRNA constructs from Harrington et al. 2017.[184] These additional base pairs may improve transcriptional efficiency and reduce off-target effects.[200,201] The two extra guanines can, however, reduce on-target efficiency.[200] Cho et al. 2014 compared four different sgRNA-targeting sites, using either G(crRNA) or GG(crRNA) created through *in vitro* transcription.[200] Two of the targets had comparable editing rates by the two sgRNA variants.[200] The other variants had significantly lower on-target effects by GG(crRNA) than G(crRNA) (18% versus 47% for one locus and 6% versus 41% for the other).[200] Despite any possible decrease in editing efficiency with an additional guanine, it is still clear that the CRISPR/Cas9 system was able to effectively cleave DNA. Furthermore, Cho et al. 2014 relied on the T7

endonuclease 1 (T7E1) assay to quantify cutting efficiency.[200] T7E1 is known to be much less sensitive than the TIDE analysis method that we used.[202] It is therefore likely that the Cho et al. 2014 study underestimated the editing efficiency of the CRISPR/Cas9.

Cho et al. 2014 and Harrington et al. 2017 used *in vitro* transcription to create their sgRNAs, meaning that they relied on the T7 promoter while we used plasmid transfection and relied on the U6 promoter to transcribe our sgRNAs. It could be that the T7 and U6 promoters create a different final sgRNA product, which may affect sgRNA association with the GeoCas9 protein and/or the target DNA region.[203] This has not however been noted in a multitude of other CRISPR studies in which U6-driven sgRNA from plasmids has been shown to be highly efficacious.

GeoCas9

GeoCas9 delivered as an RNP can efficiently cleave target DNA, including *eGFP* in HEK293-*eGFP* cells.[184] We show that when delivered as a transfected plasmid, the GeoCas9 does not cleave DNA to detectable levels.

One potential reason for reduced GeoCas9 cutting is that the optimal temperature for GeoCas9 activity may be higher than 37°C. Harrington et al. 2017 have shown that GeoCas9, while still functional at 37°C, has lower activity at this temperature compared to SpCas9.[184] In our experiments, all of our cells were incubated at 37°C, which may have reduced the efficiency of the GeoCas9 enzyme. Even though the activity of GeoCas9 is lower at 37°C, it should still produce some detectable DNA cleavage, as seen in Harrington et al. 2017.[184]

A more plausible reason for our observations is that there are differences in either post-translational processing or in protein folding of the plasmid-derived GeoCas9 proteins when compared to the RNP used by Harrington et al. 2017.[184] Post-translational modifications can affect protein folding, function, and stability.[204] For plasmid delivery, the eukaryotic HEK293-*eGFP* cells must first transcribe and then translate the protein from the plasmid DNA template. For RNP delivery, the GeoCas9 proteins are first created in *Escherichia coli* bacteria and then eluted. Post-translational modifications are known to differ between cell types. Mammalian, plant, and insect cells all glycosylate different amino acid side chains while *E. coli* cells may not glycosylate proteins at all.[204–206] Since the plasmid-derived and RNP GeoCas9 proteins were created in different types of cells, it is possible that their post-translational modifications may be different. These modifications could certainly affect protein folding and function, and could potentially explain the observed difference in DNA cleavage ability. A third possibility is that the sgRNA and GeoCas9 components do not associate with each other after transcription and translation, thus preventing normal function of the CRISPR/Cas9 system.

In Silico Modelling of GeoCas9 Activity

In addition to determining that GeoCas9 delivered as a plasmid could not cut DNA at detectable levels compared to SaCas9, we also used computer modelling to compare the two proteins in terms of major histocompatibility complex (MHC) recognition sites. For Class I and Class II MHC binding, both proteins had similar predicted structures. We do not know what these results indicate about GeoCas9's ability to induce an immune response *in vivo*.

Despite its potential to bind to MHC, GeoCas9 may still be immunological favourable to SaCas9 because pre-existing GeoCas9-targeting antibodies and T cells in the adaptive immune response are less likely to be found in humans. The similarities between GeoCas9 and SaCas9 might however indicate that MHC innate immunity goes back to a common vertebrate ancestor that might have existed long before human evolution. Additional modelling on different Cas9 variants might determine this further.

5 General Discussion

5.1 Introduction

CRISPR/Cas9 therapy holds great promise to treat retinal degeneration. Thus far, only one clinical trial using CRISPR/Cas9 to treat blindness has been approved in principle for Phase I, but it is currently still in the enrolment phase. Before CRISPR/Cas9 is adopted in the clinic on a large scale, optimizing this technology is essential. For example, determining the appropriate dose of the Cas9 and sgRNA components is crucial. Furthermore, since CRISPR/Cas9 is a biological intervention, being cognizant of the effects of therapy on the human immune system is crucial. This thesis focuses on optimizing CRISPR/Cas9 for potential AAV gene therapy, specifically by investigating factors that affect sgRNA and Cas9 expression and by exploring an alternative Cas9 protein that may target different DNA regions and evade detection by the immune system.

5.2 Promoter Orientation

In the first part of our study, we used qPCR analysis to examine the effects of U6.sgRNA promoter orientation on Cas9 and sgRNA expression levels from HEK293-*eGFP* cells transfected with SaCas9 and dSaCas-KRAB constructs. We also tested the effects of promoter orientation on target gene knock-down with the dSaCas-KRAB constructs. We found we saw no difference in gene expression levels for all of our comparisons.

Our findings are congruent with previous studies of promoter orientation on transcript levels, which show variable effects of orientation depending on the specific construct and conditions that are being assessed. Each of these vectors will have a unique sgRNA, one of several Cas9 proteins, and possibly additional elements like KRAB repressors. Therefore, it appears that for each specific

CRISPR/Cas9 gene therapy vector that is created, various promoter orientations must be individually tested. Other factors that could be investigated are the effects of the order of the transgenes within the construct.

We do not yet know whether this similarity in Cas9 and sgRNA levels is valid when the transgene is delivered in an AAV rather than through plasmid transfection and if the same results would be observed in retinal cells instead of HEK293-*eGFP* tissue culture cells. The circularised AAV genome is much smaller than the plasmids used in this assay and steric hindrance between RNA polymerase assemblies may become more of an issue as transcriptional start and termination sites converge. The transcriptional bubble refers to the open region of single stranded DNA that is unzipped by RNA polymerase II at the 3' end whilst simultaneously reforming the double strand at the 5' end during transcription. The transcriptional bubble for RNA polymerase II in eukaryotic cells is a fairly constant 15-18 nucleotide length.[207] Less is known about RNA polymerase III which transcribes the sgRNA from the U6 promoter, but electron microscopy studies indicate a similar transcription bubble size.[208] Hence it would be important at least to ensure that two genes *in cis or in trans* do not approximate closer than this as the two RNA polymerase enzymes are likely to collide. Identifying minimal spacing for regulatory elements is highly relevant to AAV gene therapy where space limits are critical, particularly for CRISPR due to the large Cas9 coding sequences. The transcriptional proteins of the nucleosome associated with the RNA polymerases are, however, likely to extend beyond the bubble and identifying a minimal safe separation zone is therefore essential in the design of any bicistronic AAV vector transgene.

Our findings are therefore relevant for cloning of transgenes for *in vivo* AAV delivery of CRISPR systems because we do not appear to have converged the coding sequences to the point where steric hindrance appears to be an issue. For the “EDIT-101” CRISPR/Cas9 therapy that is being applied to treat LCA10, Maeder et al. 2019 created a CRISPR/Cas9 construct with two sgRNAs driven by a U6 promoter arranged *in cis* with the SaCas9.[46] Based on our results, it appears that the configuration of sgRNA and SaCas9 in this study is efficient for gene editing and would not be significantly different if the U6.sgRNA and SaCas9 were oriented *in trans*.

5.3 GeoCas9

The second part of our study focused on testing a newly characterized Cas9 protein, GeoCas9, which has a unique PAM site, higher stability in blood plasma, and a theoretically lower risk of inducing an immune response than the commonly-used SpCas9 and SaCas9.[184] We demonstrate that GeoCas9 delivered through plasmid transfection to HEK293-*eGFP* cells does not detectably disrupt target DNA. Furthermore, *in silico* modelling predicts that GeoCas9 and SaCas9 have similarly-structured epitopes in terms of potential MHC Class I and II interactions.

We explored several reasons why the GeoCas9 may not be functioning as expected in our experimental setup. We confirmed the presence of both sgRNA and GeoCas9 RNA and protein. Possible reasons for lack of editing despite transcription and translation of the CRISPR/Cas9 machinery include defects in the sgRNA, GeoCas9 protein, or both.

Exploring alternative PAM sites and sgRNA scaffold sequences in our constructs could have been another approach. The described PAM site for

GeoCas9 is 5'-NNNNCRAA-3'.[184] We used 5'-NNNNCGAA-3', as it was found to be the most efficient by Harrington et al. 2017.[184] Furthermore, of all of the sgRNA scaffolds tested by Harrington et al. 2017, the scaffold that we used demonstrated the highest editing rates.[184] While the other tested PAM sites and sgRNA sequences were not as effective, they still did produce obvious DNA target disruption.[184] It seems unlikely that changing the PAM and sgRNA scaffold sequences could result in successful editing by our system when the already optimized PAM and sgRNA sequences produced no editing with our protocol. The inability of GeoCas9 to cleave DNA in our experiments, therefore, is likely due to differences in delivery and production of GeoCas9 and sgRNA (i.e. plasmid versus ribonucleoprotein delivery).

The *in silico* predictions about GeoCas9 raise the possibility that this Cas9 may still cause an immune response. We do not know, however, how these *in silico* findings translate *in vivo*. If the GeoCas9 cannot effectively edit target DNA, exploring its immunologic profile may be unnecessary.

5.4 Strategies for Optimizing CRISPR/Cas9

The aims of the work described in this thesis were to investigate two strategies for optimizing CRISPR/Cas9 for retinal gene therapy. The goal of optimization is to increase on-target editing efficiency while decreasing off-target effects. CRISPR/Cas9 optimization generally focuses on altering either the guide RNA or the Cas9.

Delivery Method

The delivery method can play a significant role in CRIPSR/Cas9 efficacy. SgRNAs transfected as plasmids cleave fewer off-target sites than do sgRNAs

created through *in vitro* transcription.[203] This finding is likely due to the fact that some *in vitro* transcribed sgRNAs may be truncated; plasmid-derived sgRNAs, on the other hand, are more likely to be complete.[203] CRISPR/Cas9 ribonucleoproteins are even better than plasmid-derived CRISPR/Cas9 systems, demonstrating less off-target cleavage.[209] Furthermore, ribonucleoproteins avoid the risk of plasmid DNA integration into the host genome.[209] While ribonucleoproteins can successfully edit retinal pigment epithelium cells, they cannot be delivered safely to the neuroretina.[112]

Using a Different Promoter

One approach is to change the promoters that drive sgRNA and Cas9 expression. In our constructs, we used the U6 promoter for sgRNA and the CMV promoter for Cas9 expression. RNA Polymerase III transcribes from U6 promoters and RNA Polymerase II transcribes from CMV promoters. A possible alternative is to use the H1 promoter instead of the U6 promoter for sgRNA expression. Compared to systems with the U6 promoter, CRISPR/Cas9 systems that rely on the H1 promoter result in similar on-target effects and fewer off-target effects.[210] This finding may be due to the H1 promoter being weaker than U6, thus resulting in a lower amount of sgRNA transcripts. [210] A single H1 promoter, which has both RNA Polymerase II and RNA Polymerase III activity, can also be used to express both the sgRNA and the Cas9.[211] The advantage of using one promoter to express both components of the CRISPR/Cas9 system is that is reduces the size of the vector.[211] The H1 promoter produces fewer Cas9 and sgRNA transcripts and higher lentiviral titres compared to the combination of U6 and CHb promoters.[211] CRISPR/Cas9 driven by the single H1 promoter is less efficient than

CRISPR/Cas9 produced from U6 and CHb, resulting in 77.6% target knockdown versus 95.9%.[211] While the editing efficiency is lower, there is likely a reduced risk for off-target effects. [211]

Another modification is to use RNA Polymerase II promoters to produce sgRNA. An advantage of RNA Polymerase II promoters is that they may allow for cell-specific, inducible, and temporally-controlled expression of the sgRNA.[212,213] For example, a rod-specific promoter could be used for Cas9 expression as a way to restrict CRISPR/Cas9 expression in only one cell type, whereas RNA polymerase III promoters that normally drive sgRNA expression from the U6 promoter are ubiquitous. However, transcripts produced by RNA Polymerase II are normally exported to the cytoplasm and undergo modifications, which may impede proper sgRNA function.[214]

In order to be expressed in a functional form, sgRNAs transcribed from RNA Polymerase II promoters must be flanked by ribonucleases. These ribonucleases allow for immediate cleavage of the sgRNA sequence from the transcript, preventing any further processing of the sgRNA.[215] Csy4 binding sites are conducive to successful sgRNA transcription production from CMV promoters.[216] The Csy4-flanked sgRNA sequences can be placed in either exons or introns[216]. If the sgRNA is placed within an exon, the addition of a 5' 110 bp sequence that forms a triple helical structure helps stabilize the transcript.[216] Another approach is to include a hammerhead ribozyme at the 5' end and an HDV ribozyme at the 3' end of the sgRNA sequence, which allows for self-cleavage of the transcript.[216] sgRNAs can also be cleaved by Drosha and Dicer if the appropriate recognition sequences flank the sgRNA region.[214] sgRNA linked by self-cleaving ribozymes or

tRNAs can be expressed from Polymerase II promoters like the muscle/heart-specific MHCK7 promoter.[217,218]

Guide RNA Modifications

Changes to the crRNA or tracrRNA sequences can improve editing efficiency and reduce off-target effects. Synthetic gRNAs (known as sgRNAs) contain shorter crRNA and tracrRNA sequences than do endogenous bacterial CRISPR systems.[18,219] Xu et al. 2017 found that increasing the crRNA and tracrRNA components to their original full length within the sgRNA can improve editing efficiency.[219] Furthermore, increasing the length of the sgRNA at the 3' end improves the efficiency of sgRNA and Cas9 binding.[220] In particular, including triple helical structures at the 3' end of sgRNAs may help stabilize them.[221] Extending the duplex region of the sgRNA by 5-10 base pairs and abolishing the poly-T transcriptional termination signal can also improve CRISPR/Cas9 efficiency.[222–224] SgRNA sequences containing a 3' G_2U_1 sequence are more stable and have higher *in vivo* editing rates and no increase in off-target effects compared to sgRNAs without this sequence.[225] The main issue with extending the sgRNA is that it may then be difficult to package within an AAV vector.

There are some changes to the sgRNA that do not require extending the length. For example, abolishing the poly-T transcriptional termination signal can improve CRISPR/Cas9 efficiency for both SpCas9 and SaCas9.[222–224] Additionally, choosing crRNA sequences with higher amounts of guanine nucleotides and fewer adenine nucleotides may improve stability and activity of the sgRNA.[226] Guanine enrichment at the 5' end may promote G-quadruplex formation, which further stabilizes the sgRNA.[226] Choosing shorter crRNAs may also be beneficial. crRNAs

that are 17-19 bp long have comparable on-target cleavage to 20 bp-long crRNAs and have lower off-target editing rates for SpCas9.[43,227] They are also more sensitive to single and double base pair mismatches.[227]

Another strategy that allows for a shorter guide RNA is to use Cas12a (Cpf1) enzymes, which require only a single crRNA without a tracrRNA. Some additional modifications to Cpf1 can improve its editing ability. For instance, changing the direct repeat sequence of the FnCpf1 crRNA enhances editing rates and reduces off-target cleavage.[228] Furthermore, including a 3' U_4AU_4 overhang improves AsCpf1 editing *in vivo*.[229] Extending the crRNA at the 5' end by up to 59 base pairs increases AsCpf1 ribonucleoprotein editing efficiency of primary mouse myoblasts.[230] This 59 bp addition still allows the AsCpf1 guide RNA to be shorter than the sgRNA sequence of SpCas9 and SaCas9.

Cas Enzyme Orthologs

Other Cas orthologs can be used, and the main advantage is that they target different regions of DNA because they each have unique PAM sequence requirements. The SpCas9 PAM site is 5'-NGG-3', and occurs approximately every 8 base pairs in the human genome.[18,26,37] NmCas9 recognizes the PAM site 5'-NNNNGATT-3', which exists every 128 base pairs.[231] Because the NmCas9 PAM site is more complex, there is a lower likelihood of off-target binding. However, it appears that NmCas9 has less on-target efficiency than SpCas9, perhaps due to its decreased ability to unwind DNA.[231] The newly-discovered ScCas9 recognizes 5'-NNG-3' as its PAM site; this PAM site is less restrictive than that of SpCas9, thus allowing ScCas9 to target more regions of DNA.[232]

The Cpf1 (Cas12a) variants are able to target AT-rich genomic regions; FnCpf1's PAM site is 5'-TTN3' while AsCpf1 and LbCpf1 recognize 5'-TTTV-3'.[233,234] Other advantages of the Cpf1 variants include that they create staggered double-stranded DNA breaks, which make it easier for the cell to perform homology-directed repair.[233] Furthermore, Cpf1 variants have a much shorter guide RNA than SpCas9 because they rely on only a crRNA sequence instead of both a crRNA and tracrRNA.[233]

Engineered Cas Enzymes

Some Cas9 orthologs have been engineered to recognize different PAM sites or to have refined DNA binding ability. Compared to wild-type SpCas9, the D1135E mutant has less binding to the non-canonical 5'-NAG-'3 PAM site and is thus more specific for the 5'-NGG-'3 PAM site.[235] The QQR1, EQR, VQR, VRQR, and VRER variants of SpCas9 recognize 5'-NAAG-3', 5'-NGAG-3', 5'-NGAN-3'/5'-NGNG-3', 5'-NGAN-3', and 5'-NGCG-3' PAM sites, respectively.[235,236] The xCas9 variant of SpCas9 has the broadest PAM site requirement, recognizing 5'-NG-3', 5'-GAA-3' and 5'-GAT-3'.[237,238] It has also higher on-target efficiency and lower off-target cleavage than wild-type SpCas9.[237] The SaCas9 KKH variant recognizes the PAM site 5'-NNNRRT-3', unlike SaCas9, which requires 5'-NNGRRT-3'.[170] The FnCas9 RHA mutant has a more relaxed PAM requirement (5'-YG-3') compared to wild-type FnCas9 (5'-NGG-3').[239] The SpCas9-HF1, eSpCas9, and HypaCas9 variants all display less off-target cleavage than wild-type SpCas9.[238,240,241]

In addition to engineering Cas9 variants, improved specificity of the CRISPR/Cas9 system can be achieved by combining the CRISPR/Cas9 system with other proteins, like FokI nucleases. In order to cleave target DNA, two FokI

nucleases must form a dimer. Tsai et al. 2014 have fused FokI nucleases to deactivated Cas9 with two sgRNAs.[242] The two sgRNAs bring the dCas-Foki proteins together, allowing the FokI subunits to dimerize and cut DNA.[242] This system can reduce off-target effects because it requires both binding of two sgRNA sequences and dimerization of the FokI domains for cleavage unlike traditional CRISPR/Cas9 systems, which simply rely on binding of one sgRNA to target DNA.[242] An alternative approach is to use two nickases, which each cleave one strand of DNA, to induce a double stranded break. Using two nickases reduces off-target cleavage by 50-1500 fold *in vitro*.[243] Furthermore, combining truncated sgRNAs with either Cas9-FokI or dCas9-Foki systems also decreases off-target effects.[227,244] Other strategies to reduce off-target cleavage include fusing Cas9 to zinc finger proteins and transcription activator-like effectors.[245] The main issue with incorporating FokI, zinc finger proteins, and transcription activator-like effectors into CRISPR/Cas systems is that the additional proteins may exceed the cargo capacity of AAV vectors.

5.5 Conclusion

The results presented in this thesis demonstrate that the orientation of the U6.sgRNA within SaCas9 and dSaCas9-KRAB AAV transgene constructs does not affect sgRNA and Cas9 expression. Furthermore, we have shown that GeoCas9 cannot cleave target DNA to detectable levels *in vitro* in HEK293-*eGFP* cells when delivered by plasmid transfection, despite previous work showing promise using RNP delivery. Future studies on retinal delivery of CRISPR/Cas9 can focus on testing the aforementioned optimizations. These experiments would be essential for successful translation of CRISPR/Cas9 into treatments for retinal disease.

6 References

1. Mojica, F. J. M., Juez, G. & Rodriguez-Valera, F. Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. *Mol. Microbiol.* **9**, 613–621 (1993).

2. Mojica, F. J. M., Ferrer, C., Juez, G. & Rodríguez-Valera, F. Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning. *Mol. Microbiol.* **17**, 85–93 (1995).

3. Jansen, R., Van Embden, J. D. A., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43**, 1565–1575 (2002).

4. Mojica, F. J. M. MicroCorrespondence: Biological significance of a family of regularly spaced repeats in the genomes of archaea, bacteria, and mitochondria. *Mol. Microbiol.* **36**, 244–246 (2000).

5. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005).

6. Lander, E. S. The heroes of CRISPR. *Cell* **164**, 18–28 (2016).

7. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).

8. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).

9. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).

10. Garneau, J. E. *et al.* The CRISPR/cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).

11. Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **9**, 467–477 (2011).

12. Nuñez, J. K., Lee, A. S. Y., Engelman, A. & Doudna, J. A. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* **519**, 193–198 (2015).

13. Nuñez, J. K. *et al.* Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* **21**, 528–534 (2014).

14. Gesner, E. M., Schellenberg, M. J., Garside, E. L., George, M. M. & MacMillan, A. M. Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol.* **18**, 688–692 (2011).

15. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355–1358 (2010).

16. Niewoehner, O., Jinek, M. & Doudna, J. A. Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. *Nucleic Acids Res.* **42**, 1341–1353 (2014).

17. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733–740 (2009).

18. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).

19. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *J. Bacteriol.* **190**, 1390–1400 (2008).

20. Horvath, P. *et al.* Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. *J. Bacteriol.* **190**, 1401–1412 (2008).

21. Marraffini, L. A. & Sontheimer, E. J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**, 568–571 (2010).

22. Samai, P. *et al.* Co-transcriptional DNA and RNA cleavage during type III CRISPR-cas immunity. *Cell* **161**, 1164–1174 (2015).

23. Osawa, T., Inanaga, H., Sato, C. & Numata, T. Crystal structure of the crispr-cas RNA silencing cmr complex bound to a target analog. *Mol. Cell* **58**, 418–430 (2015).

24. Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Classification and nomenclature of CRISPR-Cas systems: where from here? *Cris. J.* **1**, 325–336 (2018).

25. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 2579–2586 (2012).

26. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas9 systems. *Science* **339**, 819–823 (2013).

27. Ran, F. A. *et al.* In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520**, 186–191 (2015).

28. Swiech, L. *et al.* In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* **33**, 102–106 (2015).

29. Moore, J. K. & Haber, J. E. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **16**, 2164–2173 (1996).

30. Haber, J. E., Ira, G., Malkova, A. & Sugawara, N. Repairing a double-strand chromosome break by homologous recombination: Revisiting Robin Holliday's model. *Philos. Trans. R. Soc. B Biol. Sci.* **359**, 79–86 (2004).

31. Heyer, W.-D., Ehmsen, K. T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* **44**, 113–139 (2010).

32. Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife* **3**, e04766 (2014).

33. Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**, 144–149 (2016).

34. Borrelli, V. M. G., Brambilla, V., Rogowsky, P., Marocco, A. & Lanubile, A. The enhancement of plant disease resistance using CRISPR/Cas9 technology. *Front. Plant Sci.* **9**, (2018).

35. Listgarten, J. *et al.* Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. *Nat. Biomed. Eng.* **2**, 38–47 (2018).

36. Bae, S., Park, J. & Kim, J.-S. Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).

37. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).

38. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272–W276 (2016).

39. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).

40. Kim, D. *et al.* Digenome-seq: Genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* **12**, 237–243 (2015).

41. Tsai, S. Q. *et al.* CIRCLE-seq: A highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* **14**, 607–614 (2017).

42. Cameron, P. *et al.* Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods* **14**, 600–606 (2017).

43. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–198 (2015).

44. Hu, J. *et al.* Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing (LAM-HTGTS). *Nat. Protoc.* **11**, 853–871 (2016).

45. Yan, W. X. *et al.* BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun.* **8**, (2017).

46. Maeder, M. L. *et al.* Development of a gene-editing approach to restore vision loss in Leber Congenital Amaurosis type 10. *Nat. Med.* **25**, 229–233 (2019).

47. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).

48. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442 (2013).

49. Margolin, J. F. *et al.* Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4509–4513 (1994).

50. Friedman, J. R. *et al.* KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* **10**, 2067–2078 (1996).

51. Groner, A. C. *et al.* KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet.* **6**, (2010).

52. Thakore, P. I. *et al.* Highly specific epigenome editing by CRISPR/Cas9 repressors for silencing of distal regulatory elements. *Nat Methods* **12**, 1143–1149 (2015).

53. Adli, M. The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* **9**, (2018).

54. Liu, X. S. *et al.* Editing DNA methylation in the mammalian genome. *Cell* **167**, 233-247.e17 (2016).

55. McDonald, J. I. *et al.* Reprogrammable CRISPR/Cas9-based system for inducing sitespecific DNA methylation. *Biol. Open* **5**, 866–874 (2016).

56. Lei, Y. *et al.* Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nat. Commun.* **8**, (2017).

57. Mali, P. *et al.* Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833–838 (2013).

58. Maeder, M. L. *et al.* CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **10**, 977–979 (2013).

59. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10**, 973–976 (2013).

60. Chavez, A. *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).

61. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).

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

63. Friedman, D. S. *et al.* Prevalence of age-related macular degeneration in the United States. *Arch. Ophthalmol.* **122**, 564–5 (2004).

64. Klein, R., Klein, B. E. K. & Cruickshanks, K. J. The prevalence of age-related maculopathy by geographic region and ethnicity. *Prog. Retin. Eye Res.* **18**, 371–389 (1999).

65. Wong, W. L. *et al.* Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: A systematic review and meta-analysis. *Lancet Glob. Heal.* **2**, 106–116 (2014).

66. Leibowitz, H. *et al.* The Framingham eye study monograph: An ophthalmological and epidemiological study of cataract, glaucoma, diabetic retinopathy, macular degeneration, and visual acuity in a general population of 2631 adults, 1973-1975. *Surv Ophthalmol* **24**, 335–610 (1980).

67. Ferris, F. L., Fine, S. L. & Hyman, L. Age-related macular degeneration and blindness due to neovascular maculopathy. *Arch. Ophthalmol.* **102**, 1640–1642 (1984).

68. Aiello, L. P. *et al.* Suppression of retinal neovascularization in vivo by inhibition of vascular endothelial growth factor (VEGF) using soluble VEGF-receptor chimeric proteins. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 10457–10461 (1995).

69. Maberley, D. Pegaptanib for neovascular age-related macular degeneration. *Issues Emerg. Health Technol.* 1–4 (2005).

70. Edwards, A. O. *et al.* Complement factor H polymorphism and age-related macular degeneration. *Science* **308**, 421–424 (2005).

71. Fagerness, J. A. *et al.* Variation near complement factor I is associated with risk of advanced AMD. *Eur. J. Hum. Genet.* **17**, 100–104 (2009).

72. Gold, B. *et al.* Variation in factor B (BF) and complement component 2 (C2) genes is associated with age-related macular degeneration. *Nat Genet* **38**, 458–462 (2006).

73. Yates, J. R. W. *et al.* Complement C3 variant and the risk of age-related macular degeneration. *N. Engl. J. Med.* **357**, 553–561 (2007).

74. Gregory-Evans, K., Weleber, R. & Pennesi, M. Retinitis Pigmentosa and allied disorders. in *Ryan's Retina* (eds. Schachat, A. P., Sadda, S. R., Hinton, D. R., Wilkinson, C. P. & Wiedemann, P.) 861–935 (2017).

75. Puech, B., Kostriubiec, B., Hache, J. & Francois, P. Epidemiology and prevalence of hereditary retinal dystrophies in the Northern France. *J Fr Ophthalmol* **14**, 153–64 (1991).

76. Daiger, S. P., Sullivan, L. S. & Bowne, S. J. Genetic mechanisms of retinal disease. in *Ryan's Retina* (eds. Schachat, A. P., Sadda, S. R., Hinton, D. R., Wilkinson, C. & Wiedemann, P.) 711–721 (Elsevier Inc, 2018).

77. Daiger, S. P., Sullivan, L. S. & Bowne, S. J. RetNet: The retinal information network. (2019). Available at: http://www.sph.uth.tmc.edu/RetNet/.

78. Cooper, D. *et al.* HGMD : Human gene mutation database. (2018). Available at: http://www.hgmd.cf.ac.uk/.

79. Zhou, R. & Caspi, R. R. Ocular immune privilege. *F1000 Biol. Rep.* **2**, (2010).

80. Trapani, I. & Auricchio, A. Seeing the light after 25 years of retinal gene therapy. *Trends Mol. Med.* **24**, 669–681 (2018).

81. Sujirakul, T. *et al.* Multimodal imaging of central retinal disease progression in a 2 year mean follow up of Retinitis Pigmentosa. *Am J Ophthalmol* **160**, 1–25 (2015).

82. Boye, S. E., Boye, S. L., Lewin, A. S. & Hauswirth, W. W. A comprehensive review of retinal gene therapy. *Mol. Ther.* **21**, 509–519 (2013).

83. Bainbridge, J. W. B. *et al.* Effect of gene therapy on visual function in Leber's Congenital Amaurosis. *N. Engl. J. Med.* **358**, 2231–2239 (2008).

84. Bainbridge, J. W. B. *et al.* Long-term effect of gene therapy on Leber's Congenital Amaurosis. *N. Engl. J. Med.* **372**, 1887–1897 (2015).

85. Hauswirth, W. W. *et al.* Treatment of Leber Congenital Amaurosis due to RPE65 mutations by ocular subretinal injection of adeno-associated virus gene vector: short-term results of a phase I trial. *Hum. Gene Ther.* **19**, 979–990 (2008).

86. Cideciyan, A. V. *et al.* Human RPE65 gene therapy for leber congenital amaurosis: Persistence of early visual improvements and safety at 1 year. *Hum. Gene Ther.* **20**, 999–1004 (2009).

87. Jacobson, S. G. *et al.* Gene therapy for Leber Congenital Amaurosis caused by RPE65 mutations: Safety and efficacy in 15 children and adults followed up to 3 years. *Arch. Ophthalmol.* **130**, 9–24 (2012).

88. Weleber, R. G. *et al.* Results at 2 years after gene therapy for RPE65-deficient Leber Congenital Amaurosis and severe early-childhood-onset retinal dystrophy. *Ophthalmology* **123**, 1606–1620 (2016).

89. Pennesi, M. E. *et al.* Results at 5 years after gene therapy for RPE65-deficient retinal dystrophy. *Hum. Gene Ther.* **29**, 1428–1437 (2018).

90. Maguire, A. M. *et al.* Safety and efficacy of gene transfer for Leber's Congenital Amaurosis. *N. Engl. J. Med.* **358**, 2240–2248 (2008).

91. Maguire, A. M. *et al.* Age-dependent effects of RPE65 gene therapy for Leber's Congenital Amaurosis: a phase 1 dose-escalation trial. *Lancet* **374**, 1597–1605 (2009).

92. Bennett, J. *et al.* Safety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by RPE65 mutatons: A follow-on phase 1 trial Prof. *Lancet* **388**, 661–672 (2016).

93. Russell, S. *et al.* Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *Lancet* **390**, 849–860 (2017).

94. FDA approves novel gene therapy to treat patients with a rare form of inherited vision loss. (2017). Available at: https://www.fda.gov/news-events/press-announcements/fda-approves-novel-gene-therapy-treat-patients-rare-form-inherited-vision-loss.

95. MacLaren, R. E. *et al.* Retinal gene therapy in patients with choroideremia: Initial findings from a phase 1/2 clinical trial. *Lancet* **383**, 1129–1137 (2014).

96. Edwards, T. L., Jolly, J. K., Groppe, M. & Barnard, A. R. Visual acuity after retinal gene therapy for Choroideremia. *N. Engl. J. Med.* **374**, 1996–1998 (2016).

97. Xue, K. *et al.* Beneficial effects on vision in patients undergoing retinal gene therapy for choroideremia. *Nat. Med.* **24**, 1507–1512 (2018).

98. Fischer, M. D. *et al.* Changes in retinal sensitivity after gene therapy in Choroideremia. *Retina* 1 (2018). doi:10.1097/iae.0000000000002360

99. Lam, B. L. *et al.* Choroideremia gene therapy phase 2 clinical trial: 24-month results. *Am. J. Ophthalmol.* **197**, 65–73 (2019).

100. Dimopoulos, I. S. *et al.* Two-year results after AAV2-mediated gene therapy for Choroideremia: The Alberta experience. *Am. J. Ophthalmol.* **193**, 130–142 (2018).

101. Wan, X. *et al.* Efficacy and safety of rAAV2-ND4 treatment for Leber's Hereditary Optic Neuropathy. *Sci. Rep.* **6**, 2–11 (2016).

102. DiCarlo, J. E., Mahajan, V. B. & Tsang, S. H. Gene therapy and genome surgery in the retina. *J. Clin. Invest.* **128**, 2177–2188 (2018).

103. Feuer, W. J. *et al.* Gene therapy for Leber Hereditary Optic Neuropathy. *Ophthalmology* **123**, 558–570 (2016).

104. Rakoczy, E. P. *et al.* Gene therapy with recombinant adeno-associated vectors for neovascular age-related macular degeneration: 1 year follow-up of a phase 1 randomised clinical trial. *Lancet* **386**, 2395–2403 (2015).

105. Constable, I. J. *et al.* Phase 2a randomized clinical trial: Safety and post hoc analysis of subretinal rAAV.sFLT-1 for wet age-related macular degeneration. *EBioMedicine* **14**, 168–175 (2016).

106. Campochiaro, P. A. *et al.* Lentiviral vector gene transfer of endostatin/angiostatin for macular degeneration (GEM) study. *Hum. Gene Ther.* **28**, 99–111 (2017).

107. Branham, K. *et al.* Mutations in RPGR and RP2 account for 15% of males with simplex retinal degenerative disease. *Investig. Ophthalmol. Vis. Sci.* **53**, 8232–8237 (2012).

108. Fischer, M. D. *et al.* Codon-optimized RPGR improves stability and efficacy of AAV8 gene therapy in two mouse models of X-linked Retinitis Pigmentosa. *Mol. Ther.* **25**, 1854–1865

(2017).

109. Hung, S. S. C., McCaughey, T., Swann, O., Pébay, A. & Hewitt, A. W. Genome engineering in ophthalmology: Application of CRISPR/Cas to the treatment of eye disease. *Prog. Retin. Eye Res.* **53**, 1–20 (2016).

110. Gupta, R. M. & Musunuru, K. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *J. Clin. Invest.* **124**, 4154–4161 (2014).

111. Yu, W. & Wu, Z. In Vivo applications of CRISPR-based genome editing in the retina. *Front. Cell Dev. Biol.* **6**, (2018).

112. Kim, K. *et al.* Genome surgery using Cas9 ribonucleoproteins for the treatment of age-related macular degeneration. *Genome Res.* **27**, 419–426 (2017).

113. Zincarelli, C., Soltys, S., Rengo, G. & Rabinowitz, J. E. Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol. Ther.* **16**, 1073–1080 (2008).

114. Kaeppel, C. *et al.* A largely random AAV integration profile after LPLD gene therapy. *Nat. Med.* **19**, 889–891 (2013).

115. McCarty, D. M., Young, S. M. & Samulski, R. J. Integration of adeno-associated virus (AAV) and recombinant AAV vectors. *Annu. Rev. Genet.* **38**, 819–845 (2004).

116. Mays, L. E. & Wilson, J. M. The complex and evolving story of T cell activation to AAV vector-encoded transgene products. *Mol. Ther.* **19**, 16–27 (2011).

117. Sullivan, L. S. *et al.* Prevalence of disease-causing mutations in families with autosomal dominant Retinitis Pigmentosa: A screen of known genes in 200 families. *Invest Ophthalmol Vis Sci* **47**, 3052–3064 (2006).

118. Hargrave, P. A. & McDowell, J. H. Rhodopsin and phototransduction: A model system for G protein-linked receptors. *FASEB J.* **6**, 2323–2331 (1992).

119. Farrar, G. J. *et al.* Autosomal dominant Retinitis Pigmentosa: Absence of the rhodopsin proline→histidine substitution (codon 23) in pedigrees from Europe. *Am. J. Hum. Genet.* **47**, 941–945 (1990).

120. Athanasiou, D. *et al.* The molecular and cellular basis of rhodopsin retinitis pigmentosa reveals potential strategies for therapy. *Prog. Retin. Eye Res.* **62**, 1–23 (2018).

121. Illing, M. E., Rajan, R. S., Bence, N. F. & Kopito, R. R. A rhodopsin mutant linked to autosomal dominant retinitis pigmentosa is prone to aggregate and interacts with the ubiquitin proteasome system. *J. Biol. Chem.* **277**, 34150–34160 (2002).

122. Saliba, R. S., Munro, P. M. G., Luthert, P. J. & Cheetham, M. E. The cellular fate of mutant rhodopsin: Quality control, degradation and aggresome formation. *J. Cell Sci.* **115**, 2907–2918 (2002).

123. Rajan, R. S. & Kopito, R. R. Suppression of wild-type rhodopsin maturation by mutants linked to autosomal dominant retinitis pigmentosa. *J. Biol. Chem.* **280**, 1284–1291 (2005).

124. Latella, M. C. *et al.* In vivo editing of the human mutant rhodopsin gene by electroporation of plasmid-based CRISPR/Cas9 in the mouse retina. *Mol. Ther. - Nucleic Acids* **5**, e389 (2016).

125. Tsai, Y. T. *et al.* Clustered regularly interspaced short palindromic repeats-based genome

surgery for the treatment of autosomal dominant Retinitis Pigmentosa. *Ophthalmology* **125**, 1421–1430 (2018).

126. Bakondi, B. *et al.* In vivo CRISPR/Cas9 gene editing corrects retinal dystrophy in the S334ter-3 rat model of autosomal dominant Retinitis Pigmentosa. *Mol. Ther.* **24**, 556–563 (2016).

127. Giannelli, S. G. *et al.* Cas9/sgRNA selective targeting of the P23H rhodopsin mutant allele for treating Retinitis Pigmentosa by intravitreal AAV9.PHP.B-based delivery. *Hum. Mol. Genet.* **27**, 761–779 (2018).

128. Li, P. *et al.* Allele-specific CRISPR-Cas9 genome editing of the single-base P23H mutation for rhodopsin-associated dominant Retinitis Pigmentosa. *Cris. J.* **1**, 55–64 (2018).

129. Wu, W. H. *et al.* CRISPR repair reveals causative mutation in a preclinical model of Retinitis Pigmentosa. *Mol. Ther.* **24**, 1388–1394 (2016).

130. Cai, Y. *et al.* In vivo genome editing rescues photoreceptor degeneration via a Cas9/RecA-mediated homology-directed repair pathway. *Sci. Adv.* **5**, 1–13 (2019).

131. Chang, B. *et al.* In-frame deletion in a novel centrosomal/ciliary protein CEP290/NPHP6 perturbs its interaction with RPGR and results in early-onset retinal degeneration in the rd16 mouse. *Hum Mol Genet* **1**, 1847–1857 (2006).

132. Drivas, T. & Bennett, J. CEP290 and the primary cilium. in *Retinal Degenerative Diseases. Advances in Experimental Medicine and Biology* (eds. Ash, J. et al.) (2014).

133. Den Hollander, A. I. *et al.* Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber Congenital Amaurosis. *Am. J. Hum. Genet.* **79**, 556–561 (2006).

134. Stone, E. M. *et al.* Clinically focused molecular investigation of 1000 consecutive families with inherited retinal disease. *Ophthalmology* **124**, 1314–1331 (2017).

135. Ruan, G. X. *et al.* CRISPR/Cas9-mediated genome editing as a therapeutic approach for Leber Congenital Amaurosis 10. *Mol. Ther.* **25**, 331–341 (2017).

136. Sheridan, C. M. *et al.* Expression of hypoxia-inducible factor-1α and -2α in human choroidal neovascular membranes. *Graefe's Arch. Clin. Exp. Ophthalmol.* **247**, 1361–1367 (2009).

137. André, H., Tunik, S., Aronsson, M. & Kvanta, A. Hypoxia-inducible factor-1α Is associated with sprouting angiogenesis in the murine laser-induced choroidal neovascularization model. *Investig. Ophthalmol. Vis. Sci.* **56**, 6591–6604 (2015).

138. Kwak, N., Okamoto, N., Wood, J. M. & Campochiaro, P. A. VEGF is major stimulator in model of choroidal neovascularization. *Investig. Ophthalmol. Vis. Sci.* **41**, 3158–3164 (2000).

139. Kim, E. *et al.* In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. *Nat. Commun.* **8**, (2017).

140. Jo, D. H. *et al.* Long-term effects of in vivo genome editing in the mouse retina using Campylobacter jejuni Cas9 Expressed via adeno-associated virus. *Mol. Ther.* **27**, 130–136 (2019).

141. Holmgaard, A. *et al.* In vivo knockout of the Vegfa gene by lentiviral delivery of CRISPR/Cas9 in mouse retinal pigment epithelium cells. *Mol. Ther. - Nucleic Acids* **9**, 89–99 (2017).

142. Huang, X. *et al.* Genome editing abrogates angiogenesis in vivo. *Nat. Commun.* **8**, 4–11 (2017).

143. Koo, T. *et al.* CRISPR-LbCpf1 prevents choroidal neovascularization in a mouse model of age-related macular degeneration. *Nat. Commun.* **9**, 6–13 (2018).

144. Yu, W. *et al.* Nrl knockdown by AAV-delivered CRISPR/Cas9 prevents retinal degeneration in mice. *Nat. Commun.* **8**, (2017).

145. Zhu, J. *et al.* Gene and mutation independent therapy via CRISPR-Cas9 mediated cellular reprogramming in rod photoreceptors. *Cell Res.* **27**, 830–833 (2017).

146. Peirson, S. N., Butler, J. N. & Foster, R. G. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucleic Acids Res.* **31**, e73 (2003).

147. Fujii, W., Kawasaki, K., Sugiura, K. & Naito, K. Efficient generation of large-scale genome-modified mice using gRNA and Cas9 endonuclease. *Nucleic Acids Res.* **41**, (2013).

148. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839–843 (2013).

149. Ren, X. *et al.* Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in Drosophila. *Cell Rep.* **9**, 1151–1162 (2014).

150. Min, Y. L. *et al.* CRISPR-Cas9 corrects Duchenne muscular dystrophy exon 44 deletion mutations in mice and human cells. *Sci. Adv.* **5**, 1–13 (2019).

151. Grimm, D. *et al.* Fatality in mice due to oversaturation of cellular microRNA / short hairpin RNA pathways. *Nature* **441**, 537–541 (2006).

152. Shearwin, K. E., Callen, B. P. & Egan, J. B. Transcriptional interference--A crash course. *Trends Genet* **21**, 339–345 (2005).

153. Callen, B. P., Shearwin, K. E. & Egan, J. B. Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol. Cell* **14**, 647–656 (2004).

154. Chatterjee, A. *et al.* Convergent transcription confers a bistable switch in Enterococcus faecalis conjugation. *PNAS* **108**, 9721–9726 (2011).

155. Ward, D. F. & Murray, N. E. Convergent transcription in bacteriophage lambda: Interference with gene expression. *J Mol Biol* **133**, 249–266 (1979).

156. Elledge, S. J. & Davis, R. W. Position and density effects on repression by stanonary and mobile DNA-binding proteins. *Genes Dev.* **3**, 185–197 (1989).

157. Prescott, E. M. & Proudfoot, N. J. Transcriptional collision between convergent genes in budding yeast. *PNAS* **99**, 8796–8801 (2002).

158. Hobson, D. J., Wei, W., Steinmetz, L. M. & Svejstrup, J. Q. RNA polymerase II collision interrupts convergent transcription. *Mol. Cell* **48**, 365–374 (2012).

159. Crampton, N., Bonass, W. A., Kirkham, J., Rivetti, C. & Thomson, N. H. Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. *Nucleic Acids Res.* **34**, 5416–5425 (2006).

160. Adhya, S. & Gottesman, M. Promoter occlusion: Transcription through a promoter may inhibit its activity. *Cell* **29**, 939–944 (1982).

161. Kouzine, F., Liu, J., Sanford, S., Chung, H. J. & Levens, D. The dynamic response of upstream DNA to transcription-generated torsional stress. *Nat. Struct. Mol. Biol.* **11**, 1092–1100 (2004).

162. Ma, J., Bai, L. & Wang, M. D. Transcription under torsion. *Science* **340**, 1580–1583 (2013).

163. Kouzine, F., Sanford, S., Elisha-Feil, Z. & Levens, D. The functional response of upstream DNA to dynamic supercoiling in vivo. *Nat. Struct. Mol. Biol.* **15**, 146–154 (2008).

164. Brophy, J. A. & Voigt, C. A. Antisense transcription as a tool to tune gene expression. *Mol. Syst. Biol.* **12**, 854–854 (2016).

165. Kholod, N. & Mustelin, T. Elimination of transcriptional interference between tandem genes in plant cells. *Biotechniques* **31**, 328–334 (2001).

166. Eszterhas, S. K., Bouhassira, E. E., Martin, D. I. K. & Fiering, S. Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol. Cell. Biol.* **22**, 469–479 (2002).

167. Bae, J. Y., Laplaza, J. & Jeffries, T. W. Effects of gene orientation and use of multiple promoters on the expression of XYL1 and XYL2 in Saccharomyces cerevisiae. *Appl Biochem Biotechnol* **145**, 69–78 (2008).

168. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, 1–12 (2002).

169. Derveaux, S., Vandesompele, J. & Hellemans, J. How to do successful gene expression analysis using real-time PCR. *Methods* **50**, 227–230 (2010).

170. Kleinstiver, B. P. *et al.* Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* **33**, 1293–1298 (2015).

171. Woychik, N. A. & Hampsey, M. The RNA polymerase II machinery: Structure illuminates function. *Cell* **108**, 453–463 (2002).

172. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 ångstrom resolution. *Science* **292**, 1863–1876 (2001).

173. Higgins, N. P. & Vologodskii, A. V. Topological behaviour of plasmid DNA. *Microbiol. Spectr.* **3**, 1–49 (2015).

174. Szlachta, K. *et al.* Alternative DNA secondary structure formation affects RNA polymerase II promoter-proximal pausing in human. *Genome Biol.* **19**, 1–19 (2018).

175. Schek, N., Cooke, C. & Alwine, J. C. Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analyses. *Mol. Cell. Biol.* **12**, 5386–5393 (1992).

176. Gil, A. & Proudfoot, N. J. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β-globin mRNA 3′ end formation. *Cell* **49**, 399–406 (1987).

177. Allison, D. S. & Hall, B. D. Effects of alterations in the 3' flanking sequence on in vivo and in vitro expression of the yeast SUP4-o tRNAtyr gene. *EMBO J.* **4**, 2657–2664 (1985).

178. Hamada, M., Sakulich, A. L., Koduru, S. B. & Maraia, R. J. Transcription termination by RNA polymerase III in fission yeast. *J. Biol. Chem.* **275**, 29076–29081 (2000).

179. Braglia, P., Percudani, R. & Dieci, G. Sequence context effects on oligo(dT) termination

signal recognition by Saccharomyces cerevisiae RNA polymerase III. *J. Biol. Chem.* **280**, 19551–19562 (2005).

180. Gunnery, S., Ma, Y. & Mathews, M. B. Termination sequence requirements vary among genes transcribed by RNA polymerase III. *J. Mol. Biol.* **286**, 745–757 (1999).

181. Nielsen, S., Yuzenkova, Y. & Zenkin, N. Mechanism of eukaryotic RNA polymerase III transcription termination. *Science* **340**, 1577–1580 (2013).

182. Hoffmann, S. A., Hao, N., Shearwin, K. E. & Arndt, K. M. Characterizing transcriptional interference between converging genes in bacteria. *ACS Synth. Biol.* **8**, 466–473 (2019).

183. Pelechano, V. & Steinmetz, L. M. Gene regulation by antisense transcription. *Nat. Rev. Genet.* **14**, 880–893 (2013).

184. Harrington, L. B. *et al.* A thermostable Cas9 with increased lifetime in human plasma. *Nat. Commun.* **8**, 1–8 (2017).

185. Peng, J. *et al.* Efficient gene editing in adult mouse livers via adenoviral delivery of CRISPR/Cas9. *FEBS Lett.* **588**, 3954–3958 (2014).

186. Wang, D. *et al.* Adenovirus-mediated somatic genome editing of Pten by CRISPR/Cas9 in mouse liver in spite of Cas9-specific immune responses. *Hum. Gene Ther.* **26**, 432–442 (2015).

187. Annunziato, S. *et al.* Modeling invasive lobular breast carcinoma by CRISPR/Cas9-mediated somatic genome editing of the mammary gland. *Genes Dev.* **30**, 1470–1480 (2016).

188. Chew, W. L. *et al.* A multi-functional AAV-CRISPR-Cas9 and its host response. *Nat Methods* **13**, 868–874 (2016).

189. Guan, Y. *et al.* CRISPR/Cas9-mediated somatic correction of a novel coagulator factor IX gene mutation ameliorates hemophilia in mouse. *EMBO Mol. Med.* **8**, 477–488 (2016).

190. Platt, R. J. *et al.* CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014).

191. Simhadri, V. L. *et al.* Prevalence of pre-existing antibodies to CRISPR-Associated nuclease Cas9 in the USA population. *Mol. Ther. - Methods Clin. Dev.* **10**, 105–112 (2018).

192. Charlesworth, C. T. *et al.* Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019).

193. Wagner, D. L. *et al.* High prevalence of S. pyogenes Cas9-reactive T cells within the adult human population. *Nat. Med.* **25**, 295139 (2018).

194. Novik, G., Savich, V. & Meerovskaya, O. Geobacillus bacteria: Potential commercial applications. in *Industry, Bioremediation, and Bioenergy Production* (2018). doi:10.5772/intechopen.76053

195. Xie, Z. *et al.* Co-transfection and tandem transfection of HEK293A cells for overexpression and RNAi experiments. *Cell Biol. Int.* **35**, 187–192 (2011).

196. Boruah, J. L. H. *et al.* Effect of co-transfection of anti-myostatin shRNA constructs in caprine fetal fibroblast cells. *Anim. Biotechnol.* **27**, 44–51 (2016).

197. Yarrington, R. M., Verma, S., Schwartz, S., Trautman, J. K. & Carroll, D. Nucleosomes inhibit target cleavage by CRISPR-Cas9 in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 9351–

9358 (2018).

198. Knight, S. C. *et al.* Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science* **350**, 823–826 (2015).

199. Tang, L. *et al.* Efficient cleavage resolves PAM preferences of CRISPR-Cas in human cells. *Cell Regen.* **8**, 44–50 (2019).

200. Cho, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Cold Spring Harb. Lab. Press Method* **24**, 132–141 (2014).

201. Goomer, R. & Kunkel, G. The transcriptional start site for a human U6 small nuclear RNA gene is dictated by a compound promoter element consisting of the PSE and the TATA box. *Nucleic Acids Res.* **20**, 4903–4912 (1992).

202. Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P. & Pruett-Miller, S. M. A survey of validation strategies for CRISPR-Cas9 editing. *Sci. Rep.* **8**, 1–8 (2018).

203. Kim, D., Kim, S., Kim, S., Park, J. & Kim, J. S. Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq. *Genome Res.* **26**, 406–415 (2016).

204. Walsh, G. & Jefferis, R. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* **24**, 1241–1252 (2006).

205. Jarvis, D. L., Kawar, Z. S. & Hollister, J. R. Engineering N-glycosylation pathways in the baculovirus-insect cell system. *Curr. Opin. Biotechnol.* **9**, 528–533 (1998).

206. Cabanes-Macheteau, M. *et al.* N-Glycosylation of a mouse IgG expressed in transgenic tobacco plants. *Glycobiology* **9**, 365–372 (1999).

207. Choder, M. & Aloni, Y. RNA polymerase II allows unwinding and rewinding of the DNA and thus maintains a constant length of the transcription bubble. *J. Biol. Chem.* **263**, 12994–13002 (1988).

208. Han, Y., Yan, C., Fishbain, S., Ivanov, I. & He, Y. Structural visualization of RNA polymerase III transcription machineries. *Cell Discov.* **4**, (2018).

209. Kim, S., Kim, D., Cho, S., Kim, J. & Kim, J.-S. Highly efficient RNA-guide genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **128**, 1–32 (2014).

210. Ranganathan, V., Wahlin, K., Maruotti, J. & Zack, D. J. Expansion of the CRISPR-Cas9 genome targeting space through the use of H1 promoter-expressed guide RNAs. *Nat. Commun.* **5**, (2014).

211. Gao, Z., Herrera-Carrillo, E. & Berkhout, B. A single H1 promoter can drive both guide RNA and endonuclease expression in the CRISPR-Cas9 system. *Mol. Ther. - Nucleic Acids* **14**, 32–40 (2019).

212. Chen, X., Wu, J.-M., Hornischer, K., Kel, A. & Winegender, E. TiProD: the tissue-specific promoter database. *Nucleic Acids Res.* **34**, D104–D107 (2006).

213. Gauthier, N. P., Jensen, L. J., Wernersson, R., Brunak, S. & Jensen, T. S. Cyclebase.org: Version 2.0, an updated comprehensive, multi-species repository of cell cycle experiments and derived analysis results. *Nucleic Acids Res.* **38**, 699–702 (2009).

214. Xie, C. *et al.* SgRNA expression of CRIPSR-Cas9 system based on miRNA polycistrons as a versatile tool to manipulate multiple and tissue-specific genome editing. *Sci. Rep.* **7**, 1–12

(2017).

215. Gao, Y. & Zhao, Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J. Integr. Plant Biol.* **56**, 343–349 (2014).

216. Nissim, L., Perli, S. D., Fridkin, A., Perez-Pinera, P. & Lu, T. K. Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Mol. Cell* **54**, 698–710 (2014).

217. Xu, L., Zhao, L., Gao, Y., Xu, J. & Han, R. Empower multiplex cell and tissue-specific CRISPR-mediated gene manipulation with self-cleaving ribozymes and tRNA. *Nucleic Acids Res.* **45**, 1–9 (2017).

218. Ding, D., Chen, K., Chen, Y., Li, H. & Xie, K. Engineering introns to express RNA guides for Cas9- and Cpf1-mediated multiplex genome editing. *Mol. Plant* **11**, 542–552 (2018).

219. Xu, J., Lian, W., Jia, Y., Li, L. & Huang, Z. Optimized guide RNA structure for genome editing via Cas9. *Oncotarget* **8**, 94166–94171 (2017).

220. Mekler, V., Minakhin, L., Semenova, E., Kuznedelov, K. & Severinov, K. Kinetics of the CRISPR-Cas9 effector complex assembly and the role of 3′-terminal segment of guide RNA. *Nucleic Acids Res.* **44**, 2837–2845 (2016).

221. Wilusz, J. E. *et al.* A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* **26**, 2392–2407 (2012).

222. Dang, Y. *et al.* Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol.* **16**, 1–10 (2015).

223. Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).

224. Chen, B. *et al.* Expanding the CRISPR imaging toolset with Staphylococcus aureus Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res.* **44**, e75 (2016).

225. Nahar, S. *et al.* A G-quadruplex motif at the 3′ end of sgRNAs improves CRISPR-Cas9 based genome editing efficiency. *Chem. Commun.* **54**, 2377–2380 (2018).

226. Moreno-Mateos, M. A. *et al.* CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* **12**, 982–988 (2015).

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

228. Lin, L. *et al.* Engineering the direct repeat sequence of crRNA for optimization of FnCpf1-mediated genome editing in human cells. *Mol. Ther.* **26**, 2650–2657 (2018).

229. Moon, S. Bin, Kim, D. Y., Ko, J. H., Kim, J. S. & Kim, Y. S. Improving CRISPR genome editing by engineering guide RNAs. *Trends Biotechnol.* **37**, 870–881 (2019).

230. Park, H. M. *et al.* Extension of the crRNA enhances Cpf1 gene editing in vitro and in vivo. *Nat. Commun.* **9**, 1–12 (2018).

231. Lee, C. M., Cradick, T. J. & Bao, G. The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells. *Mol. Ther.* **24**, 645–654 (2016).

232. Chatterjee, P., Jakimo, N. & Jacobson, J. M. Minimal PAM specificity of a highly similar SpCas9 ortholog. *Sci. Adv.* **4**, 1–11 (2018).

233. Zetsche, B. *et al.* Cpf1 Is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).

234. Kim, H. K. *et al.* In vivo high-throughput profiling of CRISPR-Cpf1 activity. *Nat. Methods* **14**, 153–159 (2017).

235. Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).

236. Anders, C., Bargsten, K. & Jinek, M. Structural plasticity of PAM recognition by engineered variants of the RNA-guided endonuclease Cas9. *Mol. Cell* **61**, 895–902 (2016).

237. Hu, J. H. *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57–63 (2018).

238. Kleinstiver, B. P. *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).

239. Hirano, H. *et al.* Structure and engineering of Francisella novicida Cas9. *Cell* **164**, 950–961 (2016).

240. Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).

241. Chen, J. S. *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407–410 (2017).

242. Tsai, S. Q. *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* **32**, 569–576 (2014).

243. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).

244. Wyvekens, N., Topkar, V. V., Khayter, C., Joung, J. K. & Tsai, S. Q. Dimeric CRISPR RNA-guided FokI-dCas9 nucleases directed by truncated gRNAs for highly specific genome editing. *Hum. Gene Ther.* **26**, 425–431 (2015).

245. Bolukbasi, M. F. *et al.* DNA-binding-domain fusions enhance the targeting range and precision of Cas9. *Nat. Methods* **12**, 1150–1156 (2015).

7 Appendix

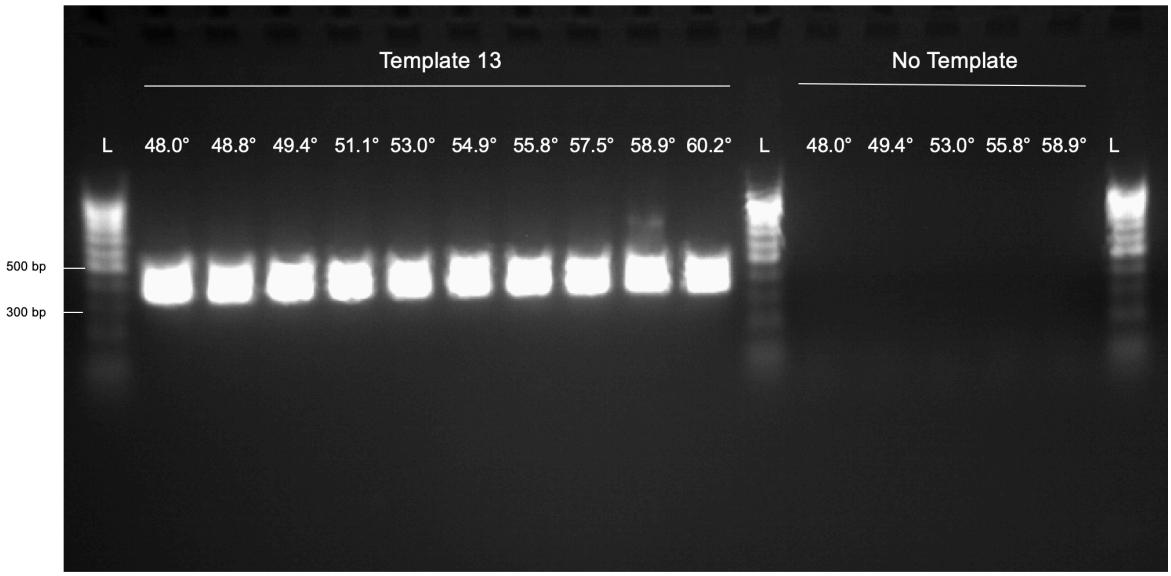

Figure 7.1 Optimizing Annealing Temperature for PCR Amplification of U6.sgRNA

The U6.sgRNA region was amplified using the KOD Hot Start DNA Polymerase, as described in chapter 2, with 30 PCR cycles using the primer pairs NotI_hU6_F2 and KpnI_gRNA_R1. The different annealing temperatures are shown above each well on the gel electrophoresis. The bands produced by the PCR are of the appropriate size (365 bp), when compared to the DNA ladder (L). There was no amplification of DNA in the no template control samples, indicating that there is no contamination of the PCR reagents. The annealing temperature did not seem to make a difference in the quality of DNA produced by the PCR. An annealing temperature of 55°C was chosen for all subsequent PCR amplifications with this primer pair.

Protocol #	Amount of DNA Added in ng	Comb Size	Final DNA Concentration	260/280	260/230
1	1000 ng	10 well	5.0	2.17	0.01
2	2000 ng	16 well	21.5	2.17	0.06

Figure 7.2 Optimizing Digestion of the Vector Plasmid

The vector plasmid containing dSaCas9 and KRAB was digested with the NotI and KpnI restriction enzymes and purified, as detailed in chapter 2. Two different protocols were tested. The first one used 1000 ng of vector DNA for the restriction digest and a 10-well agarose gel comb to create wells in the 1% agarose gel for gel electrophoresis. The second protocol used 2000 ng of vector DNA in the reaction and a 16-well comb for gel electrophoresis. Protocol #1 yielded a lower final DNA concentration compared to protocol #2. Both protocols resulted in similar 260/280 and 260/230 values, which indicate purity of the sample. Having more initial template DNA resulted in a higher amount of DNA after gel purification. Using a smaller gel comb (16- versus 10-well) did not appear to change the purity of the final sample.

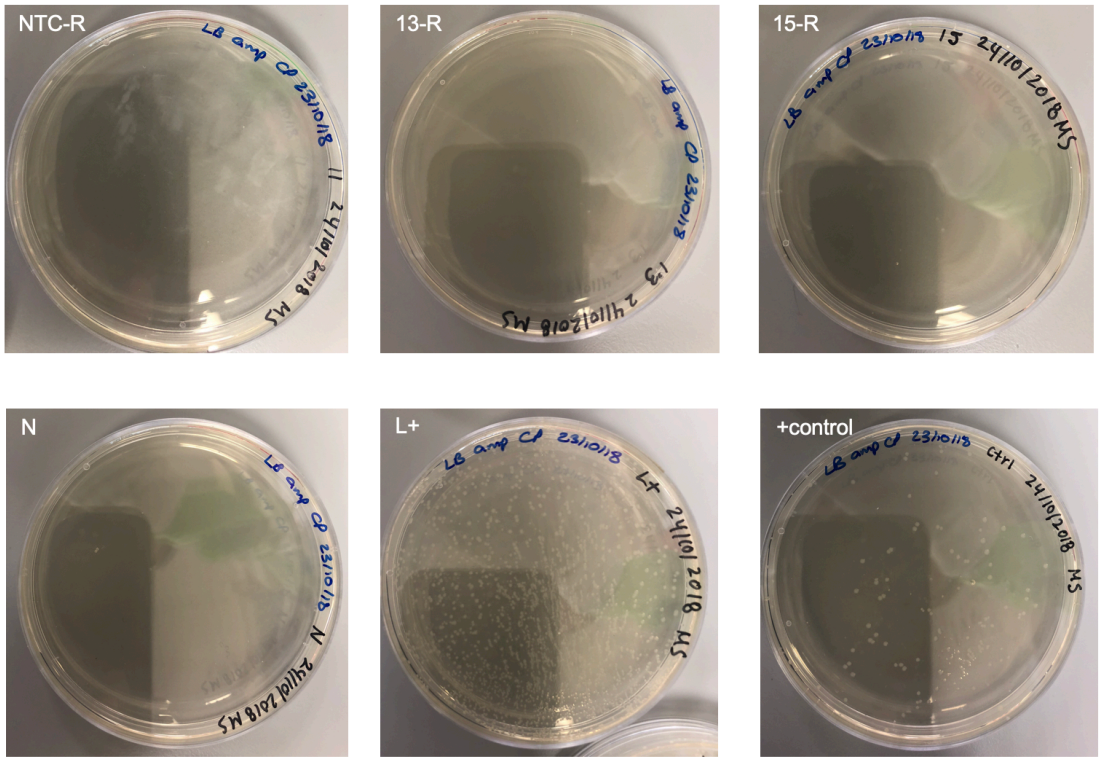

Figure 7.3 Optimizing Ligation of the Digested Inserts (U6.sgRNA) into the Digested DSaCas9-KRAB Vector (5:1 insert:vector ratio with 20 ng vector)

For these ligations, a 5:1 insert:vector ratio was used with 20 ng vector. The ligated plasmids were then transformed into XL-10 Ultracompetent *Escherichia coli* cells (Strategene, USA) and placed on the above ampicillin-containing LB agar plates. There is no growth of bacterial colonies on the plates with inserts NTC, 13, and 15 ligated into the vector (plates NTC-R, 13-R, and 15-R). This lack of growth is either due to a failure of the bacterial transformation or a failure in the ligation step.

The transformation reaction was successful, as indicated by growth of bacterial colonies on the positive control plate (+control). The ligase enzyme is functioning properly, as indicated by presence of bacterial colonies on the plate with linearized plasmid incubated with ligase (L+). There is no re-circularization of the plasmid, as there are no bacterial colonies on the plate with digested vector and no insert (N). The lack of bacterial growth on plates NTC-R, 13-R, and 15-R is therefore due to a failure in the ligation reaction.

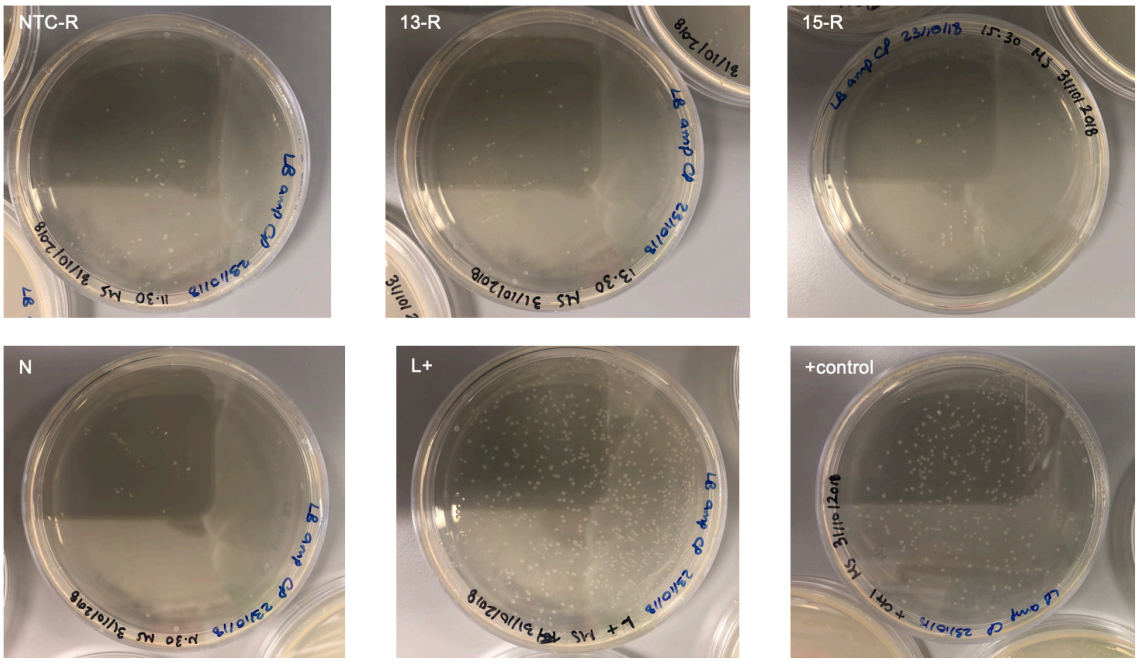

Figure 7.4 Optimizing Ligation of the Digested Inserts (U6.sgRNA) into the Digested DSaCas9-KRAB Vector (10:1 insert:vector ratio with 30 ng vector)

For these ligations, a 10:1 insert:vector ratio was used with 30 ng vector. The ligated plasmids were then transformed into XL-10 Ultracompetent *Escherichia coli* cells (Strategene, USA) and placed on the above ampicillin-containing LB agar plates. There is some growth of bacterial colonies on the plates with U6.sgRNA inserts NTC, 13, and 15 ligated into the vector (plates NTC-R, 13-R, and 15-R). It appears that these inserts have been successfully ligated into the backbone. In this reaction, there is very little re-circularized vector, as indicated by the minimal colony growth on the plate with digested vector, ligase, and no insert (N). The ligase step is functioning properly, as indicated by bacterial growth on the plate with linearized plasmid incubated with ligase (L+). Increasing the ratio of insert to vector from 5:1 (see Figure 7.3) to 10:1 allowed for a successful ligation reaction.

Plasmid	Starter Culture Time (hours)	Temperature of Starter and Overnight Cultures	Concentration in ng/uL	260/280	260/230
NTC-R	8	37°C	27.8	1.92	1.97
13-R	8	37°C	138.0	1.92	2.27
15-R	8	37°C	75.9	1.98	2.12
NTC-R	12	37°C	593.3	2.03	2.45
13-R	12	37°C	773.3	2.07	2.41
15-R	12	37°C	741.6	2.08	2.39
NTC-R	8	30°C	1353.8	2.05	2.47
13-R	8	30°C	1489.1	2.07	2.42
15-R	8	30°C	1443.9	2.06	2.45

Figure 7.5 Optimizing Maxiprep of Plasmids for the DSaCas9-KRAB Study

Optimizations to the incubation time of the starter culture and the incubation temperature of the overnight culture for the maxiprep were required for high plasmid yield and high-quality sequencing results. There was a low plasmid yield following the standard maxiprep protocol (8 hour incubation of the starter culture at 37°C and then inoculation of an overnight culture followed by a 12-16 hour incubation at 37°C). There was a higher plasmid yield following an increased starter culture incubation (12 hours). There was also a higher plasmid yield following incubation of both starter cultures (8 hours) and overnight cultures at 30°C. The restriction digests of the above samples are shown in Figure 7.6.

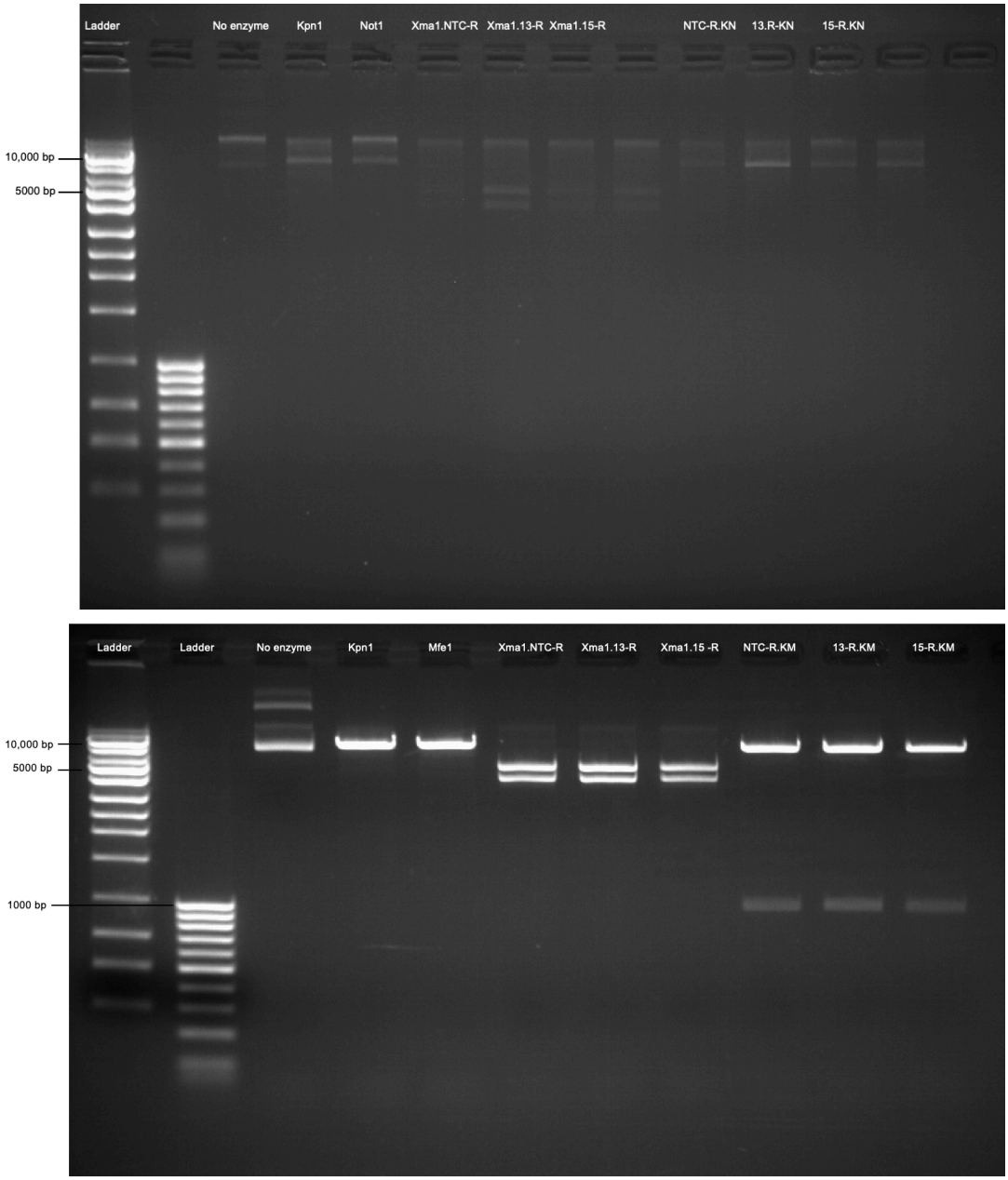

Figure 7.6 Restriction Digests of Maxiprepped Plasmids for the DSaCas9-KRAB Study

Restriction digests were performed on maxiprepped samples to determine if the plasmids have retained the expected restriction sites. If the restriction sites are present, then full Sanger sequencing of the plasmid is performed.

The top gel shows restriction digests for samples maxiprepped using a 12 hour starter culture and incubation at 37°C. Incubation with KpnI and NotI enzymes alone should create a linearized plasmid, and should thus produce one 9030 bp band on the agarose gel. XmaI should produce two bands, approximately 4000-5000 bp in length. None of the restriction enzymes have created the expected

pattern on the agarose gel. Therefore, the sequence of the maxiprepped plasmids is incorrect.

The bottom gel depicts restriction digests for samples maxiprepped using an 8 hour starter culture and incubation at 30°C. Both KpnI and MfeI produced the expected linearized plasmid band that is 9030 bp long. Digests with XmaI produced the expected two bands, which are each approximately 4000-5000 bp in length.

KpnI and “K”=plasmid digested with KpnI

NotI and “N”= plasmid digested with NotI

MfeI and “M”= plasmid digested with MfeI

XmaI= plasmid digested with XmaI

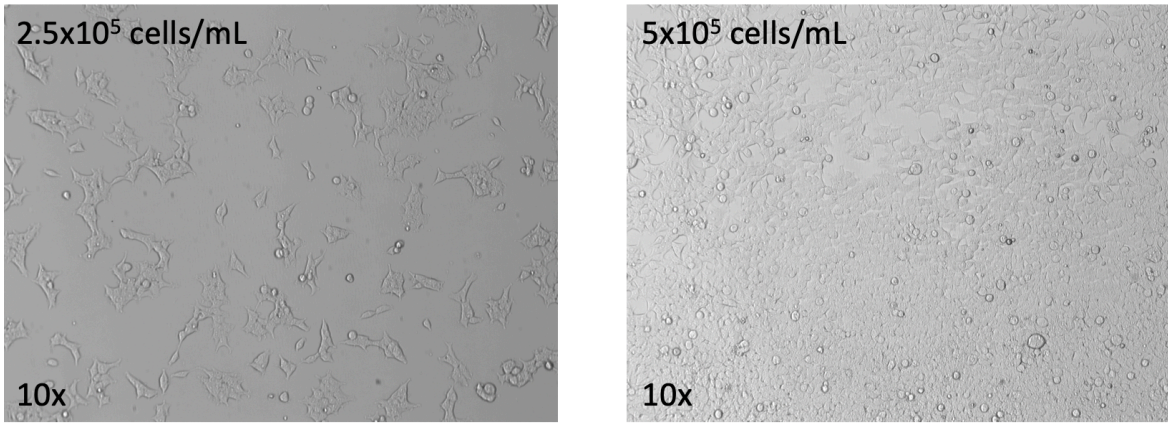

Figure 7.7 Optimizing Seeding Density of HEK293-*eGFP* Cells in Corning 24-well plates

Different seeding densities of HEK293-*eGFP* cells in 24-well plates were tested. The above images show brightfield images at 10x of HEK293-eGFP cells seeded at two different densities in 24-well plates, 24 hours after seeding. Each well received 0.5 mL of a cell suspension that contained either $2.5x10^5$ cells/mL or $5x10^5$ cells/mL. The cells should be approximately 70-80% confluent 24 hours after seeding for optimal plasmid transfection. The cells seeded a $5x10^5$ cells/mL appear to be ~80% confluent while the cells seeded at $2.5x10^5$ cells/mL are approximately 30% confluent. It was therefore determined that the cells seeded at a density of $5x10^5$ cells/mL were more suitable for plasmid transfection.

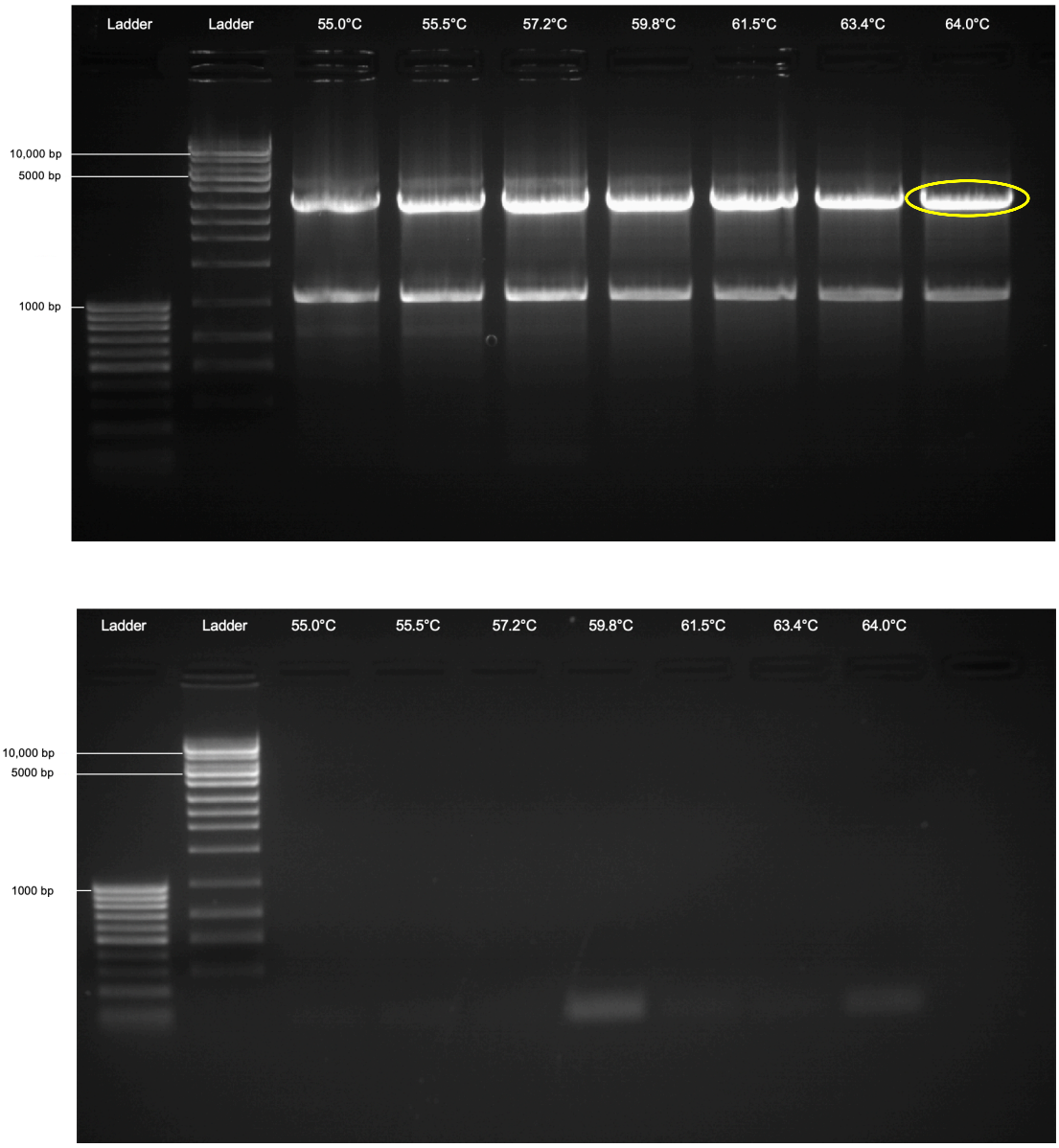

Figure 7.8. Optimizing Gradient PCR to Amplify the GeoCas9 region from Addgene Plasmid pET-MBP-NLS-Geo_st (#87703, Addgene)

The GeoCas9 region from pET-MBP-NLS-Geo_st plasmid (#87703, Addgene) was amplified using primers GeoCas9 KpnI FW6 and GeoCas9 HindIII RV5 at several different annealing temperatures, as indicated above the gel electrophoresis lanes. The PCR products were then placed on a 1% agarose for separation of the products at 100 V for ~1 hour. The top gel shows the amplified samples and the bottom gel shows the no DNA template controls.

All annealing temperatures allowed for amplification of a ~3000 bp band, which is the size of the GeoCas9. There was no contamination, as indicated by the lack of bright bands on the no template control gel. The band amplified at 64.0°C (denoted by the yellow oval) was excised and gel extracted for subsequent ligation.

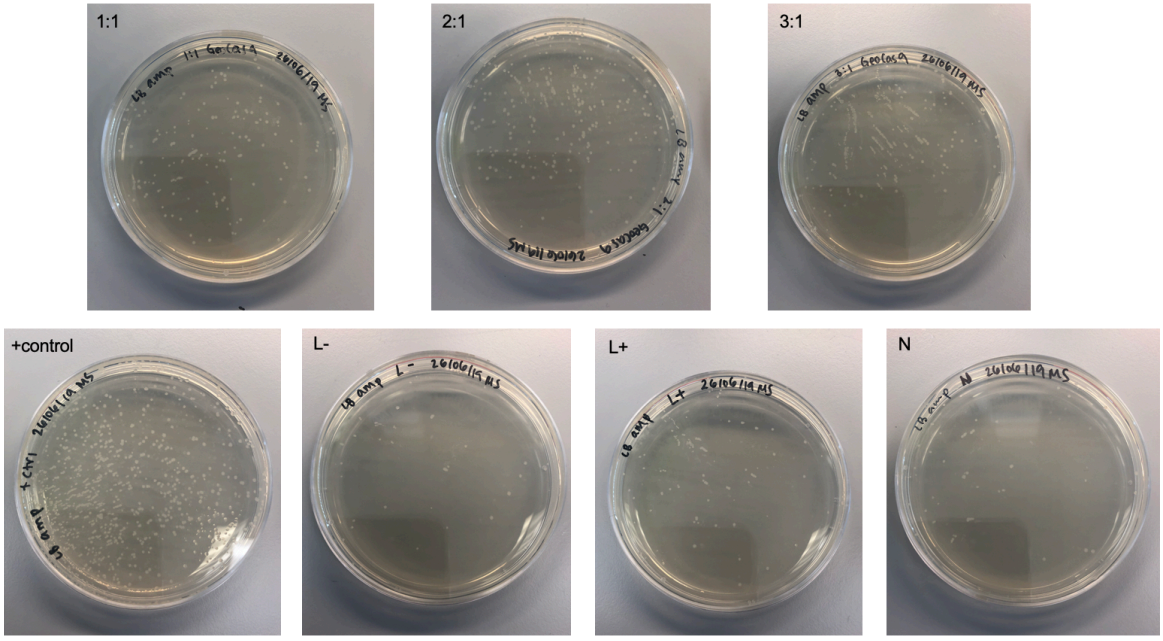

Figure 7.9 Optimizing Ligation of the Digested GeoCas9 Fragment into the PX404 *Campylobacter jejuni* Cas9 Plasmid Vector (#68338, Addgene)

The amplified GeoCas9 fragment and the PX404 *Campylobacter jejuni* Cas9 plasmid vector (#68338, Addgene) were both digested with KpnI and HindIII restriction enzymes. The digested fragments were then ligated with a total of 10 ng of PX404 *Campylobacter jejuni* Cas9 Plasmid Vector plasmid. Three different ratios of GeoCas9 fragment: PX404 *Campylobacter jejuni* Cas9 plasmid vector were used (1:1, 2:1, and 3:1). The ligated plasmids were then transformed. The results of the transformation are shown in the image above.

The presence of bacterial growth on the plate with the positive control (+control) pUC18 plasmid shows that the transformation reactions were successful. Colonies for miniprep were picked from the plate with the 2:1 ratio, which was the experimental plate with the greatest colony growth.

In this reaction, there is some re-circularized vector, as indicated by the colony growth on the plate with digested vector, ligase, and no GeoCas9 insert (N). There is also some undigested vector, as indicated by colony growth on the plate with digested vector without ligase (L-). The ligase step is functioning properly, as indicated by bacterial growth on the plate with linearized plasmid incubated with ligase (L+).

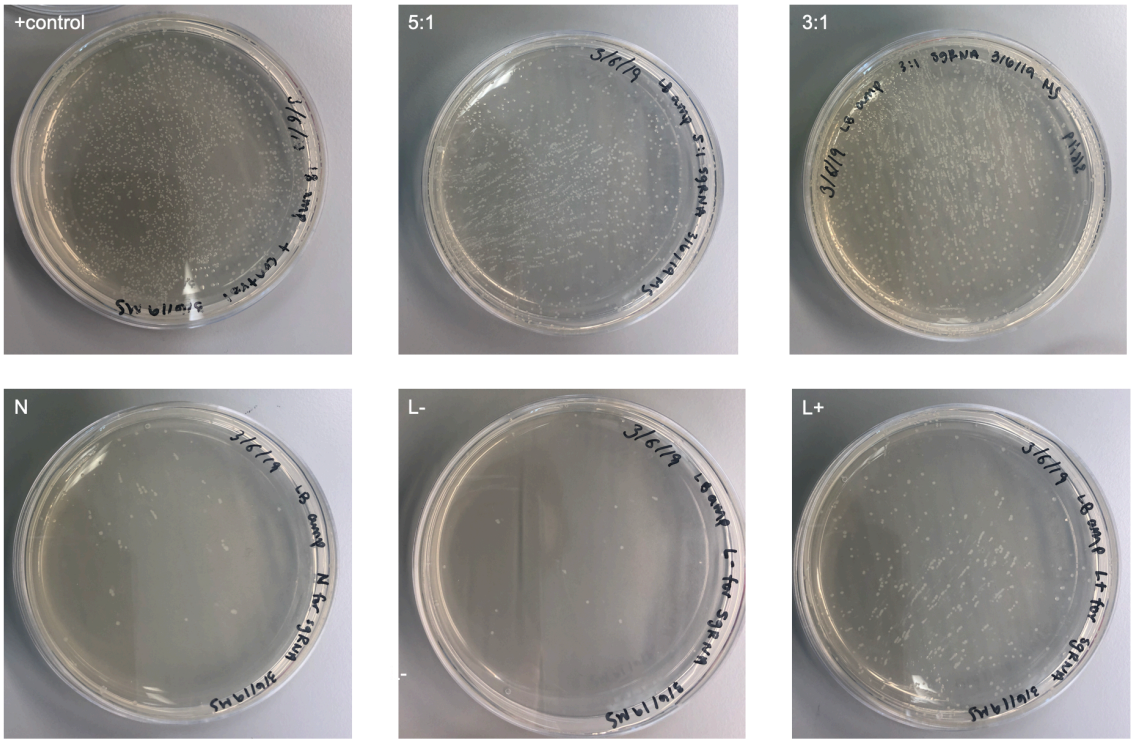

Figure 7.10 Optimizing Ligation of the Non-Targeting GeoCas9 SgRNA into the pU6-Cj-sgRNA Plasmid (#89753, Addgene) Vector

Oligos coding for the non-targeting GeoCas9 sgRNA were ordered, annealed, and phosphorylated. The oligos were then ligated with 30 ng of pU6-cj-sgRNA plasmid (#89753, Addgene) vector that has previously been digested with the SalI and NdeI restriction enzymes. Two different ratios of oligo:linearized plasmid vector were used (3:1 and 5:1). The ligated plasmids were then transformed. The results of the transformation are shown in the image above.

The presence of bacterial growth on the plate with the positive control (+control) pUC18 plasmid shows that the transformation reactions were successful. There are bacterial colonies on the plates with ligated plasmid in both the 3:1 and 5:1 ratio. It appears that these inserts have been successfully ligated into the backbone. In this reaction, there is very little re-circularized vector, as indicated by the minimal colony growth on the plate with digested vector, ligase, and no oligo insert (N). There is also very little undigested vector, as indicated by minimal colony growth on the plate with digested vector without ligase (L-). The ligase step is functioning properly, as indicated by bacterial growth on the plate with linearized plasmid incubated with ligase (L+).

Subsequently, crRNA oligos targeting either the eGFP gene in HEK293-eGFP cells, the rhodopsin gene, or the VEGFA gene were cloned into the above plasmids by restriction with the BsmbI restriction sites.

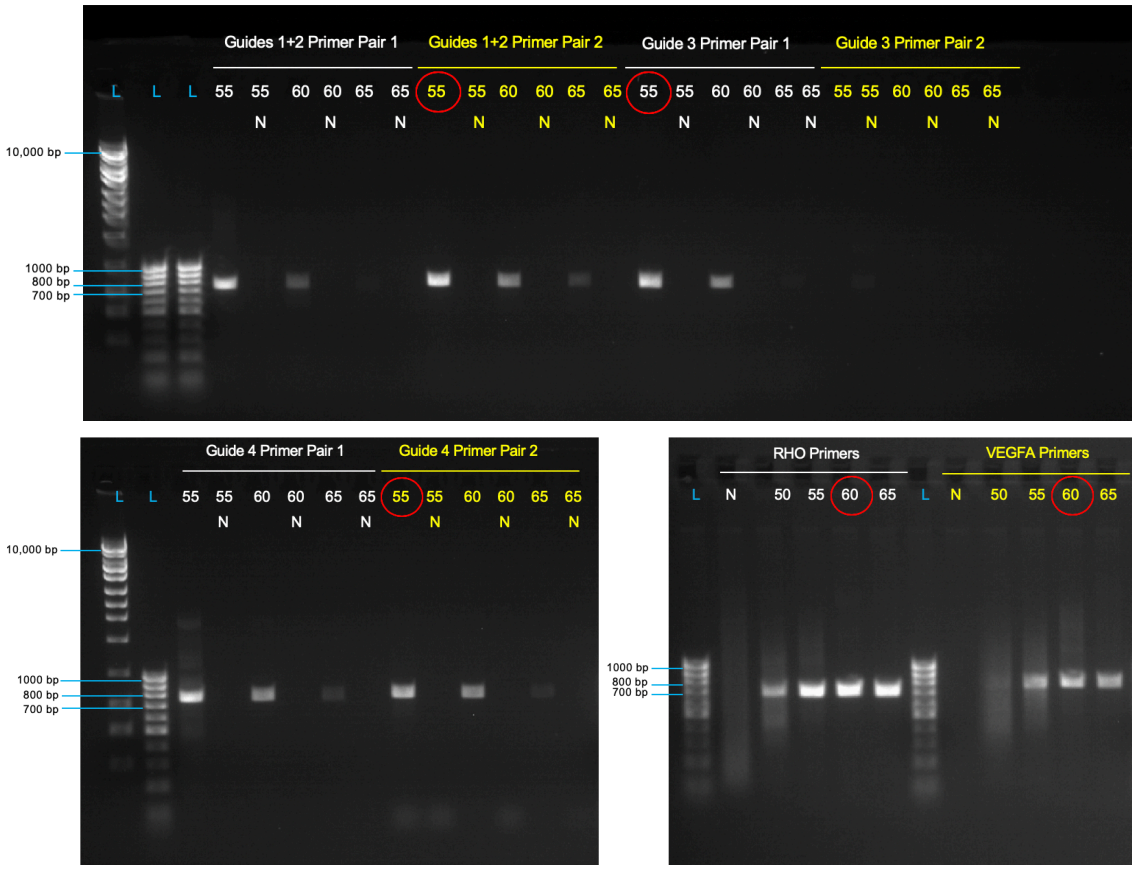

Figure 7.11 TIDE Primer Optimization

After CRISPR/Cas9-edited DNA is extracted, the target region must be PCR amplified by a pair of primers to produce a ~700 bp fragment. This amplified fragment is then sent for Sanger sequencing and the sequence is subsequently used for TIDE analysis.

The purpose of this experiment was to find the ideal annealing temperature and primer pair for each of the guide RNAs (N=no template control; L=ladder).

The top panel shows PCR of the regions targeted by G1+G2 and G3. For each target region, two primer pairs were ordered (1 and 2). DNA was amplified at three different temperatures (55°C, 60°C, 65°C).

The bottom left panel shows PCR of the region targeted by G4. For this guide RNA, two primer pairs were ordered (1 and 2). The DNA regions were amplified at three different temperatures (55°C, 60°C, 65°C).

The bottom right panel shows PCR of the regions targeted by RHO and VEGFA. DNA was amplified at four different temperatures (50°C, 55°C, 60°C, 65°C) and the no template control (N) was amplified at 50°C.

The chosen PCR conditions and primer pairs are indicated by the red circles.

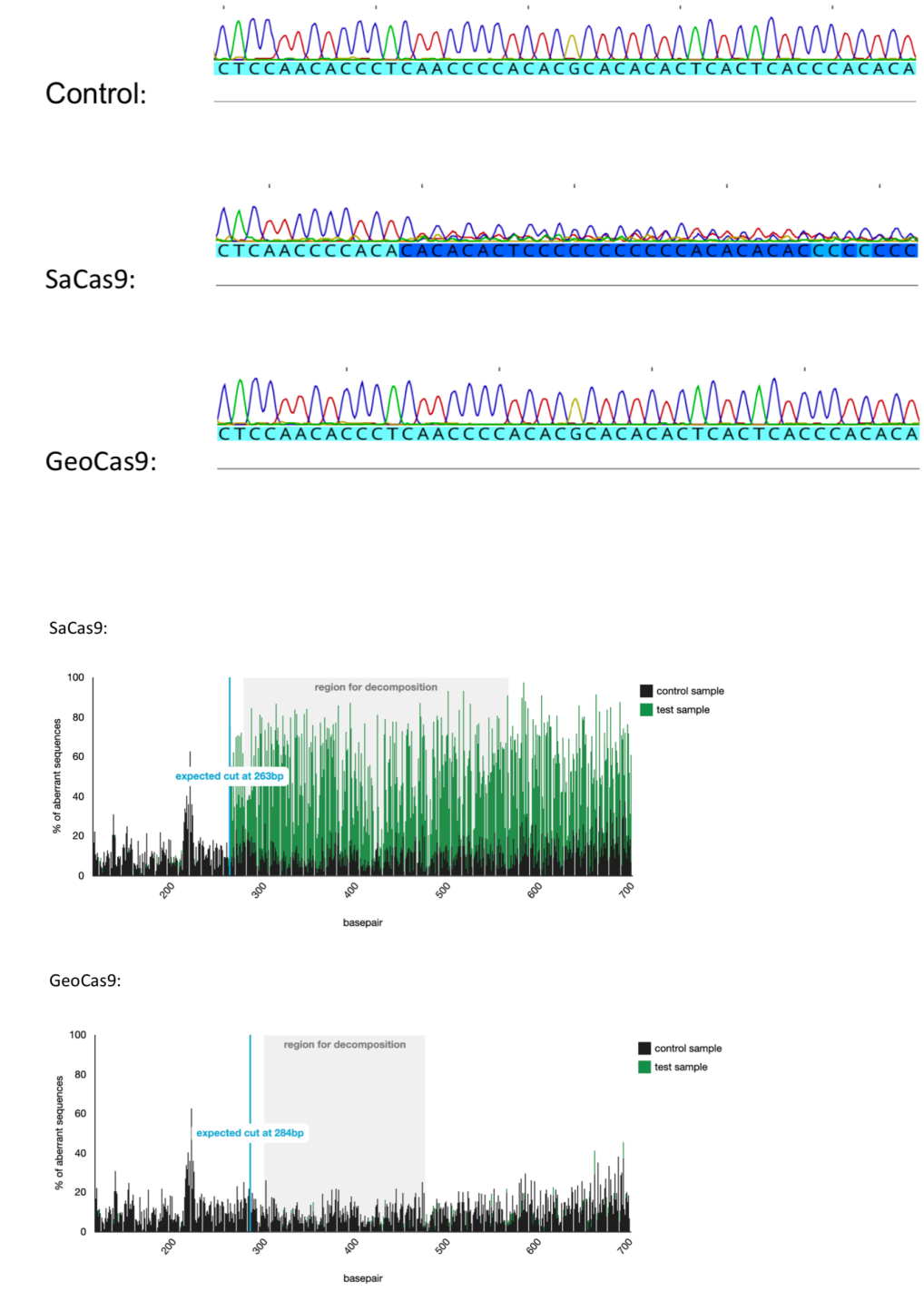

Figure 7.12 TIDE Analysis of SaCas9 and GeoCas9 Editing of VEGFA

The unedited control DNA sample and the CRISPR/Cas9 edited samples were sent for Sanger sequencing (top three sequences). The sequences were uploaded onto the TIDE analysis website (https://tide.deskgen.com/), which aligns the control sequence with the edited sequences in order to determine the editing rate of the Cas9 protein (bottom two panels). In this case, the editing rate of the GeoCas9 was 0.0% and the editing rate of the SaCas9 was 68.1%.

Table 4. List of CrRNA Sequences

Chapter	Plasmid Code	sgRNA Target	crRNA Sequence
3	RHO-F	Rhodopsin	TGCCATTACCTGGACCAGCCG
3	RHO-R	Rhodopsin	TGCCATTACCTGGACCAGCCG
3	VEGFA-F	*VEGFA*	GGGTGAGTGAGTGTGTGCGTG
3	VEGFA-R	*VEGFA*	GGGTGAGTGAGTGTGTGCGTG
3	NTC-F	Non-targeting control	GAAGCACTGCACGCCGTAGGT
3	NTC-R	Non-targeting control	GAAGCACTGCACGCCGTAGGT
3	13-F	*eGFP*	GCTGCCGTCCTCGATGTTGTG
3	13-R	*eGFP*	GCTGCCGTCCTCGATGTTGTG
3	15-F	*eGFP*	CTTGTACAGCTCGTCCATGCC
3	15-R	*eGFP*	CTTGTACAGCTCGTCCATGCC
4	G1	*eGFP*	CTGCTTCATGTGGTCGGGGTAG
4	G2	*eGFP*	CACGACTTCTTCAAGTCCGCCA
4	G3	*eGFP*	GATGCGGTTCACCAGGGTGTCG
4	G4	*eGFP*	GCCGAGAGTGATCCCGGCGGCG
4	RHO	Rhodopsin	GAACCATGCCATCATGGGCGTT
4	VEGFA	*VEGFA*	TGAGGACGTGTGTGTCTGTGTG
4	SaNTC	Non-targeting control	GGAGACCACGGCAGGTCTCA
4	SaVEGFA	*VEGFA*	GGGTGAGTGAGTGTGTGCGTG

Table 5. List of PCR Primers

Chapter	Primer Name	Primer Sequence
3	NotI_hU6_F2	ATGCGGCCGCCCGAGGGCCTATTTCCCATG
3	KpnI_hU6_R1	ATGGTACCCCGCAAAAATCTCGCCAACA
4	GeoCas9 KpnI FW6	GAATAGGTACCGCCACCATGCCCAAG
4	GeoCas9 HindIII RV5	GCAAGGAAGCTTTTAGTCACGAGTAGATTG

Table 6. List of Oligonucleotides

Chapter	Oligonucleotide Name	Oligonucleotide Sequence
4	AgeI-KpnI-HindIII-EcorI FW	CCGGTGGTACCGATCGATTAAAGCTTG
4	AgeI-KpnI-HindIII-EcorI RV	AATTCAAGCTTTAATCGATCGGTACCA
4	6x His tag FW	CGCCACCATGCATCATCACCATCACCACGGTAC
4	6x His tag RV	CGTGGTGATGGTGATGATGCATGGTGGCGGTAC
4	Geocas9 sgrna scaffold (ordered as a double stranded sequence)	GACTAGGTCGACTAGTCAATAATCAATGT CAAAAAAAAAGGATGGGGAATGCCCGC CAAAAGCGGGCGACAGGCGATCCCCAAC GCCGCGGGTCAGTCTGCCTTAGGCAGAA AGCCCTTATCATAGTAACCCTGATTTCTC AGGGGAACTATGACGGAGACGGGTCAT CCCGTCTCCCCCGGTGTTTCGTCCTTTC CACAAGATATATAAAGCCAAGAAATCGA AATACTTTCAAGTTACGGTAAGCATATGCGTAGT

Table 7. Primers for CDNA Synthesis for QPCR

Chapter	Primer Name	Primer Sequence
3	sgRNA_R1	ATCTCGCCAACAAGTTGACG
4	GeosgRNA2	AGGATGGGGAATGC

Table 8. Primers for TIDE Analysis

Chapter	Primer Name	Primer Sequence	Plasmid Associated with Primer
4	GeoCas9 TIDE Guide 1 FW 2	GATAGCGGTTTGACTCACG	G1 and G2
4	GeoCas9 TIDE Guide 1 RV 2	GCTTGTCGGCGGTGATATAG	G1 and G2
4	GeoCas9 TIDE Guide 3 FW 1	CGGTGGGAGGTCTATATAAGC	G3
4	GeoCas9 TIDE Guide 3 RV 1	GTGCTCAGGTAGTGGTTG	G3
4	GeoCas9 TIDE Guide 4 FW 2	GTCTATATCACCGCCGACAAG	G4
4	GeoCas9 TIDE Guide 4 RV 2	CAACACCACGGAATTGTC	G4
4	GeoCas9 RHO primer FW 1	AGGCCTCCTCAAATCCCTCTC	RHO
4	hRHO+2699R	GAGTGGGACCCAGTTCCAAG	RHO
4	VEGFA_F1	GGCTCCAACAGGTCCTCTTC	VEGFA
4	VEGFA_R1	CACCAAGGTTCACAGCCTGA	VEGFA

Table 9. Custom TaqMan Probe Sequences

Chapter	Primer Name	Primer Sequence
3	SA.GRNA_SCAFF_F	GTTTTAGTACTCTGGAAACAGAATCTACT
3	SA.GRNA_SCAFF_M	AAACAAGGCAAAATGC
3	SA.GRNA_SCAFF_R	GCCAACAAGTTGACGAGATAAACAC
3	SACAS9_F	GCAAAGAGAACGCCAAGTACCT
3	SACAS9_M	ATCGAGAAGATCAAGCTGC
3	SACAS9_R	CACTTGCCTTCCTGCATGTC
4	GEO.GRNA_SCAFF_F	CTATGATAAGGGCTTTCTGCCTAAGG
4	GEO.GRNA_SCAFF_M	ATCGCCTGTCGCCCGC
4	GEO.GRNA_SCAFF_R	GGGAATGCCCGCCAAAA
4	GEOCAS9_F	CTGAACCCTACTGGTCACGATATTG
4	GEOCAS9_M	CACAGCTTGAACTTCA
4	GEOCAS9_R	CAGCGACCGTTCTGTTCAGA