Engineering Virus-Like Particles for the Delivery of Genome Editing Enzymes

by

Beth Rousseau

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in the University of Michigan
2022

Doctoral Committee:

Associate Professor David L Turner, Chair
Professor Emeritus Roberta Fuller
Professor Daniel Goldman
Associate Professor Kenneth Kwan
Professor Michael Uhler

Beth Rousseau

bapoppe@umich.edu

ORCID iD: **0000-0002-6751-6605**

Dedication

This thesis is dedicated to everyone who refused to self-censor during the COVID-19 pandemic. There are too many free-thinkers to name individually, but I do want to give a special acknowledgement to the authors of the Great Barrington Declaration, Doctors Jay Bhattacharya, Martin Kulldorff, and Sunetra Gupta. They contributed to the discussion on how public policy should be crafted with respect to COVID-19 at a time when questioning the policies in place was met with censorship and great risk to one's career. Many of the public health policies adopted during the COVID-19 pandemic violated our civil liberties. They were an unacceptably authoritarian, illiberal response to the presence of a pathogen. In order to protect the rights of the people, further restrictions must be placed on the powers of state and federal governments.

Acknowledgements

Dr. Huanqing Zhang helped with the creation of the mouse dermal fibroblast line, the collection of embryonic murine retinas, with cloning the KRAS guide RNA plasmid, and with ordering many of the reagents for this work.

The Uhler, Kwan, and Goldman labs were all very kind in lending a few of the reagents needed to complete this work.

The Freddolino and Baldridge labs kindly allowed us the use of the MACSquant Vyb flow cytometer.

Table of Contents

Dedication ii

Acknowledgements iii

List of Tables vii

List of Figures viii

Abstract x

Chapter 1 Introduction 1

1.1 Adaptive Immunity in Microbes 2

1.2 RNA-targeting with Cas9 5

1.3 Genome Editing with Cas9 6

1.4 Prime Editing 10

1.5 Mechanisms of Cas9 Delivery into Cells 12

1.5.1 Virus-Like Particles for the Transduction of Cas9 Ribonucleoproteins 17

1.6 The Murine Leukemia Virus 17

1.6.1 Murine Leukemia Virus Lifecycle 19

1.6.2 Examples of Protein Transduction by MLV VLPs 21

1.7 Properties of MLV Gag that Govern Particle Formation 25

1.8 Unresolved Questions Concerning Protein Transduction with MLV VLPs 29

Chapter 2 Programmable RNA Cleavage and Recognition by a Natural CRISPR-Cas9 System from Neisseria Meningitidis 31

2.1 Introduction 31

2.2 A natural RNA-guided ribonuclease activity of NmeCas9 *in vitro* 34

2.3 RNA-guided RNA cleavage by NmeCas9 is programmable 38

2.4 Type II-C anti-CRISPRs inhibit *in vitro* RNA cleavage by NmeCas9 42

2.5 RNA-guided, sequence-specific binding of NmeCas9 to RNA targets *in vitro* 43

2.6 Expanded targeting capacities for CRISPR-Cas9 systems 47

2.7 Potential RNA-targeting applications 48

2.8 Biological Implications of NmeCas9 RNA targeting 48

Chapter 3 Development of Virus-Like Particles Derived from the Moloney Murine Leukemia Virus for the Delivery of Genome Editing Enzymes 50

3.1 Introduction 50

3.2 Results 53

3.2.1 Generation of VLPs for Cas9 Delivery 53

3.2.2 Gag Codon Optimization Improves Cas9 VLP Titer 57

3.2.3 Integrase and Reverse Transcriptase Are Not Necessary for Cas9 VLP Function 62

3.2.4 GagCO-PE Can Edit a Target Gene in Transiently Transfected HEK293 Cells 70

3.2.5 MLV VLPs Deliver Functional Prime Editing Enzymes 73

3.3 Discussion 81

3.4 Methods 89

3.5 Nucleic Acid Sequences 97

Chapter 4 Dual Purpose Virus-Like Particles Derived from the Moloney Murine Leukemia Virus 99

4.1 Introduction 99

4.2 Results 100

4.2.1 Dual Guide RNA Delivery with Cas9 VLPs 100

4.2.2 Cas9 VLP Efficacy Is Not Improved by the Addition of the ET-208 Nuclear Export Sequence to the Guide RNA 108

4.2.3 Cre-Cas9 VLPs Can Deliver Two Enzymes Simultaneously 112

4.2.4 Mutating the Cleavage Site Between Gag and Cas9 Reduces Cas9 VLP Titer 115
4.2.5 Comparison of Three Different Protease Cleavage Sites in GagCO-Cas9 118
4.2.6 The Effect of Altering the Stoichiometry in Cas9 VLPs Made with Gag-Stop-Prot . 121
4.2.7 MMLV VLPs Deliver Cre recombinase Efficiently 124
4.2.8 MMLV VLP Delivery into Embryonic Murine Retinas In Vitro..... 125
4.3 Discussion: 129
4.3.1 Challenges and Potential Solutions to Multiplexed Cas9 VLPs 129
4.3.2 Challenges and Potential Solutions to Delivering Multiple Functional Enzymes 130
4.4 Methods 131
Chapter 5 Conclusions and Future Directions 135
5.1 Implications of RNA-targeting by NmeCas9 136
5.2 Implications of Efficient Delivery of Cas9 by Virus-Like Particles 138
5.3 Considerations in the Design of Gag-Fusion Proteins 139
Bibliography 146

List of Tables

Table 1 Summary of MLV VLP Publications 25

Table 2 The sequences of guide RNAs and prime editing guide RNAs..... 97

Table 3 The sequences of oligonucleotides used to amplify genomic DNA for TIDE sequencing. 98

List of Figures

Figure 1 CRISPR-Cas9 Adaptive Immune System of Streptococcus pyogenes against bacteriophages..... 3

Figure 2 Cas9 catalyzes a double-stranded break in DNA. 7

Figure 3 The Murine Leukemia Virus genome and virion. Created with BioRender.com. 18

Figure 4 The simplified lifecycle of the Murine Leukemia Virus. Created with BioRender.com. 20

Figure 5 Particle assembly and budding are driven by different parts of Gag. 26

Figure 6 NmeCas9 possesses a natural RNA-guided ribonuclease activity in vitro..... 37

Figure 7 NmeCas9-catalyzed RNA cleavage is programmable. 41

Figure 8 Type II-C anti-CRISPRs block RNA cleavage by NmeCas9. 43

Figure 9 dNmeCas9 binds ssRNA target in vitro, in a crRNA-guided manner..... 46

Figure 10 Cas9 VLPs activate a reporter in the Ai9 MDF cell line..... 57

Figure 11 Codon optimization significantly improves Cas9 VLP titer. 58

Figure 12 Cas9 VLPs are efficient..... 61

Figure 13 Testing Protease Only Cas9 VLPs. 64

Figure 14 VLPs can deliver functional Cas9 with mutation of RT and deletion of IN. 66

Figure 15 Further characterization of GagPolRT$^-$IN$^-$ VLPs. 69

Figure 16 GagCO-PE functions when directly transfected into HEK293T cells. 72

Figure 17 Prime editing VLPs introduce the HA epitope tag into lamin in HEK293T cells. 75

Figure 18 Western blots show GagPE production in HEK293T cells and processing by the viral protease in PE VLP supernatants..... 79

Figure 19 PE VLPs introduce the HA tag into β-actin. 80

Figure 20 Cas9 VLPs can target two genomic loci simultaneously. 103

Figure 21 Dual delivery of Rpl22 and Wtap sgRNA is less efficient than the independent delivery of either guide. 106

Figure 22 Addition of the ET208 aptamer nuclear export sequence to the sgRNA does not alter the efficacy of Cas9 VLPs. 111

Figure 23 Gag-Cre-Cas9 VLPs can deliver both enzymes to MDF cells. 114

Figure 24 Mutating the cleavage site between Gag and Cas9 reduces Cas9 VLP titer. 117

Figure 25 Using different viral protease sites between Gag and Cas9 alters the titer of Cas9 VLPs. 120

Figure 26 Adjusting the stoichiometry of VLP components did not improve Cas9 VLP titer... 123

Figure 27 VLPs efficiently deliver Cre into Ai9 MDF cells. 124

Figure 28 Cre VLPs can activate the tdTomato reporter in mouse retinal cells. 128

Abstract

Genome editing with Cas9 is a powerful method of investigating the roles of genes in biology. *S. pyogenes* Cas9 catalyzes a precise, blunt, double-stranded break in DNA when directed toward a genomic locus complementary to a programmable guide RNA and adjacent to the sequence 5'-NGG-3'. However, prolonged expression of the Cas9 ribonucleoprotein (RNP) in cells can increase off-target cleavage events. Transient Cas9 RNP transduction can mitigate the risk of off-target events by reducing both the length of time that cells are exposed to Cas9 RNPs and controlling the dose of RNP administered. Common methods of delivering Cas9 RNPs include electroporation, lipid-based transfections, nanoparticles, and virus-like particles (VLPs). VLPs are particles similar in form and function to a virus but lacking a viral genome. We chose to construct VLPs for Cas9 delivery because they require no specialized equipment to transduce cells, are inexpensive, can be scaled up easily, have relatively low toxicity, and can be pseudotyped with different viral envelope proteins which can enable delivery to a wide range of cells. VLPs derived from the Murine Leukemia Virus (MLV) are known to serve as a vehicle for the efficient transduction of proteins, including Cas9 and other genome editing enzymes. Gag and GagPol are the two polyproteins that comprise the interior of an MLV virion. Gag expressed independently of other viral transcripts and proteins spontaneously forms a VLP. A protein fused to the C-terminus of Gag is loaded into MLV VLPs concomitantly with their creation. In an effort to improve VLP efficacy, we have made novel Moloney MLV VLPs with codon optimized Gag. Codon optimization improves VLP titer by a factor of ten. Another optimization that we have made to VLPs is the elimination of functional reverse transcriptase and integrase domains from GagPol. Although these domains are important for the lifecycle of the virus, we have found them to be inessential for Cas9 delivery by VLP. We have also found that VLPs can deliver two

enzymatically active cargo proteins at once, Cre and Cas9. Finally, we show that delivery of prime editing enzymes can be achieved with VLPs.

Chapter 1 Introduction

The functions of myriad genes have been painstakingly characterized over decades often through cloning, protein purification, and finally in vitro assays to assess the binding and/or enzymatic functions of proteins. Direct gene manipulation with zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and transposases made it easier to elucidate the functions of genes and their products (Cai, Bak, and Mikkelsen 2014; Bogdanove and Voytas 2011). However, ZFNs and TALENs can be difficult to reprogram to specific genomic loci, and transposases do not always disrupt in a specific genomic locus. Fortunately, the process for elucidating gene function has been greatly simplified with the advent of CRISPR genome editing technology derived from the adaptive immune systems of microbes. Due to the ease of reprogramming CRISPR gene editors, it is straightforward to target most genomic loci for differential regulation, knock-out, base editing, and even epitope tagging. This technology is now used in laboratories around the world for research purposes. Even more incredibly, it is already in the clinic as a therapeutic to help combat cancer and other debilitating diseases (Rees et al. 2021). Despite these advances, implementing CRISPR technology for genome editing can still be challenging, often due to the difficulty of delivering the molecular genome editing tools into cells where they can be effective. The overarching goal of this thesis has been to engineer a vehicle to efficiently deliver these genome editing tools so that they may be easier to use in the lab and perhaps even in the clinic.

1.1 Adaptive Immunity in Microbes

Like so many wonderful discoveries in science, the first observation of a bacterial immune system was discovered incidentally, and its implications were not fully understood until years later. In 1987, Ishino et. al. published their findings on a protein coding gene from *E. coli* called *iap* (Ishino et al. 1987). A genomic fragment known to contain the *iap* gene was sequenced end to end. At the 3' end of the gene was a series of five repeats of 29 nucleotides with 32 nucleotides of unique sequence separating the repeats. They hypothesized that perhaps the sequences were for the purpose of mRNA stabilization. Today we know these peculiar repeats as CRISPRs, Clustered Regularly Interspaced Short Palindromic Repeats which are essential to the adaptive immune system of microbes (Bhaya, Davison, and Barrangou 2011). CRISPR systems are composed of CRISPR arrays and adjacent encoded CRISPR-associated proteins. There are six types of CRISPR systems classified differently depending on their suites of CRISPR-associated proteins (Cas1, Cas2, etc.) and RNAs (Koonin, Makarova, and Zhang 2017). The type II system is the one found in *S. pyogenes* and *N. meningitidis* (Mir, Edraki, et al. 2018). All subsequent descriptions of CRISPR systems will be in regard to type II as it is of primary relevance to this body of work.

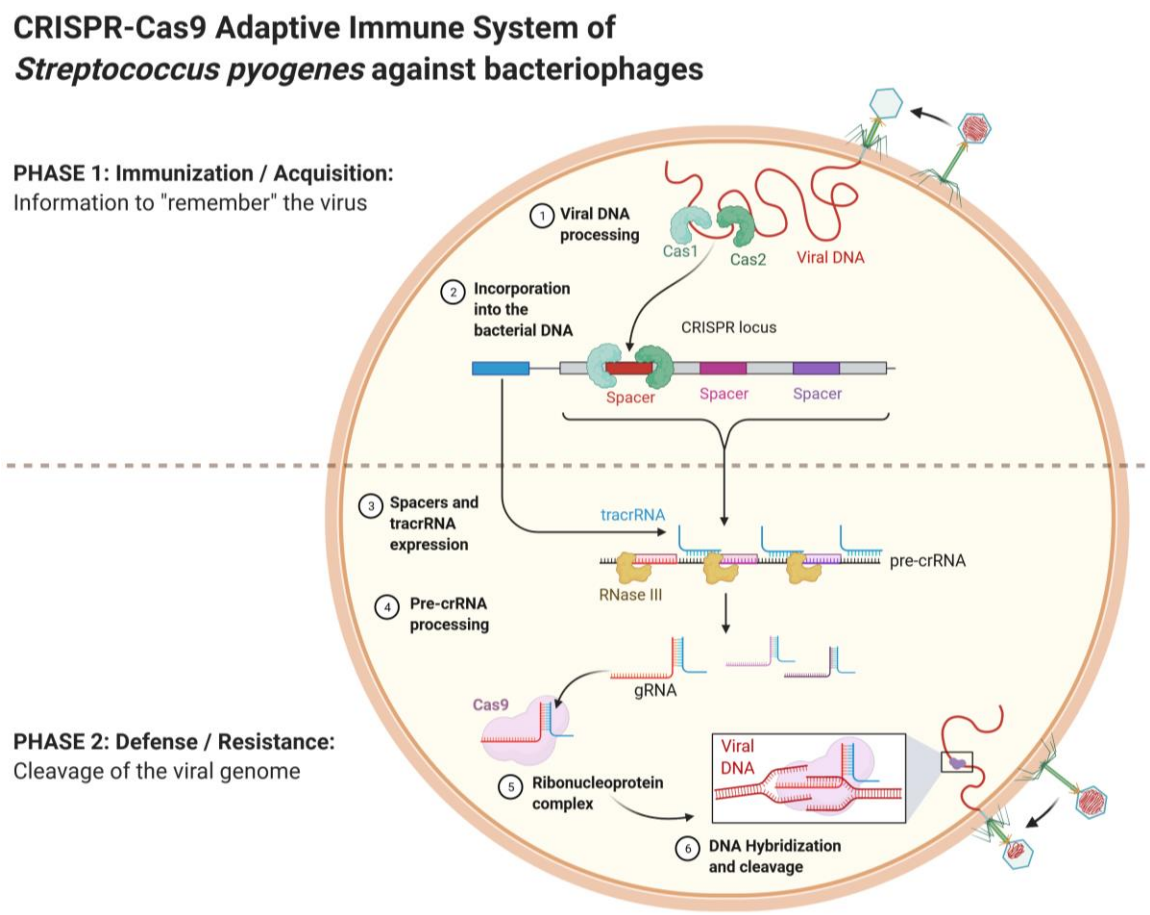

Figure 1 CRISPR-Cas9 Adaptive Immune System of Streptococcus pyogenes against bacteriophages. The adaptive immune response of prokaryotes has two phases. In the first, a sequence of bacteriophage DNA is incorporated into the CRISPR array. In the second, the endonuclease Cas9 is guided by an RNA derived from the sequence of the bacteriophage to the invader's DNA where it stops the infection through catalyzing a double-stranded DNA break. Created with BioRender.com.

When bacteria are attacked by bacteriophages, there are no immune cells to rush to their rescue. They must provide their own adaptive immunity (Fig. 1). The Cas1 and Cas2 enzymes are first on the scene (Koonin, Makarova, and Zhang 2017; Bhaya, Davison, and Barrangou 2011; Chylinski et al. 2014). They excise a small piece of the invading bacteriophage's double-stranded DNA (dsDNA) and insert the DNA as a new unique sequence, called a spacer, into the CRISPR array. Next, an RNA polymerase transcribes tracrRNA and pre-CRISPR RNA (crRNA). The tracrRNA is complementary to the repeat portion of the crRNA and forms a duplex with the crRNA. RNase III separates the long polymer of pre-crRNA into individual units

of tracrRNA and crRNA hybrids. The tracrRNA/crRNA hybrids are loaded onto Cas9 via specific interactions between the Cas9 protein and the tracrRNA/crRNA. The Cas9 protein/RNA complex will scan dsDNA for a Protospacer Adjacent Motif (PAM) which is a short sequence of DNA that is recognized by the PAM-interacting domain of the Cas9 enzyme (Chylinski et al. 2014, 9; Mir, Edraki, et al. 2018, 9). Although the crRNA originates from a DNA sequence in the bacterial genome, the placement of the crRNA into a CRISPR array ensures that the PAM sequence will not be adjacent to the spacer sequence which prevents self-targeting. The identity of the PAM sequence is specific to the origin of the Cas9. For example, *S. pyogenes* Cas9 recognizes the PAM 5'-NGG-3' whereas *C. jejuni* Cas9 recognizes 5'-NNNNACA-3'. Upon binding a PAM sequence in the bacteriophage DNA, Cas9 locally melts the double-stranded DNA which allows strand invasion by the crRNA (Mir, Edraki, et al. 2018; Chylinski et al. 2014; Sternberg et al. 2014). If the crRNA is complementary to the DNA, then Cas9 catalyzes a double-stranded break in the bacteriophage DNA which abrogates infection by the bacteriophage.

Although originally part of an adaptive immune system, Cas9's potential as a programmable genome editor has been exploited for research and therapeutic purposes (Jinek et al. 2013; Rees et al. 2021). The Cas9 protein is an RNA-guided, sequence specific nuclease. The ability of Cas9 to cut DNA at a site defined primarily by its RNA guide has led to numerous new methods for genome editing, both in mammalian cells and in other systems. Most of the genome editing experiments presented in this thesis have been conducted in mammalian cells with the well-characterized and commonly used SpyCas9 (Rees et al. 2021). This is important to note because Cas9 varies in both its physical and enzymatic properties between different species of microbe. It can vary in length from the compact 984 amino acid (aa) *C. jejuni* Cas9, to *N.*

meningitidis Cas9 at 1082 aa, to the rather large SpyCas9 at 1368 aa (E. Kim et al. 2017; Jinek et al. 2012; Hou et al. 2013). Size is important in selecting a Cas9 enzyme for genome editing. Smaller is generally considered better for cellular delivery (Schmidt et al. 2021). Another consideration is cleavage efficiency and the rate of off-target cleavage (Vakulskas and Behlke 2019). Finally, a key feature of Cas9 that can be limiting in its application as a genome editor is the requirement of a PAM sequence (Nishimasu et al. 2018). The exact PAM sequence is specific to the Cas9 variant. For example, SpyCas9 prefers to bind to the sequence 5'-NGG-3'. After PAM binding, then the guide RNA can attempt to anneal to the sequence on the strand opposite the strand containing the PAM sequence. All of the above are considerations when selecting a Cas9 protein for genome editing. We chose ultimately to use SpyCas9 for the studies here, despite its large size, due to the large number of targets already known to be efficiently edited by SpyCas9, as well as the ease of designing guides for additional uncharacterized genomic targets.

1.2 RNA-targeting with Cas9

An RNA-guided enzyme that cleaves both DNA and RNA could be very useful for both turning off gene expression and also quickly clearing RNA transcripts that may be very stable and linger long after the gene has been disrupted, thus preventing the effects of gene disruption from presenting in a timely fashion. Although initially defined as an RNA-guided DNA endonuclease, SpyCas9 has also been shown to bind single-stranded RNA in vitro and in vivo (O'Connell et al. 2014; Nelles et al. 2016; Batra et al. 2017). Moreover, it can be tricked into cleaving an RNA target with the addition of a PAMmer which is a DNA oligomer that binds to the target RNA adjacent to the sequence complementary to the guide RNA so that it presents a PAM sequence. Cas9's PAM-interacting domain engages with the PAM sequence of the

PAMmer and induces cleavage of the ssRNA target. Of interest was whether there were any Cas9 variants that could target and cleave ssRNA without the need for an exogenous PAMmer supplied in trans. In Chapter 2, we show that *N. meningitidis* (Nme)Cas9 can cleave ssRNA targets site specifically in vitro without the use of a PAMmer (Rousseau et al. 2018). Other labs have shown that *S. aureus* and *C. jejuni* Cas9s (Strutt et al. 2018; Dugar et al. 2018) can also target and cleave ssRNA transcripts without the need for a PAMmer. Additionally, we demonstrated that ssRNA cleavage can be blocked by anti-CRISPR proteins that are also known to block scission of DNA by NmeCas9. These naturally occurring enzymes are potentially new tools for answering tough questions in biology.

1.3 Genome Editing with Cas9

For Cas9 to recognize and cleave at a specific genomic site, it must be bound to a guide RNA. In the native host, Cas9 binds hybrids of tracrRNA and crRNA. The duplex RNA acts as a handle for Cas9 to hold onto while the single-stranded RNA sequence in the crRNA is what Cas9 attempts to hybridize to DNA after binding a PAM sequence. For simplicity in genome editing, additional nucleotides were introduced to link tracrRNA and crRNA into a chimera called a single-guide RNA (sgRNA) (Jinek et al. 2012). This sgRNA can be produced as a single transcript in vitro or intracellularly with the use of an appropriate promoter. The sgRNA sequence is recognized as a contributing factor to the cleavage efficiency and specificity of Cas9 (Briner et al. 2014; J.-P. Zhang et al. 2016). SpyCas9 normally binds to a sgRNA of approximately 100 nucleotides (nt) in length, of which 20 nt at the 5’ end determine the target specificity to target DNA. Perfect complementarity to the DNA is best for inducing cleavage. There are several programs that will allow for the quick and reasonably accurate prediction of a Cas9 sgRNA sequence that will efficiently target a desired locus (Sanson et al. 2018; Doench et

al. 2016). Although not used herein, many chemical modifications of synthetic sgRNAs have been made in an effort to enhance their stability, protect against degradation, reduce off-target cleavage, and even minimize immune responses (Q. Chen, Zhang, and Yin 2021).

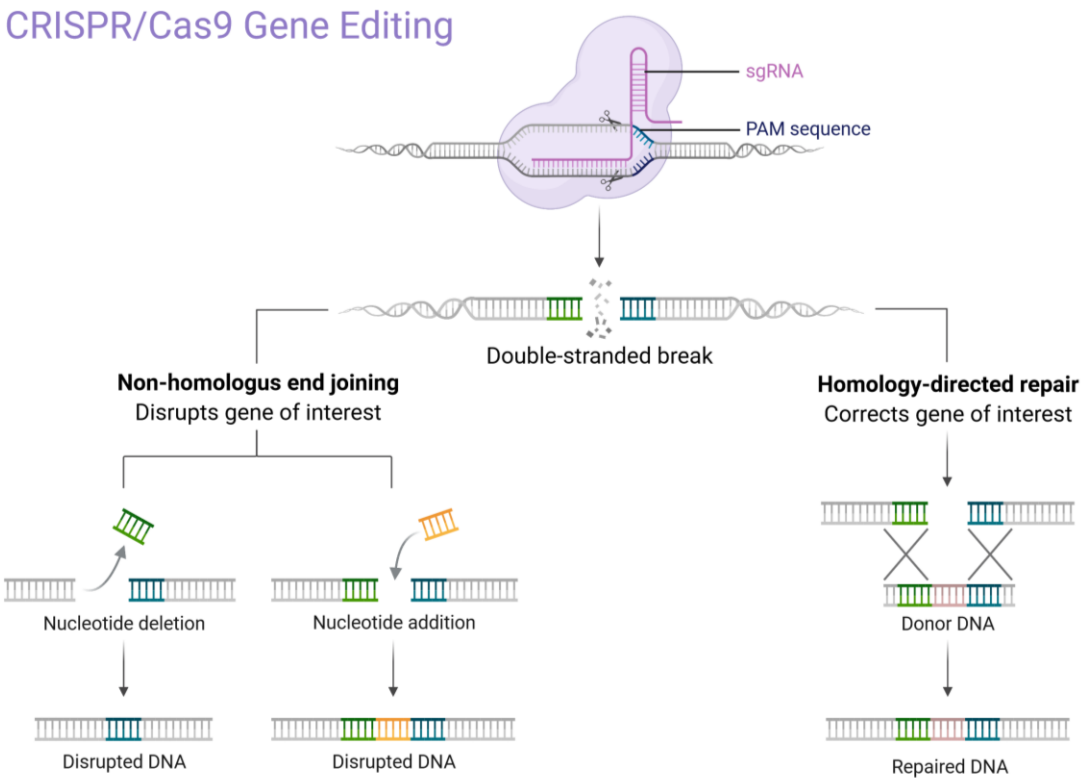

Figure 2 Cas9 catalyzes a double-stranded break in DNA.
Endogenous cellular machinery repairs the break by either non-homologous end joining or homology-directed repair. Created with BioRender.com.

When Cas9 creates a double-stranded break in the genomic DNA of a cell (Fig. 2), the repair of the break depends on endogenous cellular machinery. One cellular mechanism for repair of a double-stranded DNA break is non-homologous end-joining (NHEJ) (Chang et al. 2017). A set of DNA-binding, kinase, nuclease, and ligase proteins assemble at the site of a double-stranded break where they repair the break by joining the two ends together regardless of whether the repaired DNA is identical to the original sequence. This often results in insertions or

deletions (indels). In a protein coding gene, indels within the coding region can cause frameshifts that result in the effective knockout of the gene. It is also possible to create a deletion in a target with Cas9 by creating a pair of double-stranded breaks within a few thousand bases, which can then be rejoined by NHEJ (Canver et al. 2014). Another repair mechanism for double-stranded breaks is homology-directed repair (HDR), in which a separate homologous DNA template is used as a donor repair template (X. Li and Heyer 2008). In HDR, the broken DNA is separated into single strands, coated by replication protein A (RPA), and then guided by Rad51 as it attempts to find homology through strand invasion. The donor template can be a sister chromosome, or it can be a synthetic donor DNA molecule that is supplied in trans. Donor DNA that is designed for HDR mediated repair will have at least one homology arm that matches the genomic sequence flanking a desired repair or edit site. The HDR donor can introduce a modified or edited sequence at the repair site, such as a deletion, insertion, or modification/recoding of the targeted genomic locus. Due to these different repair pathways, their molecular requirements, and their dependence on the cellular repair enzymes, the outcomes of DNA editing with Cas9 are not always identical, nor homogenous within an edited population of cells. Biasing the repair of Cas9 induced breaks to favor certain outcomes has been an active area of research (H. Yang et al. 2020).

As alluded to above, Cas9 can cleave at sites that are not perfect matches to the guide RNA sequence and/or do not have a perfect PAM. These off-target cleavage events can lead to repair at the site by NHEJ which can be detected, even if occurring only at a low frequency, by next-generation sequencing methods such as genome-wide unbiased identification of DSBS enabled by sequencing (GUIDE-seq) (Tsai et al. 2015). This method takes advantage of the fact that a short, blunt programmed dsDNA can be incorporated with reasonable efficiency into a

double-stranded break when it is repaired by NHEJ. The programmed dsDNA acts as an identifier of locations that were cleaved by Cas9 after the DNA of an edited population of cells has been sheared, ligated with adapters, and deep sequenced. By GUIDE-seq and other methods, it has been shown that Cas9 has off-target events at rates high enough to be a cause for concern if Cas9 were employed as a therapeutic in vivo. Efforts to reduce the off-target events include modifying the guide RNA, using Cas9 variants with higher fidelity (Kleinstiver et al. 2016; Slaymaker et al. 2016; Vakulskas et al. 2018; J. S. Chen et al. 2018), or limiting the dose or duration of exposure to Cas9 (Senturk et al. 2017). Delivery of Cas9 by RNP is an example of having tighter control over the amount of Cas9 delivered to cells and reducing the length of exposure. Cas9 RNPs do not persist for a prolonged period, whereas Cas9 synthesized from an introduced viral or nucleic acid vector will usually persist for longer. RNP delivery has proven to be successful in reducing or mitigating undesired off target cleavage events (Vakulskas and Behlke 2019).

The dsDNA cleavage ability of Cas9 has been widely exploited, but Cas9 is also useful simply for its ability to bind at a specific genomic sequence when programmed with an appropriate sgRNA. The Cas9 RNP can bind to a genomic locus even if the HNH or RuvC nuclease domains have been inactivated by point mutations. Cas9 with one nuclease domain mutated is called a nickase (nCas9), while disruption of both nuclease domains yields a “dead” Cas9 (dCas9). When dCas9 is directed to a specific genomic locus by an sgRNA, the binding is sufficiently stable that other proteins fused to dCas9 can have observable effects on the genomic locus. For example, dCas9 fused to a protein that functions as a transcriptional activator, such as VP64, will enhance transcriptional activation when targeted near a promoter (Prashant Mali et al. 2013), effectively creating a synthetic, RNA-guided transcription factor. Furthermore, dCas9 can

be fused to a cytosine or adenine base editor, enzymes which are capable of generating all four transition mutations (Kantor, McClements, and MacLaren 2020). The great advantage to a dCas9 base editor is that there is no need to supply a repair template in trans if the mutation needed is a simple nucleotide conversion. In addition, the base editors avoid Cas9 induced double-stranded breaks. If NHEJ creates an indel at a Cas9 target site, the site cannot be targeted a second time by the same sgRNA, while also introducing new and potentially undesired mutations. dCas9 base editors avoid both of these issues, and Cas9 base editing has been used to create mouse models of breast cancer (Annunziato et al. 2020), repair mutations leading to Cystic Fibrosis in patient derived cells (Geurts et al. 2020), or create herbicide-resistant watermelon (Tian et al. 2018).

1.4 Prime Editing

In order to alter the genetic code in a specific manner using HDR, a DNA repair template containing the desired alteration needs to be supplied in trans with Cas9 and an sgRNA. A recently developed method called prime editing now allows a DNA sequence change to be inserted directly into the genome without a DNA repair template (Anzalone et al. 2019). This overcomes the challenge of delivering and co-localizing the repair template to the genomic site of Cas9 cleavage. Cas9 has RuvC and HNH nuclease domains, each of which will reliably cut only one of the DNA strands that are separated by a bound Cas9 RNP. These domains function independently and each can be disrupted by mutation. A Cas9 nickase (nCas9) bearing the H840A mutation has only a functional RuvC domain, which cleaves the strand of DNA not bound to the guide RNA. The prime editor was created through an ingenious fusion of nCas9 to the reverse transcriptase (RT) from the Murine Leukemia Virus. RT is a DNA polymerase that can use RNA as a template. In the virus, RT synthesizes DNA as directed by a viral RNA template during viral replication. During prime editing, a prime editing guide RNA (pegRNA)

provides an RNA template which is localized a site nicked by nCas9, allowing RT to synthesize a modified DNA sequence that can be used to repair the nick. The pegRNA is a sgRNA that has been elongated on its 3' end to have a primer binding site (PBS), a desired edit/repair template, and a homology arm that will promote the integration of a newly synthesized strand of DNA into the targeted genomic locus. At the beginning of a prime editing event, Cas9 binds to a genomic locus as directed by the pegRNA. The RuvC domain of Cas9 cleaves the unbound strand of DNA. The DNA is now free to hybridize with the primer binding site of the pegRNA. Reverse transcriptase binds to the DNA:RNA hybrid and begins to polymerize new DNA onto the end of the existing broken DNA strand complementary to the sequence encoded in the pegRNA. Reverse transcriptase stops at or near an RNA:RNA duplex that located at the 3' end of an unextended guide RNA. The sequence encoded in the pegRNA that is newly synthesized into the DNA will include a homology arm to promote the integration of the new DNA strand into the genome. Flap equilibration is followed by repair of the nicked site by endogenous enzymes. If the edited flap is extruded and cleaved, then the site can be restored to the original sequence by the other flap, and it can then potentially be targeted for prime editing again. Anzalone et. al. demonstrated that prime editing can be used to introduce point mutations, short insertions or deletions, or even longer insertions such as LoxP sites into a mammalian genome. The shorter the insertion, the more efficient the editing appeared to be. The potential uses of this technology are numerous, if the prime editing components can be delivered to cells efficiently. One obvious use is that it allows the rapid creation of cell or animal models with a mutated gene, or the correction of an existing mutation (e.g. in cells from a patient with a genetic pathology). Prime editing has now been demonstrated in mouse embryos (Liu et al. 2020), human T cells (Petri et al. 2021), and organoids (Schene et al. 2020), as well as in other organisms. Delivery of the

prime editor has been achieved by nucleic acid delivery (Liu et al. 2020), ribonucleoprotein (RNP) delivery (Petri et al. 2021), and by dual-AAV vectors (Zhi et al. 2021).

1.5 Mechanisms of Cas9 Delivery into Cells

The Turner laboratory is interested in studying retinal development and previously has used genetically modified mice to study genes crucial to retina development (Zhuang et al. 2020). Mouse models are expensive and time consuming in their creation and maintenance. Efficient genome editing in somatic cells could permit the use of common mouse strains for experiments where embryonic or postnatal retinas could be edited with Cas9. Commercially available mouse strains such as the Ai9 or Ribotag can be especially useful in conjunction with Cas9 mediated genome editing. We created a Mouse Dermal Fibroblast (MDF) line derived from a mouse homozygous for the Ribotag and heterozygous for the Ai9 reporter (See Chapter 3). The MDF line is maintained in inexpensive media and acts as a proving ground for our development of a delivery system for Cas9 and other genome editing enzymes. Although MDF cells do not have the exact same transfectability as primary retinal progenitor cells, the MDF cells are, like primary cells, a difficult to transfect cell type. Furthermore, while our delivery system has been largely optimized with MDF cells, it is not cell-type specific and it should be applicable for other cell types as well. It is well established that retroviruses and retroviral vectors can infect a wide range of cells types (Bouard, Alazard-Dany, and Cosset 2009). We have only performed one pilot experiment in retinal explants (see Chapter 4.2.8), but the initial result was promising and leads us to believe that further optimization of VLPs could make them a suitable vehicle for the transduction of genome editing enzymes into retinas as a standard practice.

Cas9 and an associated guide RNA can be delivered into cells either encoded in nucleic acids or pre-assembled as ribonucleoproteins. There are less common methods of delivery for

either of these molecules, but the discussion here will focus on the most common methods of delivery: viral vectors, virus-like particles, electroporation, and traditional transfection reagents (e.g. calcium phosphate, polyethyleneimine, or commercially available lipid-based transfection reagents such as lipofectamine or VirusGen). We chose not to pursue less common modes of Cas9 delivery such as microinjection, mechanical cell deformation, gold nanoparticles, and others (reviewed in (Lino et al. 2018)). Despite limiting our consideration to common delivery mechanisms, there was still a great breadth and depth of literature to consider in choosing the delivery mechanism that would best suit our research interests and potentially be useful beyond the delivery of Cas9 into MDF cells.

Nucleic acid delivery includes the co-delivery of Cas9 mRNA and in vitro transcribed or synthetic sgRNA, which is usually achieved via electroporation or traditional transfection reagents. We did not pursue this option because our long-term vision was to deliver not only Cas9 but also multiple enzymes simultaneously. This could prove difficult with separate mRNAs and guide RNAs delivered by electroporation or transfection reagents. Furthermore, neither electroporation or traditional transfection reagents can target delivery to specific cell types, which was another goal that we hoped to achieve in the long-term. Electroporation and transfection reagents are also known to rarely transfect all of the cells in a given population which we hoped to achieve with a higher efficiency delivery method. Nucleic acid delivery also includes the delivery of plasmid DNA expression vectors for Cas9 and the sgRNA. Similar to mRNA, plasmids are often delivered by transfection reagents which can be toxic to primary cells, or by electroporation, which can be harsh on delicate cells. Furthermore, due to multiple rounds of transcription and multiple rounds of translation, Cas9 can be produced from plasmid DNA in cells at high levels for a prolonged period. Off-target cleavage events (see Ch 1.2) by

Cas9 appear more frequently when Cas9 is expressed for a prolonged period of time. Expression can even become permanent if plasmid DNA is integrated into the genome inadvertently. This problem is not limited to transfected plasmid DNA, but also is a concern with viral vector delivered nucleic acids.

Viral vectors have been widely used for gene transfer, and there are many examples from the literature of Cas9 delivered by Adeno-Associated Virus (AAV) or lentiviral/retroviral vectors (reviewed in (Xu et al. 2019)). Viral vectors can be an attractive option for Cas9 delivery. AAVs have low immunogenicity and different serotypes can encourage delivery to different tissues thereby enabling targeted delivery. However, AAVs are limited in the size of their payload to about 4.7 kb. Dual AAVs where the desired payload can be split into two virions and recombined post-infection help to expand the effective carrying capacity. Lentiviral vectors are retroviral vectors derived from the Human Immunodeficiency Virus (HIV). They can be very efficient delivery mechanisms and can be pseudotyped with envelope proteins distinct from those native to the virus (Bouard, Alazard-Dany, and Cosset 2009). The use of pseudotyped lentivectors allows infection of a wide range of cell types and could potentially allow targeting to specific cell types, based on their receptors. However, lentiviruses, except for very specialized ones, integrate into the host genome. Retroviral vectors derived from the Murine Leukemia Virus (MLV) have also been used to deliver Cas9 (Williams et al. 2016). Normally they will integrate their payloads into the host genome, but one variation can deliver Cas9 as an RNA (Knopp et al. 2018). MLV virions can also be pseudotyped like lentiviral vectors (Schnierle et al. 1997; Lindemann et al. 1997). Despite their advantages in efficiency, low immunogenicity and potential for targeted delivery, viral vectors have some drawbacks as a delivery mechanism for Cas9. They are delivering Cas9 in the form of a nucleic acid which means that expression will be

prolonged, and as discussed above, that can increase the risk of off-target cleavage events. Lentiviral and retroviral vectors usually integrate into the host genome and potential disruption of a gene at the integration site is an additional mechanism for off-target effects. Furthermore, there are mechanisms for episomal DNA to integrate into the host genome, and AAV has a preferred integration site on chromosome 19 in human cells (Samulski et al. 1991). Integration of AAV actually has been observed with AAV delivery of Cas9. The AAV genome was integrated into the Cas9 double-stranded break site (Hanlon et al. 2019). Unintentional integration of DNA into a Cas9 cleavage site could create novel gene products that are harmful to cells. For these reasons, we decided not to attempt to engineer viruses for Cas9 delivery.

Aside from nucleic acid-based delivery, Cas9 can also be delivered as a pre-assembled RNP. Protein transduction is the delivery of proteins or RNPs into cells. This may initially seem to be a disadvantage because the protein cannot be amplified through transcription and translation. However, the transient presence of a protein can be an advantage when the protein potentially produces additional off-target events over time, such as Cas9 (Vakulskas and Behlke 2019). It is also easier to control the amount of Cas9 transduced into a cell and to have temporal control over Cas9 if it is transduced as an RNP. Importantly, a transduced protein can function immediately, without the temporal delay required for the transcription and/or translation of nucleic acids. There are other proteins that may also be effective when transduced in small quantities such as Cre recombinase, Flp recombinase, transcription factors, and other Cas9-derived genome editors such as the prime editor and base editors. Protein transduction is a method of delivering Cas9 that has been shown to work, but the question for us was which method of delivery could be optimized in MDF cells. Different methods of getting Cas9 past the

cell membrane include transfection reagents, electroporation, extracellular vesicles, and virus-like particles.

The plasma membrane is composed of a lipid bilayer impassable to all but the smallest solutes or solutes with specific transporters. The plasma membrane can be breached through physical methods such as microinjection or electroporation, but these require specialized equipment and may be impractical to scale and efficacy can vary greatly depending on the cell type. Vehicles such as synthetic nanoparticles, extracellular vesicles, or virus-like particles (VLPs) can deliver proteins without a causing a sudden disruption to the plasma membrane and are quite scalable. Each vehicle has advantages and disadvantages. Some synthetic nanoparticles have special properties such as magnetism to allow guidance to specific cells, but nanoparticles are also known to cause toxicity (Gul et al. 2019). Extracellular vesicles (EVs) are small, spherical structures with a lipid-bilayer known to be a biologically relevant method of transport for proteins from cell to cell (Y. Yang et al. 2018). VSV-g, an envelope protein from the vesicular stomatitis virus, can independently encourage the formation of EVs and promotes the fusion of vesicles with cells by receptor binding. The coexpression of VSV-g with Cas9 and gRNA will lead to the budding of EVs laden with Cas9 RNPs from producer cells (Montagna et al. 2018). These VSV-g coated EVs ("gesicles") can be harvested from supernatant and used to infect a wide variety of cell types. However, EVs do not necessarily load specific proteins during their formation, meaning that they might load other cellular proteins and RNAs which dilute the concentration of Cas9 RNPs in a complex mixture of EVs that cannot be easily separated. Additionally, Cas9 is normally expressed with a peptide nuclear localization signal which means that most of the protein should be nuclear localized, not incidentally adrift near the cell membrane (Lange et al. 2007). Furthermore, some envelope proteins useful for cell targeting

may be poorer at encouraging EV formation. Therefore, EVs on their own are not an optimal method for Cas9 protein transduction for all applications. This left us to consider engineering VLPs as an efficient vehicle for Cas9 RNP delivery.

1.5.1 Virus-Like Particles for the Transduction of Cas9 Ribonucleoproteins

VLPs can be made from many different types of viruses, but this introduction will be limited to the discussion of VLPs derived from retroviruses. Retrovirus-derived VLPs are made in mammalian cells by expressing the structural proteins that will polymerize at the plasma membrane and bud off as particles, without the need for a viral genome or non-structural viral proteins (Johnson et al. 2014). VLPs have an advantage over nanoparticles in that the protein to be transduced does not require purification prior to combination with the delivery vehicle. VLPs also provide an advantage over extracellular vesicles in that they can be designed to load specific cargo proteins concomitant with their formation. VLPs for protein transduction have been derived from the Avian Sarcoma Leukosis Virus (Kaczmarczyk et al. 2011), the Human Immunodeficiency Virus (Cai and Mikkelsen 2014; Cai, Bak, and Mikkelsen 2014; Cai et al. 2016; Skipper et al. 2018; Robert et al. 2017), and the Murine Leukemia Virus (Mangeot et al. 2019; D.-T. Wu and Roth 2014; Bobis-Wozowicz et al. 2015). VLPs derived from the Murine Leukemia Virus are a promising method for protein transduction and will be discussed below in terms of how they are made, design considerations, and their most recent examples to date.

1.6 The Murine Leukemia Virus

The Murine Leukemia Virus (MLV) is a member of the genus *gammaretrovirus* from the family *Retroviridae* (Rein 2011). MLV has been used as a model to study retroviruses, as a vector for the introduction of transgenes into mammalian cells, and more recently as a vehicle for

protein transduction. All retroviruses have *Gag*, *Pol*, and *Env* genes. Some retroviruses, like HIV, are more complex and have additional genes important to their life cycles. MLV has been used for decades as a model system for retroviruses because it is very simple, composed of only *Gag*, *Pol*, and *Env*. Gag is the main structural protein of the virus. MLV Gag is 538 amino acids in length and 65 kDa, as such it is often referred to as Gag Pr65. It is a polyprotein composed of four proteins, matrix (MA), p12, capsid (CA), and nucleocapsid (NC). There is no crystal structure for the protein, but small-angle X-ray scattering data suggests that it has an elongated rod-like structure (Datta et al. 2011).

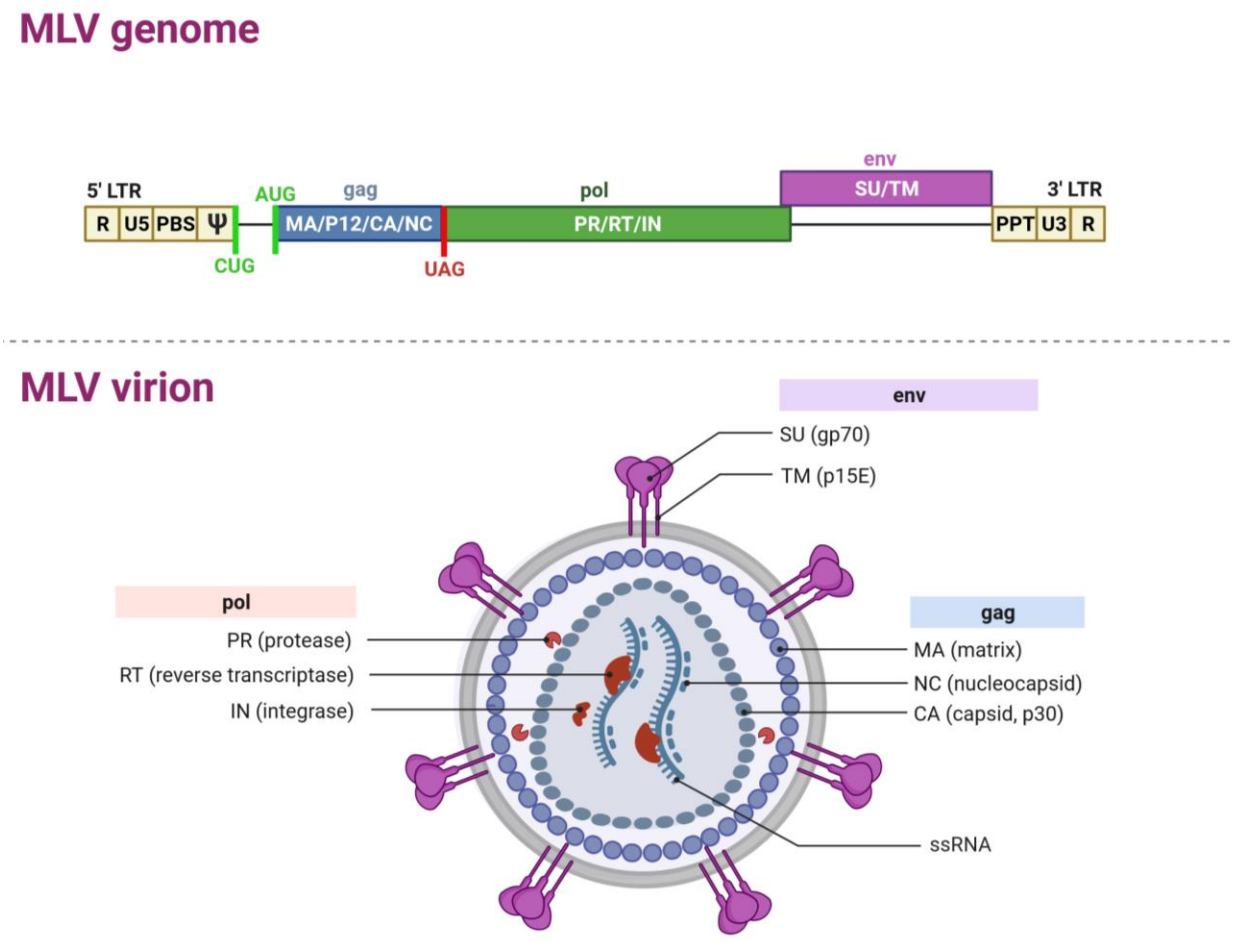

Figure 3 The Murine Leukemia Virus genome and virion. Created with BioRender.com.

Pol is produced from the same mRNA transcript as Gag. The ribosome reads through the Gag stop codon ~5% of the time to translate the Pol domain, creating one long polyprotein called GagPol. Ribosome readthrough of the stop codon is promoted by a pseudoknot in the mRNA, near the end of the Gag coding region. The Pol polyprotein contains protease (Prot), reverse transcriptase (RT), and integrase (IN) proteins. The envelope protein, Env, is produced from a separate subgenomic transcript. Env forms a trimer and localizes to the exterior of the plasma membrane (Fig. 3). After virions bud off, Env maturation is triggered by proteolytic processing of the C-terminal domain by the viral protease, allowing Env to bind to a receptor on cells, and promoting cell entry of virions.

1.6.1 Murine Leukemia Virus Lifecycle

The lifecycle of the MLV is typical of a retrovirus. The earliest event, or step 1, of the lifecycle is the binding of a virion to a cell surface receptor. This is mediated by the envelope protein. There are five classes of MLV with distinct Env proteins that are differentiable based on their receptor usage: ecotropic, amphotropic, 10A1, polytropic, and xenotropic (Battini, Heard, and Danos 1992; Lu and Roth 2001). One of the strains that uses the ecotopia Env is called the Moloney Murine Leukemia Virus (MMLV) and it is the strain used in this body of work. The SU domain of Env binds to the mouse cationic amino acid transporter 1 (mCAT1) receptor. The ecotropic Env is specific to rodents and will not bind to human CAT1. As a result of this specificity most studies of the MMLV have taken place in mice or in mouse cells such as the NIH/3T3 line. Binding of Env to a receptor results in fusion between the virion and the cell, step 2.

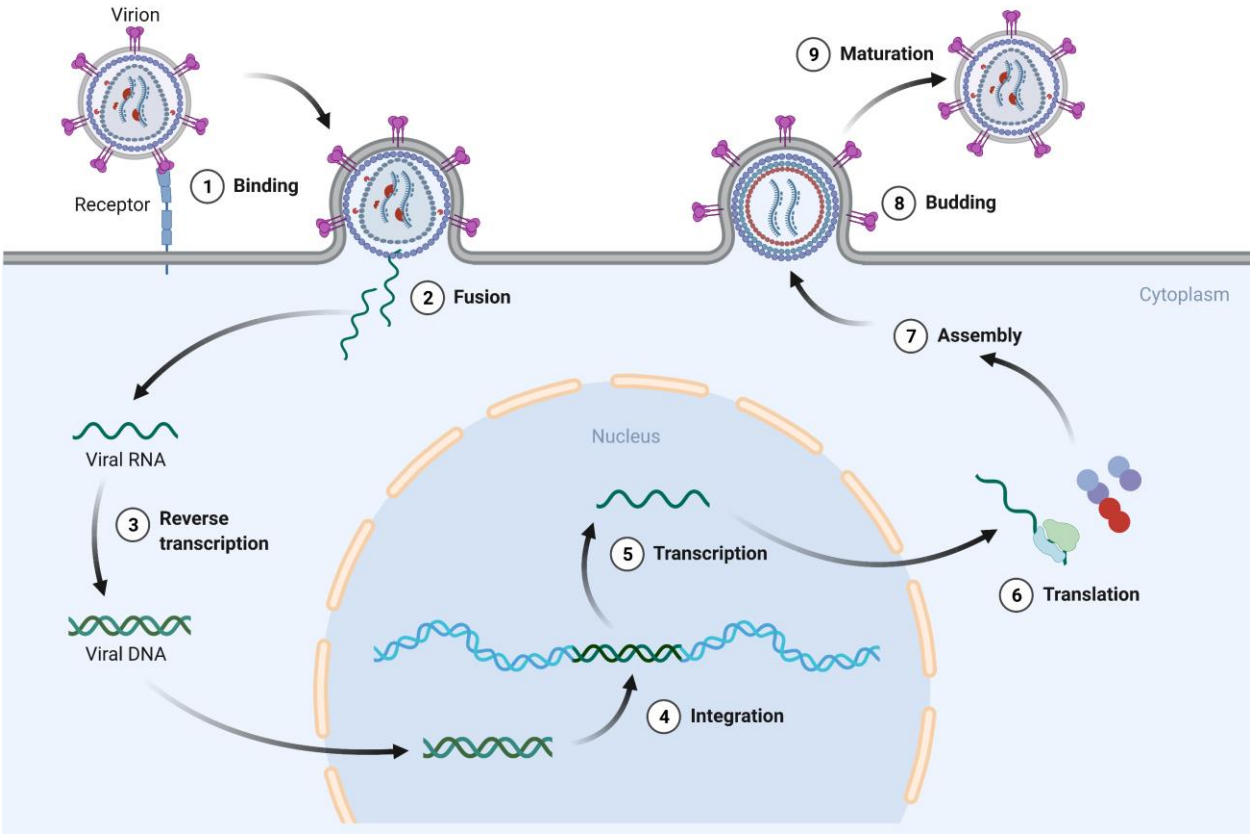

Figure 4 The simplified lifecycle of the Murine Leukemia Virus. Created with BioRender.com.

Once inside the cytoplasm, the core of the particle will dissociate and the viral RNA will replicate as shown in step 3. The newly created dsDNA that has been reverse transcribed by the RT will enter the nucleus. In step 4, the viral integrase incorporates the viral DNA into the genome. The newly inserted proviral DNA can then be transcribed by cellular proteins like any other gene in step 5. Many copies of the ssRNA will leave the nucleus where it can be translated into new protein in step 6. The new viral proteins are trafficked to the plasma membrane where they assemble into a new particle along with two copies of the viral genome (Roy et al. 1990). Finally, the particle will assemble at the cell membrane and be released in step 8. In step 9 the particle matures through protealytic processing of the Gag, GagPol and Env polyproteins. In

addition, daughter cells of the infected cell will inherit the integrated provirus as permanent genomic element.

1.6.2 Examples of Protein Transduction by MLV VLPs

There are four publications that show protein delivery by MLV VLPs that informed the experiments executed in this body of work. Their findings in terms of the proteins delivered, sites of fusion, pseudotypes, and cell types transduced are summarized in Table 1. In 2010, Voelkel et. al. showed that GFP and Flp recombinase could be delivered via MLV VLPs (Voelkel et al. 2010). They examined different locations in Gag for optimal cargo protein insertion. They found that insertion of the Flp recombinase after MA (a construct called MA.Prot.Flpo) yielded more protein transduction than insertion at the end of Gag (a construct called NC.Prot.Flpo). However, the MA.Prot.Flpo construct included the Pol domain whereas the NC.Prot.Flpo construct did not. This is significant because the Pol mRNA has sequences that promote protein translation (Bartels and Luban 2014) that could have increased production of the MA.Prot.Flpo protein which was likely the real cause of the high VLP titer. Indications of this can be inferred from data presented in Chapter 3. It is possible that the regulatory elements in the Pol mRNA increase protein production either through increasing export of GagPol mRNA from the nucleus or from improving translation by the ribosome, but we did not investigate the relative attribution of either of these possible effects. The use of the original codons in Voelkel et. al. means that the precise location where a protein is fused to Gag may have an impact on any negative regulatory sequences in Gag and therefore will confound the analysis of the location in Gag that is best for introducing fusion proteins. The disruption of regulatory elements that affect protein translation may have an equal or greater effect on VLP efficacy than alterations the Gag-fusion's ability to polymerize into a particle.

An interesting observation was that the Flp recombinase functioned whether or not a protease site was present to allow Flp to be cleaved from Gag. However, Flp appeared to function better with the addition of a protease site. The protease site was RSSLY/PALTP (a duplication of the MA/p12 site) in MA.Prot.Flpo. The native site QTSLL/TLDD was used in NC.Prot.Flpo. There may be a difference in protease cleavage efficiency depending on which cleavage site is used and where a cargo protein is fused to Gag, but we cannot disambiguate contributions of either of those factors from this data. Finally, they found that VLPs made with MA.Prot.Flpo and SF11tCD34 (a vector that codes RNA for CD34 that can be packaged into particles and integrated into a host genome) to deliver both the Flp protein and an integrating CD34 transgene. Dual delivery of genes and proteins could be very useful depending on the application. A viral vector delivers genes and proteins together, but here the protein delivered is not encoded by the viral vector.

In 2015, Bobis-Wozowicz et. al. compared MLV vectors designed to deliver DsRed-Express or a zinc finger nuclease (ZFN) as an integrating transgene (wild-type virus), a transgene expressed from episomal DNA, translated from transiently delivered RNA, or transduced as a protein (Bobis-Wozowicz et al. 2015). Only protein transduction will be discussed here. The vectors in which DsRed or ZFN were cloned came from Voelkel et. al. They found that transduced protein was far more stable when target K562 cells were maintained at 30°C versus 37°C. Temperature may be one factor that it is important to VLP stability and infectivity, but it has not been probed extensively by any of the existing literature. Although it is difficult to compare between the different assays used in Bobis-Wozowicz et. al. versus Voelkel et. al., the transient delivery of GFP-targeting ZFN proteins did not appear to be as efficient as the transient delivery of Flp recombinase. This could be due to the efficiency of the ZFN, the method of

concentrating the VLPs, the method of transduction, or the cell type transduced and its compatibility with the envelope protein pseudotyping the VLPs. All of this is to say that there are technical challenges and sources of error that could make VLP generation difficult, but also there is clearly room for improvement and tailoring depending on the cells that are being targeted for transduction, the protein transduced, and the measured outcome of protein transduction.

In 2014, Wu and Roth used MLV VLPS to transduce transcription factors Sox2, Oct4, Klf4, and c-Myc (D.-T. Wu and Roth 2014). All fusion proteins were put after the MA domain and an extra RSSLY/PALTP site was put in between the MA and cargo protein presumably to encourage complete excision of the cargo protein as in the MA.Prot.Flpo constructs. They found that the transcription factors were transduced into a reporter line in sufficient quantity to activate a GFP reporter in line with the level of activation observed when using MLV wild-type virus to integrate the transcription factor into the genome. They found that unlike GFP, the transcription factors translocated to the nucleus of the producer cell strongly enough that very few VLPs were released from the plasma membrane of the producer cells. They inserted nuclear export signals after the transcription factor and after the NC domain to improve targeting to the plasma membrane and VLP production. Furthermore, they found that coexpression of GagPol facilitated VLP release from the plasma membrane. A difference between Wu and Roth versus the other studies is that a lentiviral vector was used to make stable producer cell lines for the VLPs. This did not obviously produce higher titers of VLPs and may not be an advantage versus transiently transfecting HEK293T cells to make VLPs, as was done in the other papers. The use of transient transfection facilitates creating VLPs from varying combinations of plasmids, which may be particularly useful when creating VLPs to deliver proteins like Cas9 that can target a wide range of specific sequences. Making a stable packaging line was the motivation behind the choice of

envelope protein (EA6-3X) (O'Reilly and Roth 2003) because it can be expressed without inducing syncytia and death of the packaging line, a potential problem with the VSV-g protein. However, the EA6-3X envelope may not have been an optimal choice versus the more robust VSV-g. Our experiments have shown that VSV-g is a more effective envelope protein compared to the MMLV ecotropic envelope when infecting mouse dermal fibroblasts (see Chapter 3.2.1).

In 2019 Mangeot et al. was the most extensive study to date of protein transduction using MLV VLPs and showed several new aspects including the use of the VLPs in vivo, the first use of the Friend MLV rather than the Moloney MLV which had been used in all prior studies, and the exploration of new Env pseudotypes (Mangeot et al. 2019). The Friend MLV differs from Moloney slightly on the protein level and more substantially on the RNA level with many third position changes. The use of the Friend MLV possibly circumvented problems with the production of Moloney MLV Gag fusion proteins that are poorly translated. The exact sequences in Gag that impede translation have not been identified, but codon optimization of Moloney Gag improves protein production by as much as 30-fold (Bartels and Luban 2014) (see Chapter 3.2.1). Mangeot et al. chose to use VSV-g and BaEVRLess envelope proteins to pseudotype their VLPs, as they found that the combination of the two was better than either one on its own for transducing HEK293Ts or immortalized bone marrow macrophages. The other envelope proteins did not appear to work as well as the combination of VSV-g and BaEVRLess for macrophage transduction. The choice of envelope protein is an important one in the creation of VLPs as they have nearly no transducing ability without an envelope protein, and different envelope proteins can have differing efficiencies depending on the specific target cell type. The use of multiple envelope proteins could be important for overcoming receptor saturation or to realize the benefit of envelopes that may behave cooperatively. Finally, Mangeot et al. documented novel in vivo

applications of VLPs. Retroorbital injections into mice with Cas9 VLPs targeting *Hpd* successfully edited hepatocytes. Mouse zygotes were also edited to produce albino mice with up to 100% editing of alleles. VLPs show strong promise for more in vivo genome editing applications.

Table 1 Summary of MLV VLP Publications

Reference	Proteins Delivered	Fusion Sites	Pseudotypes	Cells Transduced
(Mangeot et al. 2019)	Cre, Cas9, dCas9-VPR	After NC	VSV-g, BaEVRLess, Influenza NA/HA/M, MuLV Amphotropic, Baculovirus gp64	HEK293T, U2OS, HepG2, hiPSCs, Mouse Bone marrow-derived macrophages (BMDMs), Human primary hepatocytes, CD34+ cells purified from human cord-blood, Mouse zygotes, Mouse hepatocytes in vivo.
(D.-T. Wu and Roth 2014)	GFP, Sox2, Oct4, c-Myc, Klf4, MazEF	After MA	EA6-3X chimeric Env (N261I/E311V/G552R)	Mouse Embryonic Fibroblasts (SNL), HeLa
(Bobis-Wozowicz et al. 2015)	DsRed-Express, ZFN	After MA, After NC	VSV-g	K562, U2OS, Murine ESCs (BK4-G3)
(Voelkel et al. 2010)	GFP, Flp	After MA, After NC, After IN	RD114-TR, VSV-g, Ecotropic Env	HT1080, HT1080mCAT, MEF (SC1), murine iPSCs

1.7 Properties of MLV Gag that Govern Particle Formation

VLPs are similar to virions with the key difference being that there is no viral genome inside a VLP. Although retrovirus-derived VLPs can contain the entire complement of viral

proteins, the only MLV protein both necessary and sufficient to make VLPs is Gag. Gag will polymerize at the cell membrane and bud off into the supernatant without the need for Pol, Env, or genomic RNA. Gag polymerizes side-to-side to form a lipid bilayer bound particle that is roughly 100nm in diameter and contains 1,500 – 2,000 Gag monomers. Gag's ability to independently form particles has been the key to all protein transduction through VLPs. Gag assembles at the cell membrane and buds off into the supernatant. Figure 5 below generally describes retrovirus assembly and budding with an emphasis on which parts of Gag are important for targeting to the cell membrane, Gag multimerization, and Gag higher order assembly.

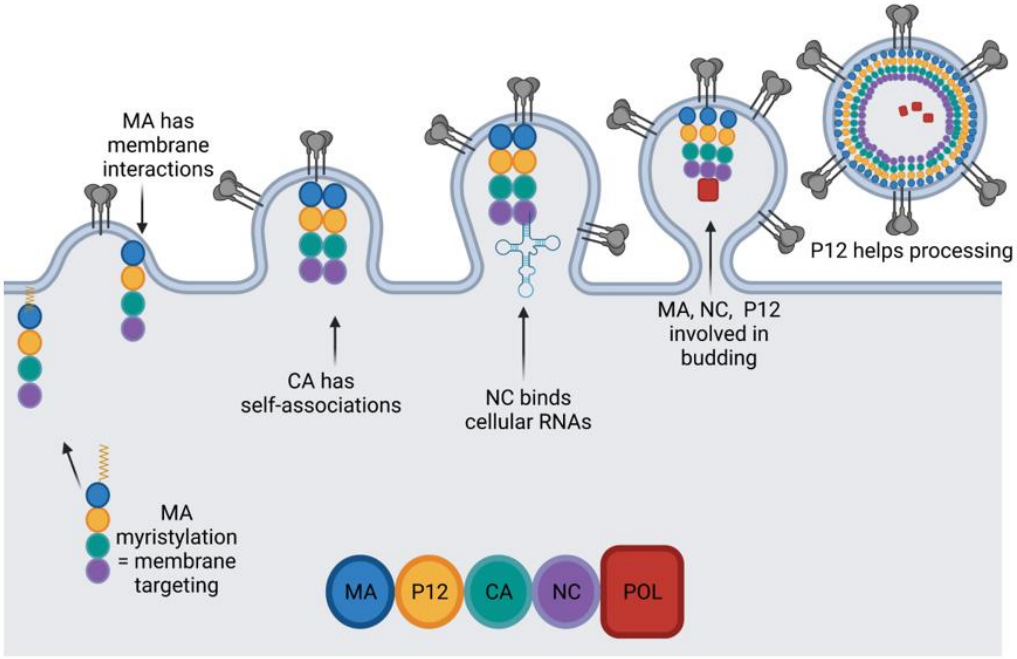

Figure 5 Particle assembly and budding are driven by different parts of Gag.
MA is myristylated which directs translocation to the cell membrane. MA has further interactions with phospholipids at the cell membrane. CA self-associates to drive higher order Gag assembly. NC has a zinc finger domain and binds the viral genome, but will also bind cellular RNAs in the absence of a viral genome. MA, NC, and P12 have all been implicated in particle budding. P12 is required for viral protease cleavage of Gag after budding. Created with BioRender.com.

VLP cargo proteins have been fused to the end of Gag, or placed within Gag, so that they are loaded into the VLPs concomitant with VLP creation. However, Gag and any proteins fused

to Gag will not be processed into separate proteins without the viral protease. VLPs reported in the literature thus far have included the viral GagPol polyprotein in order to incorporate a functional viral protease. Gag is composed of the matrix (MA), P12, capsid (CA), and nucleocapsid (NC) (Rein 2011). All studies to date have fused a VLP cargo protein to the entire Gag protein as the vehicle for loading cargo proteins into VLPs. It may be useful to briefly address each of the four proteins that compose Gag Pr65. They each have unique functions during virion assembly and different stages of the viral life-cycle. Gag is composed of the matrix (MA), P12, capsid (CA), and nucleocapsid (NC). All studies to date have used the entire Gag protein as the vehicle for loading cargo proteins into VLPs. A brief overview of the roles and functions of each of these four proteins in the native MLV can give a better understanding of how they may influence cargo protein loading and delivery. Each Gag-derived protein has unique functions during virion assembly and different stages of the viral life-cycle. Below, each protein is discussed in terms of its role in the assembly of virus-like particles with an emphasis on known mutations that may have implications for improving the function of VLPs. There also are some host factors that have been identified as partners for each protein, and in some cases they could be significant in modulating VLP creation or delivery, but that is beyond the scope of this discussion.

MA's function is primarily structural, and mutations have been made to determine which segments of MA are critical for particle formation. The factor most critical for MLV virion formation is myristylation. The matrix protein is myristylated at the second glycine after the removal of the first methionine at the N-terminus, which directs the Gag protein to the plasma membrane (Soneoka, Kingsman, and Kingsman 1997; F. Li et al. 2013). Mutations to a polybasic patch spans residues 17 to 34 can cause severe impairment of Gag localization to the

cell membrane, reduced binding to phosphatidylinositol-(4,5)-bisphosphate, and subsequently low viral titer (Hamard-Peron et al. 2010). Beyond localization to the cell membrane, there is also the question of multimerization. Mutant Gag that cannot be myristylated will not localize to the cell membrane, nor does it form puncta in cells. Mutant Gag that can still be myristylated but does not localize to the cell membrane due to mutations in the polybasic region can still form puncta in cells, depending on the mutation (F. Li et al. 2013; Soneoka, Kingsman, and Kingsman 1997; Hamard-Peron et al. 2010). This suggests that there is likely a combination of factors in MA governing multimerization of Gag, and the absence of MA would cause particles to not form. The deletion of MA and replacement with a membrane targeting domain enables targeting of Gag to the cell membrane, but not release of the particles (F. Li et al. 2013). This would suggest a functional role for MA in virion budding as well as membrane targeting.

P12 is important for virion assembly, protein processing, reverse transcription, and is part of the pre-integration complex. P12 is also called the late assembly domain or L-domain. Inside of the P12 domain is the PPPY motif. Virions can be made without the PPPY motif, but they are not infectious and Gag appears to be cleaved inefficiently (Yuan et. al. 2000). The PPPY motif can be moved to the MA or NC domains which then restores processing of Gag. Interestingly, the MLV P12 domain can be substituted for the L-domains of other retroviruses (e.g. HIV-1 or RSV) and still produce functional, replication-competent viruses (Bing Yuan et al. 2000). Disruption of the PPPY domain also gives rise to budding defects (B. Yuan 1999).

CA has self-associates to promote the multimerization of Gag (Andrawiss et al. 2003). CA also contributes to the structures of immature viral particles by self-associating in hexamers or pentamers to allow for the curvature of the assembling particles (Mayo, McDermott, and Barklis 2002). Deletions or mutations have been made to various parts of this protein. Some are

harmful to virion production (Auerbach, Brown, and Singh 2007). NC is primarily responsible for binding to the genomic RNA through its zinc finger domain during particle formation and for viral DNA synthesis (Gonsky, Bacharach, and Goff 2001). Although viral genomic RNA is dispensable for particle formation, the NC domain cannot be deleted itself and mutation of the zinc finger will cause defects in budding. The implication is that NC binds cellular RNAs (mostly mRNAs) in the absence of genomic RNA and that binding creates a scaffold for particle formation. Certain NC mutations can even shift the preference in packaging from mRNAs to rRNAs. Although several NC mutations have been made, none appear to improve particle formation but they could have implications for the directed packaging of certain RNA species (Muriaux et al. 2002; 2004).

1.8 Unresolved Questions Concerning Protein Transduction with MLV VLPs

Although protein transduction with VLPs derived from the MLV has been quite successful, there are still questions that remain open in terms of the mechanism by which the VLPs assemble and deliver proteins, and whether or not there are modifications that can be made to the VLPs in order to improve their efficacy. In the literature to date, MLV VLPs were made using GagPol coexpressed with the Gag fusion protein (the VLP cargo). The naturally occurring ratio of Gag to GagPol proteins in the virion is 95:5 which is about the same as the ratio of proteins translated from the single GagPol mRNA transcript. Johnson et al. found that when Gag and GagPol with a null protease are coexpressed from separate vectors, the ratio of Gag to GagPol in the particles appears to change as measured indirectly by western blot (Johnson et al. 2014). However, the absolute amount of GagPol released into the supernatant from producer cells drops dramatically as the ratio of Gag to GagPol decreases. Therefore, there is a balance in maximizing the number of particles released and the amount of GagPol in the particles. It is

similarly possible that different Gag fusion proteins will experience this inverse relationship and that optimizing the ratio of coexpressed Gag to Gag fusion protein will help improve VLP efficacy and/or cargo delivery. It is also possible that some fusion proteins can form VLPs without the coexpression of Gag. Expression of different ratios of the Gag, GagPol, and Gag proteins for VLP production can test this hypothesis. Additionally, because GagPol has been coexpressed with Gag fusion proteins, all VLPs in the literature to date have co-delivered the integrase and reverse transcriptase proteins. These proteins are essential to the lifecycle of the virus, but may not be necessary for the delivery of cargo proteins. The data presented in Chapters 3 and 4 show that there are multiple ways of improving Cas9 VLP titers. The most significant improvement made was to codon optimize *Gag*.

Chapter 2 Programmable RNA Cleavage and Recognition by a Natural CRISPR-Cas9 System from Neisseria Meningitidis

This chapter was adapted from Rousseau, Beth A., Zhonggang Hou, Max J. Gramelspacher, and Yan Zhang. 2018. "Programmable RNA Cleavage and Recognition by a Natural CRISPR-Cas9 System from Neisseria Meningitidis." *Molecular Cell* 69 (5): 906-914.e4. Author contributions: B.A.R, Z.H. and Y.Z. designed and conducted experiments. M.J.G. and Z.H. purified proteins. Z.H. and Y.Z. wrote the manuscript. All authors edited the manuscript. The experiments described in this chapter were done in the laboratory of Dr. Yan Zhang.

2.1 Introduction

The microbial CRISPR systems enable adaptive defense against mobile elements, and also provide formidable tools for genome engineering. The Cas9 proteins are Type II CRISPR-associated, RNA-guided DNA endonucleases that identify double-stranded DNA targets by RNA-DNA complementarity and protospacer adjacent motif (PAM) recognition. Here we report that the Type II-C CRISPR-Cas9 from *Neisseria meningitidis* (Nme) is capable of programmable, RNA-guided, sequence-specific cleavage and recognition of single-stranded RNA targets. Importantly, unlike the *Streptococcus pyogenes* (Spy) Cas9, NmeCas9's natural ribonuclease activity does not rely on any PAM-presenting oligonucleotide (PAMmer) co-factor supplied in *trans*. Furthermore, we have defined the functional determinants, mechanistic features, and specificity constraint for *in vitro* RNA cleavage by NmeCas9, and also show that nuclease-null dNmeCas9 binds to RNA targets complementary to the CRISPR RNA (crRNA) guide, in a PAM- and PAMmer oligo- independent manner. Finally, we demonstrate that

NmeCas9-catalyzed RNA cleavage can be blocked by three families of Type II-C anti-CRISPR proteins. These results fundamentally expand the targeting capacities of CRISPR-Cas9, and highlight the potential utility of NmeCas9 as a single platform to target both RNA and DNA.

Clustered, regularly interspaced, short, palindromic repeats (CRISPR) loci and their associated (*cas*) genes constitute an adaptive defense system widespread in bacteria and archea that limits horizontal genetic transfer (Barrangou et al. 2007; Marraffini and Sontheimer 2008). CRISPR spacers, which are acquired from invader's genome and integrated between CRISPR repeats, specify the nucleic acid targets for CRISPR interference (Barrangou et al. 2007; Marraffini 2015). The CRISPR locus is transcribed and processed into small RNAs called CRISPR RNAs (crRNAs), which guide the Cas protein effectors to destroy complementary targets (Brouns et al. 2008; Wright, Nuñez, and Doudna 2016). Based upon *cas* gene content, the diverse CRISPR-Cas systems are categorized into two Classes, six major Types and nearly thirty subtypes (Koonin, Makarova, and Zhang 2017), and the majority of these systems confer interference by DNA targeting (Marraffini and Sontheimer 2008; Wright, Nuñez, and Doudna 2016; Garneau et al. 2010).

The Cas9 proteins, which are the single protein effectors of Type II CRISPRs, generally function as RNA-guided DNA endonucleases (Jinek et al. 2012; Gasiunas et al. 2012), and provided revolutionary tools for programmable genome engineering in eukaryotes and prokaryotes, with the Type II-A SpyCas9 being the most commonly used (Cho et al. 2013; Cong et al. 2013; Hwang et al. 2013; Jiang et al. 2013; Jinek et al. 2013; P. Mali et al. 2013). When programmed by crRNA and another RNA cofactor called tracrRNA (Deltcheva et al. 2011), SpyCas9 targets double-stranded DNA (dsDNA) through the recognition of PAM sequence, DNA unwinding, R-loop formation, and the triggering of DNA scission (Sternberg et al. 2014).

The HNH and RuvC nuclease domains of Cas9 cleave the crRNA-complementary and non-complementary target strands, respectively (Jinek et al. 2012; Gasiunas et al. 2012). Catalytically inactive "dead" Cas9s (dCas9s), which can bind to DNA targets without inducing any DNA breaks (Jinek et al. 2012), have also been harnessed as eukaryotic genome-binding platforms to deliver effector domains for the regulation, imaging or modification of the specific chromosomal loci (Wright, Nuñez, and Doudna 2016). In addition, the newly discovered Type II anti-CRISPR (Acr) proteins, encoded by mobile genetic elements (MGEs), are potent Cas9 inhibitors and can block Cas9-mediated DNA cleavage and gene editing (Harrington et al. 2017; Pawluk et al. 2016; Rauch et al. 2017). These and other Cas9-based DNA manipulation tools are transforming biomedical research.

There are a few CRISPR systems capable of RNA targeting. For example, the Type VI effector Cas13 is a promiscuous RNase, which upon activation by crRNA-guided RNA recognition, degrades nearby RNAs non-specifically (Abudayyeh et al. 2016; East-Seletsky et al. 2016). The Type II-B Cas9 from *Francisella novicida* (Fno) employs the tracrRNA, rather than crRNA, as a guide to silence an endogenous transcript, yet the protein encoding the ribonuclease activity remains unidentified (Sampson et al. 2013). Type III CRISPRs use large multi-protein complexes to confer immunity through transcription-dependent co-degradation of the DNA and its transcripts (Samai et al. 2015). And a separate Type III-associated RNase, Csm6, provides signal-activated, non-specific RNA clearance (Jiang, Samai, and Marraffini 2016; Kazlauskiene et al. 2017; Niewoehner et al. 2017).

Several reports indicated that Cas9s lack crRNA-guided RNA cleavage activity (Gasiunas et al. 2012; Ma et al. 2015). SpyCas9, however, can be tricked into cleaving RNAs *in vitro* by an exogenously supplied PAMmer DNA oligo that hybridizes to the ssRNA target-

flanking region and presents the DNA PAM on the opposing strand. Specific binding of SpyCas9 to the RNA target can also be achieved with a longer PAMmer that extends into the crRNA-complementary region (O'Connell et al. 2014). These *in vitro* findings open up doors to develop CRISPR-based RNA-targeting tools to recognize and manipulate specific RNAs *in vitro* and *in vivo*. Indeed, a PAMmer- and dSpyCas9-based strategy has enabled the pull-down of RNAs from cell extracts (O'Connell et al. 2014), and the visualization of stress granule RNAs in mammalian cells (Nelles et al. 2016). In addition, dSpyCas9 was repurposed to eliminate disease-associated, toxic repetitive RNAs (Batra et al. 2017), and more recently, the Cas13 RNase was adopted for *in vivo* knockdown and editing of mammalian transcripts (Abudayyeh et al. 2017; Cox et al. 2017).

Despite these advances, it is unclear if any other Cas9 ortholog has PAMmer-inducible or even intrinsic crRNA-programmable RNase activity. Here we explore the RNA-targeting potential of Nme CRISPR-Cas9, a Type II-C system previously shown to limit DNA natural transformation in its native context *N. meningitidis* (Y. Zhang et al. 2015) and has been adopted as an eukaryotic gene-editing platform (Hou et al. 2013; Esvelt et al. 2013; Lee, Cradick, and Bao 2016; Y. Zhang 2017). We find that NmeCas9 has a natural, PAMmer oligo-independent, programmable RNase activity *in vitro*. We describe the functional determinants, mechanistic features, and specificity constraints for RNA cleavage by NmeCas9, and show that this activity can be blocked by three families of Type II-C Acr proteins. We also find that catalytically inert dNmeCas9 binds to RNA targets in a sequence-specific manner.

2.2 A natural RNA-guided ribonuclease activity of NmeCas9 *in vitro*

The Type II-C CRISPR-*cas* locus from *N. meningitidis* strain 8013 consists of three *cas* genes (*cas1*, *cas2*, *cas9*), a *tracrRNA* locus, and a CRISPR array (Y. Zhang et al. 2013) (Fig.

6A). We expressed recombinant, FLAG-tagged NmeCas9 in *E. coli* and isolated it by heparin, ion exchange, and size exclusion chromatography (Rousseau et al. 2018). *In vitro* RNA cleavage assays were performed using purified protein, *in vitro*-transcribed tracrRNA and crRNA, and a fluorescent end-labeled ssRNA oligonucleotide bearing a target region complementary to the spacer of crRNA-sp25 (Fig. 6B). We found that NmeCas9 efficiently catalyzes *in vitro* cleavage of the RNA substrate, resulting in one prominent labeled cleaved product, and that this reaction requires the cognate crRNA, the tracrRNA, and Mg^{2+} or Mn^{2+} (Fig. 6C) (Rousseau et al. 2018). A DNA guide containing sequences identical to crRNA-sp25 can not support RNA cleavage, indicating that NmeCas9's RNase activity is strictly RNA-guided (Rousseau et al. 2018).

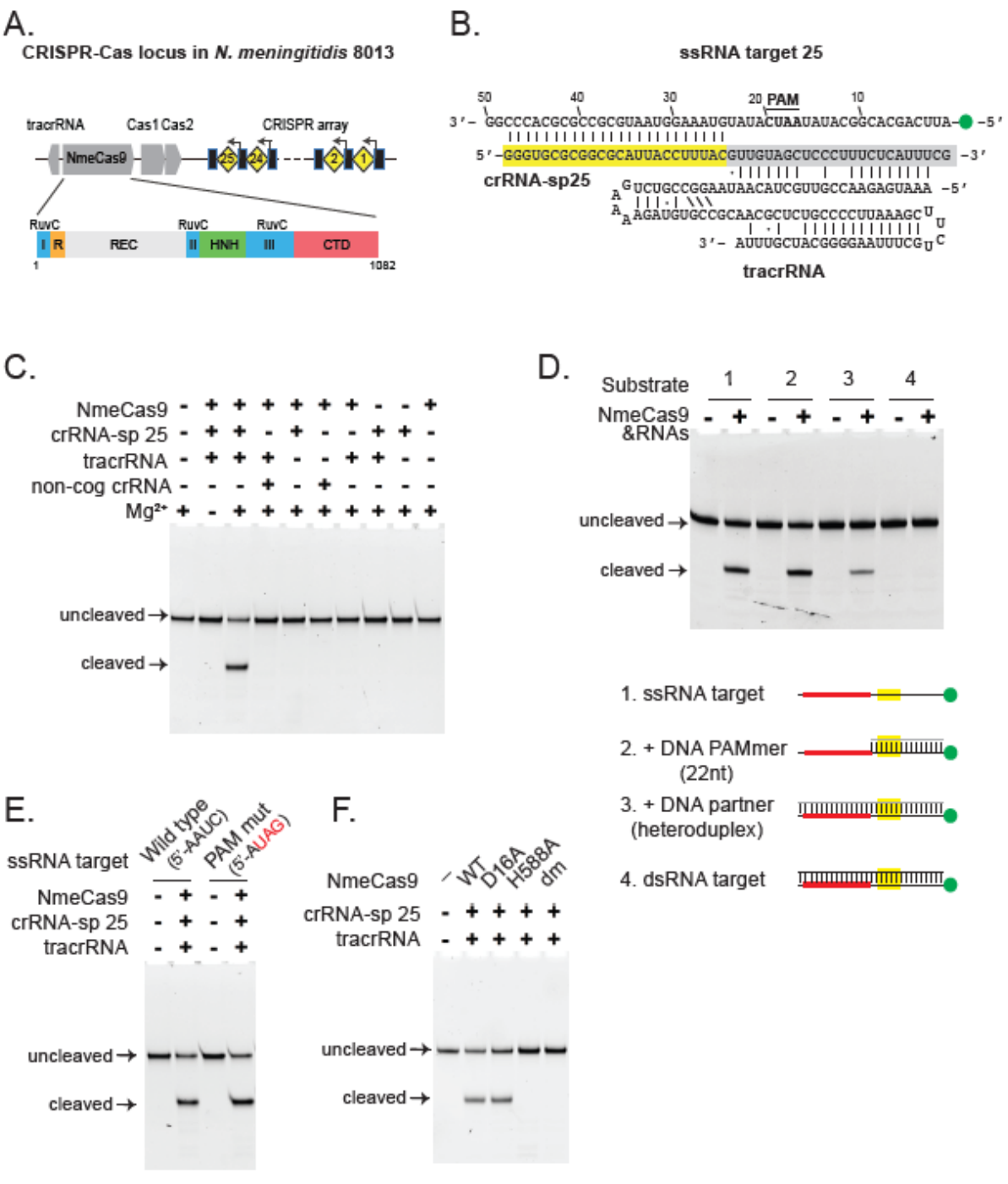
A.
CRISPR-Cas locus in N. meningitidis 8013
tracrRNA
Cas1 Cas2
CRISPR array
NmeCas9
25
24
2
1
RuvC
I
R
REC
RuvC
II
HNH
RuvC
III
CTD
1
1082
B.
ssRNA target 25
50
40
30
20 PAM
10
3′– GGCCCACGCGCCGCGUAAUGGAAAUGUAUACUAAUAUACGGCACGACUUA –5′
5′–GGGUGCGCGGCGCAUUACCUUUACGUUGUAGCUCCCUUUCUCAUUUCG –3′
crRNA-sp25
A GUCUGCCGGAAUAACAUCGUUGCCAAGAGUAAA –5′
A AGAUGUGCCGCAACGCUCUGCCCCUUAAAGC U U
3′– AUUUGCUACGGGGAAUUUCG U C
tracrRNA
C.
NmeCas9 - + + + + + + - - +
crRNA-sp 25 - + + - + - - + + -
tracrRNA - + + + - - + + - -
non-cog crRNA - - - + - + - - - -
Mg2+ + - + + + + + + + +
uncleaved
cleaved
D.
Substrate 1 2 3 4
NmeCas9 &RNAs - + - + - + - +
uncleaved
cleaved
1. ssRNA target
2. + DNA PAMmer (22nt)
3. + DNA partner (heteroduplex)
4. dsRNA target
E.
ssRNA target
Wild type (5′-AAUC)
PAM mut (5′-AUAG)
NmeCas9 - + - +
crRNA-sp 25 - + - +
tracrRNA - + - +
uncleaved
cleaved
F.
NmeCas9 / WT D16A H588A dm
crRNA-sp 25 - + + + +
tracrRNA - + + + +
uncleaved
cleaved

Figure 6 NmeCas9 possesses a natural RNA-guided ribonuclease activity in vitro.
(A) Schematics depicting the CRISPR-Cas9 of N. meningitidis strain 8013, and the domain organization of NmeCas9. Individual elements are not drawn to scale. Black rectangles, CRISPR repeats; yellow diamonds, CRISPR spacers; grey boxes, cas1, cas2, cas9 and tracrRNA genes; black arrows, transcription driven by repeat-embedded promoters. CTD, C-terminal domain; R, arginine-rich motif ; REC, recognition domains; HNH, HNH domain; RuvC, RuvC domains. (B) A schematic for the complex of the crRNA-sp25, the tracrRNA and the ssRNA target 25. Yellow, crRNA spacer; grey, crRNA repeat; bold, PAM equivalent (DNA PAM for NmeCas9 is N_4GATT on the non-target strand and $AATCN_4$ on the target strand); green, FAM label. (C) NmeCas9 cleaves ssRNA target efficiently in vitro, and this reaction requires the cognate crRNA, the tracrRNA and divalent metal Mg^{2+}. Non-cog crRNA, crRNA-sp23. The uncleaved probe and cleavage products were indicated. (D) An exogenously supplied PAMmer DNA oligo has a modest effect on NmeCas9-catalyzed ssRNA cleavage. The panel of pre-annealed nucleic acid substrates (depicted at the bottom) were assayed for in vitro RNA cleavage as in (C). Black, RNA; grey, DNA; red, crRNA-complementary sequence; yellow, PAM equivalent region; green, FAM label. (E) NmeCas9 cleaves ssRNA in a PAM-independent manner. A PAM mutant RNA substrate as analyzed for cleavage as in (C). Red, mutated sequences in the PAM equivalent region. (F) An intact HNH domain is required for RNA cleavage by NmeCas9. Nuclease domain active site mutants of NmeCas9 were tested as in (C). Dm, double mutant (D16A +H588A).

Strikingly, unlike SpyCas9, NmeCas9 catalyzed efficient RNA cleavage without any PAMmer oligo co-factor (Fig. 6C), reflecting a fundamental distinction between the RNase activities of these two Cas9 orthologs. We also tested whether a similar PAMmer strategy would modulate RNA cleavage by NmeCas9, by pre-annealing the same ssRNA target to various top strand partners (Fig. 6D) and analyzing the resulting substrates. RNA cleavage was minimally enhanced by a 22 nucleotide (nt) DNA PAMmer, but greatly impeded or completely abrogated on a fully annealed DNA-RNA heteroduplex or on a double-stranded RNA substrate (Fig. 6D, upper) (Rousseau et al. 2018). These results reveal that NmeCas9-mediated RNA cleavage is specific for ssRNA targets and is independent of any PAMmer provided in trans.

Next, we tested if NmeCas9's ribonuclease activity requires a PAM equivalent (5'-$AAUCN_4$-3' for NmeCas9) within the ssRNA target, by assaying a mutant RNA substrate with 3 nt mutations (AAUC to AUAG) introduced into the PAM motif (Fig. 6E). The same triple mutation was sufficient to abolish dsDNA cleavage by NmeCas9 (Y. Zhang et al. 2015). This PAM mutant RNA substrate was cleaved as efficiently as the wild type counterpart (Fig. 6E), indicating that RNA cleavage by NmeCas9 is PAM-independent. All Cas9 enzymes described to

date employ the HNH and RuvC domains to cut the crRNA-complementary and non-complementary strands of the dsDNA targets, respectively (Jinek et al. 2012; Gasiunas et al. 2012). For NmeCas9, active site residues D16 in the RuvC domain and H588 in the HNH domain were previously shown to be essential for dsDNA targeting *in vivo* and *in vitro* (Y. Zhang et al. 2015; 2013). To test the involvement of these two active sites in RNA cleavage, we purified and analyzed three mutant NmeCas9 proteins [D16A, H588A, and the double mutant (dm, D16A+H588A)] (Rousseau et al. 2018). RNA cleavage was abolished for the H588A and dm proteins, but not for the D16A mutant (Fig. 6F), suggesting that the HNH domain of NmeCas9 mediates the ssRNA cleavage, and therefore is capable of both DNA and RNA scission. In addition, a time-course experiment revealed that RNA cleavage at 37°C became detectable within one minute and plateaued after 30 minutes, and occurs more slowly at room temperature (Rousseau et al. 2018). RNA cleavage was robust under a wide range of NmeCas9-RNA ribonucleoprotein (RNP) concentrations or monovalent salt concentrations (KCl) (Rousseau et al. 2018). By analyzing serial crRNA mutants with 5' truncations (in the 24 nt spacer portion) or 3' truncations (in the 24 nt repeat portion), we found that RNA cleavage is reduced with a 18-20 nt spacer and completely lost with a 16 nt spacer, and that the first 8 nts of the crRNA repeat is sufficient to support robust RNA cleavage by NmeCas9 (Rousseau et al. 2018).

2.3 RNA-guided RNA cleavage by NmeCas9 is programmable

To assess the programmability of the NmeCas9 ribonuclease, we first attempted to redirect RNA cleavage to different positions within the same ssRNA target. We created two variants of the crRNA-sp25, walk-2 and walk+3, which match different regions of the same target (Fig. 7A). These two variants were analyzed alongside the wild type counterpart in a

cleavage site mapping experiment. Both variants directed *in vitro* RNA cleavage, and importantly their cleavage sites moved in concert with the guide-target complementarity (Fig. 7B). By comparing the NmeCas9 cleavage products with RNase T1 and hydrolysis ladders, we found that the wild type crRNA-sp25 and the two variants predominantly directed cuts after A27, G25 and G30, respectively (Fig. 7B), indicating that NmeCas9 catalyzes RNA scission between the 3rd and the 4th nts of the crRNA-paired target region proximal to the 5' end of target (Fig. 7A). This is consistent with the DNA double strand breaks (DSBs) between the 3rd and the 4th nts generated by NmeCas9 (Y. Zhang et al. 2015) and other Cas9 orthologs (Jinek et al. 2012; Gasiunas et al. 2012).

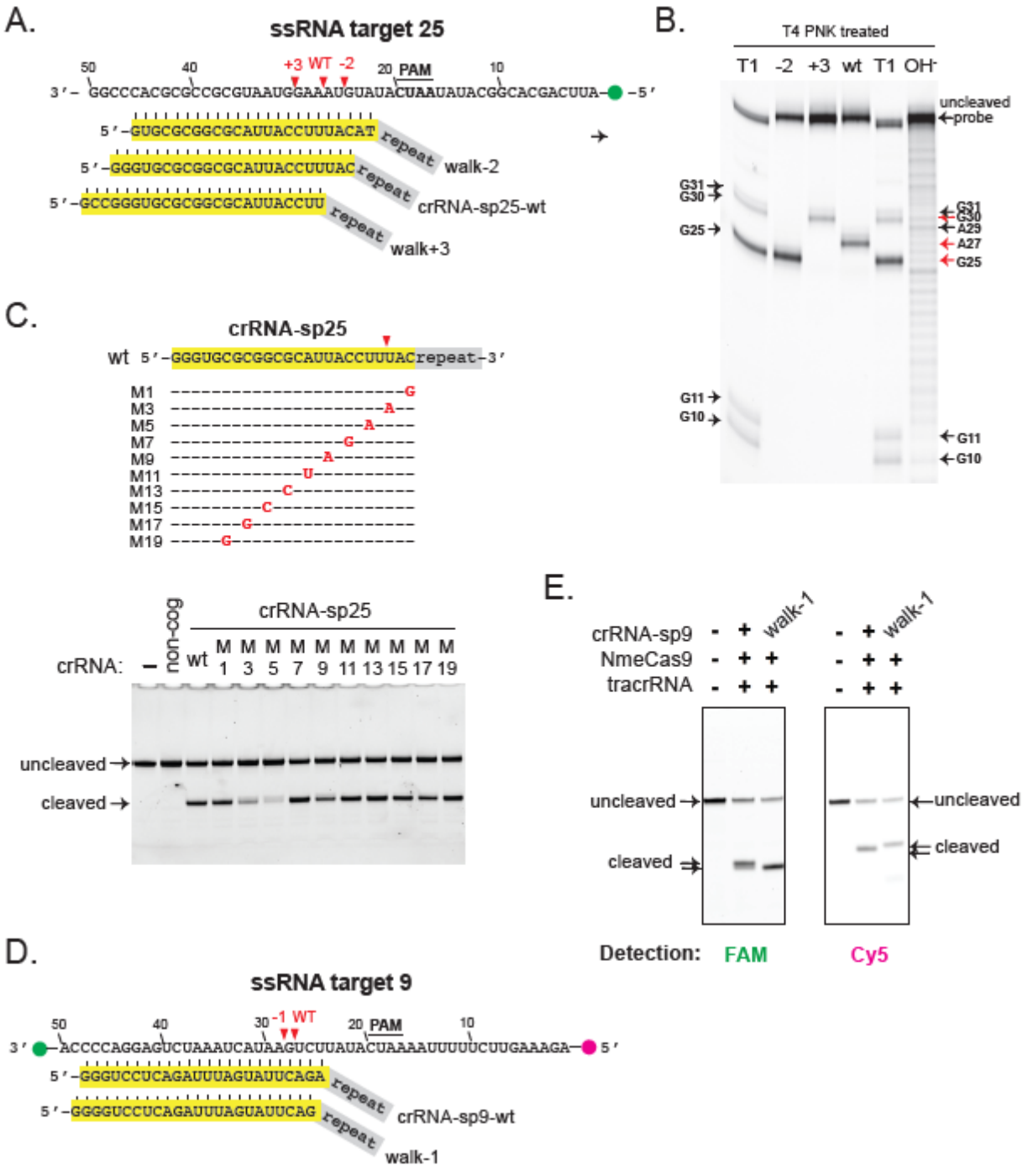
A.
ssRNA target 25
+3 WT -2
PAM
crRNA-sp25-wt
walk-2
walk+3
B.
T4 PNK treated
T1 -2 +3 wt T1 OH-
uncleaved
probe
G31
G30
G25
G11
G10
A29
A27
C.
crRNA-sp25
wt
repeat
M1
M3
M5
M7
M9
M11
M13
M15
M17
M19
crRNA:
non-cog
crRNA-sp25
uncleaved
cleaved
D.
ssRNA target 9
-1 WT
crRNA-sp9-wt
walk-1
E.
crRNA-sp9
NmeCas9
tracrRNA
walk-1
Detection:
FAM
Cy5

Figure 7 NmeCas9-catalyzed RNA cleavage is programmable.
(A) A schematic depicting the ssRNA target 25 and three matching crRNAs, crRNA-sp25-wt, walk-2, and walk+3). Yellow, crRNA spacer; grey, crRNA repeat; bold, PAM equivalent; green, FAM label; red arrows, predominant RNA cleavage sites mapped out in (B) (B) RNA cleavage site mapping experiment. SsRNA target 25 and the three crRNAs used are shown in (A). The NmeCas9 cleavage products and RNase T1 and hydrolysis ladders were subjected to 3' de-phosphorylation by T4 Polynucleotide kinase, and separated by 15% denaturing PAGE. T1, RNase T1 ladder; OH^-, hydrolysis ladder. -2, walk-2; +3, walk+3; wt, crRNA-sp25. Sites of G cleavage by RNase T1 are indicated; sites of NmeCas9 cleavage (G30, A27, G25 for the three crRNAs, respectively) are marked by red arrows. (C) Serial single nt mutants of crRNA-sp25 were analyzed for NmeCas9-catalyzed RNA cleavage. M1 through M19, single nt mutation introduced into every other position of the crRNA spacer. The location and sequence of each mutation (in red) are shown at the top. Yellow, crRNA spacer; grey, crRNA repeat; red arrow, RNA cleavage site. Non-cog, crRNA-sp23. (D) A schematic depicting the ssRNA target 9 and the two matching crRNAs (crRNA-sp9-wt and walk-1). Yellow, crRNA spacer; grey, crRNA repeat; bold, PAM equivalent; red arrows, predicted RNA cleavage sites; green, FAM label; magenta, Cy5 label. (E) NmeCas9's ribonuclease activity is re-programmable on a different RNA substrate. The two crRNAs shown in (D) were assayed for in vitro cleavage on ssRNA target 9. The same denaturing gel is subjected to FAM (shown on the left) and Cy5 (shown on the right) detection.

Next, we sought to investigate the specificity rule governing NmeCas9's tolerance for mismatches in the crRNA-RNA target complementarity. We created and tested serial crRNA mutants each bearing a single nt deviation from the wild-type sequence at every odd position within the spacer (Fig. 7C, upper panel). Only the two single mismatches at the 3rd and 5th nt, which are close to the cleavage site, abrogated *in vitro* RNA cleavage; whereas the other mutations either didn't affect cleavage or (e.g. at the 9th or 17th nt) only caused modest defects (Fig. 7C, lower panel). We also analyzed crRNA mutants with multiple mismatches and found that RNA cleavage was diminished by short (2-4 nts) mutations clustered around the cleavage site, but was only partially reduced by 2-3 nts of mismatches in regions away from the cleavage site (Rousseau et al. 2018). CrRNAs with 4 or more nts of mismatches all exhibited severe cleavage defects (Rousseau et al. 2018). Overall, the NmeCas9 ribonuclease has certain degree of tolerance for guide-target mismatches that are not next to the cleavage site.

Programmable RNA cleavage by NmeCas9 was also observed with a different dual fluorophore-labelled RNA substrate bearing the target sequence for spacer 9 of *N. meningitis* strain 8013 (Fig. 7D). RNA cleavage guided by crRNA-sp9 resulted in one predominant 5'

product and one major 3' product (Fig. 7E, Cy5- and FAM-labeled respectively). For a variant of crRNA-sp9, walk-1, which has the guide-target pairing region shifted by 1 nt, the cleavage site moved in concert (Figs. 7D-7E). Collectively, these results demonstrate that NmeCas9 ribonuclease is crRNA-guided and programmable.

2.4 Type II-C anti-CRISPRs inhibit *in vitro* RNA cleavage by NmeCas9

In light of the recent discovery of Type II Acrs that inhibit Cas9-mediated genetic interference, DNA cleavage and genome editing, (Harrington et al. 2017; Pawluk et al. 2016; Rauch et al. 2017; Hynes et al. 2017) we wondered whether the three families of Type II-C Acrs also inhibit NmeCas9-catalyzed RNA cleavage. We purified four Acr proteins, AcrIIC1$_{Nme}$, AcrIIC2$_{Nme}$, AcrIIC3$_{Nme}$ and AcrIIA4 (Fig. 8A) and analyzed them in our RNA cleavage assay. Notably, pre-incubation of NmeCas9 with increasing amounts of AcrIIC1$_{Nme}$, AcrIIC2$_{Nme}$, or AcrIIC3$_{Nme}$ all resulted in blockage of RNA cleavage dose-dependently, and near complete inhibition was achieved with all the three Acr proteins at or above 3-fold molar excess over NmeCas9 (Fig. 8B). In contrast, cleavage was not affected by increasing amounts of AcrIIA4 (Fig. 8B), an control Acr that specifically blocks DNA cleavage by two Type II-A Cas9s, SpyCas9 and *Listeria monocytogenes* (Lmo) Cas9 (Rauch et al. 2017). In a quality control plasmid cleavage experiment, the three Type II-C Acrs all prevented linearization of a target plasmid DNA by NmeCas9 (Fig. 8C), consistent with previous reports (Harrington et al. 2017; Pawluk et al. 2016; Rauch et al. 2017). In summary, these results showed that in addition to their known roles in blocking DNA targeting, AcrIIC1$_{Nme}$, AcrIIC2$_{Nme}$, and AcrIIC3$_{Nme}$ can all block NmeCas9-catalyzed *in vitro* RNA cleavage as well.

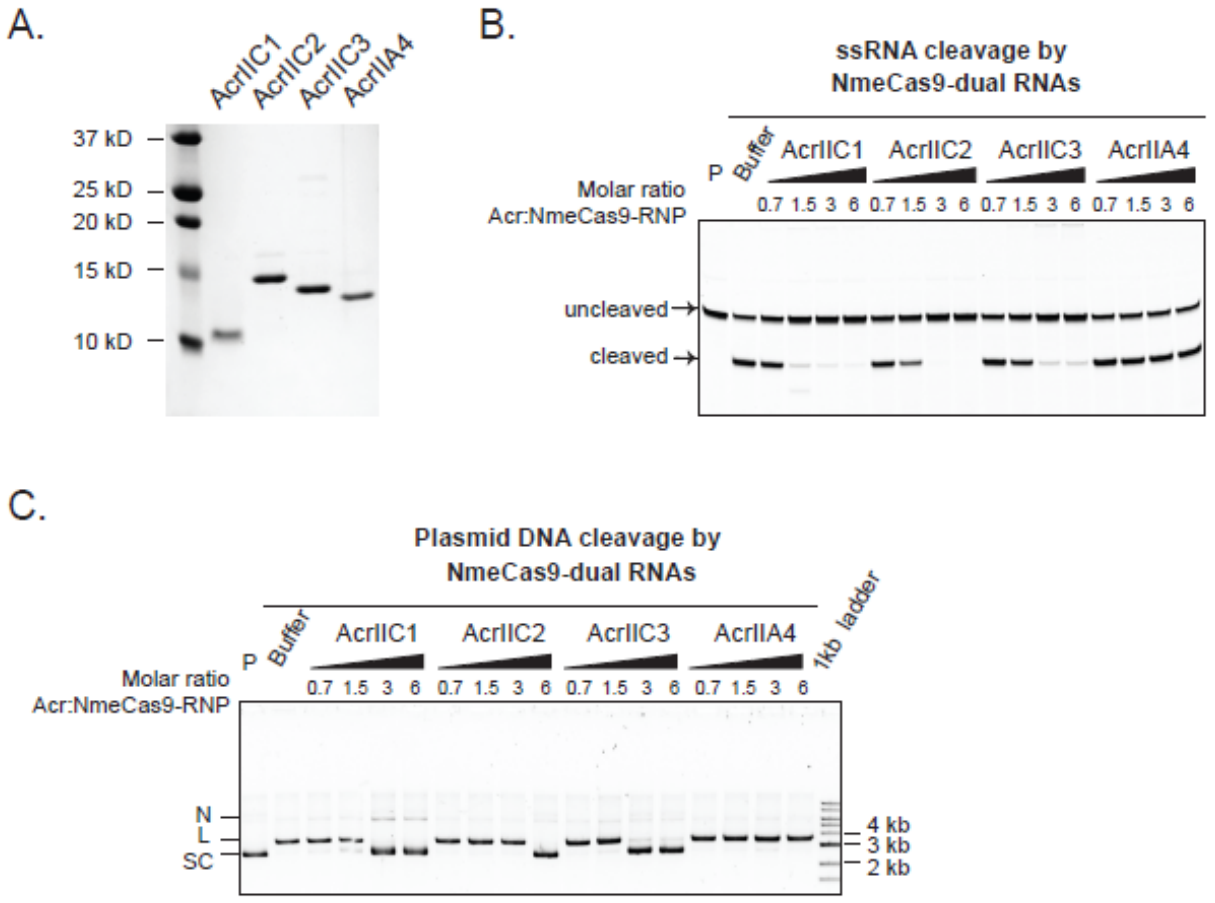

Figure 8 Type II-C anti-CRISPRs block RNA cleavage by NmeCas9.
(A) A Coomassie-stained 15% SDS-PAGE of purified Acr proteins. The predicted molecular weight of AcrIIC1Nme, AcrIIC2Nme, Flag-AcrIIC3Nme, and AcrIIA4 (a Type II-A control Acr) are about 9.8, 14.4, 14.6, and 10.4 kDa, respectively. Molecular weight markers are indicated. (B) Three Type II-C anti-CRISPR proteins AcrIIC1, AcrIIC2, AcrIIC3, but not AcrIIA4, inhibit NmeCas9-catalyzed RNA cleavage. NmeCas9 was incubated with increasing amounts of various Acr proteins for 10 minutes, assembled with the tracrRNA and crRNA for another 10 minutes, and then mixed with fluorescently-labeled RNA substrate (25nM) to license in vitro RNA cleavage. Each Acr protein was used at 0.75, 1.5, 3, and 6 fold molar equivalents relative to NmeCas9 RNP (500nM). The cleavage reactions were analyzed by denaturing PAGE. (C) All three purified Type II-C Acr proteins, but not AcrIIA4, are potent inhibitors for NmeCas9-mediated plasmid DNA cleavage. N, nicked; L, linearized plasmid, about 3.6 kb; SC: supercoiled. Molecular size markers are indicated. Reactions were done as in (B) except that 8 nM of plasmid DNA was added as the substrate, and the reactions were analyzed by agarose gel and SYBR Safe staining.

2.5 RNA-guided, sequence-specific binding of NmeCas9 to RNA targets *in vitro*

To investigate how NmeCas9 engages ssRNA targets, we turned to electrophoretic mobility shift assays (EMSAs). Divalent metals were omitted from the reactions to render NmeCas9 catalytically inactive. We observed robust mobility shifts forming on the fluorescently-labeled RNA target only when NmeCas9, a matching crRNA, and the tracrRNA

were included (Fig. 9A, lane 2 from the left, top bands). Importantly, these shifts were greatly reduced when a non-cognate crRNA was used instead (Fig. 9A, lane 3 from the left), indicative of sequence-specific, stable binding of NmeCas9 RNP to the ssRNA target. There were a few other shifts with intermediate mobility in all the binding reactions containing the cognate crRNA-sp25 (Fig. 9A, lanes 2, 4, 7 and 8 from the left), likely reflecting binding events mediated by RNA-RNA interactions only. The same PAM mutant RNA substrate used in Fig. 6E was also analyzed in EMSA experiments and exhibited no binding defects at all (Fig. 9B), suggesting that NmeCas9 recognizes its ssRNA target in a PAM-independent way. This is in stark contrast to Cas9-mediated dsDNA binding, where PAM recognition is a prerequisite (Sternberg et al. 2014; Jinek et al. 2012).

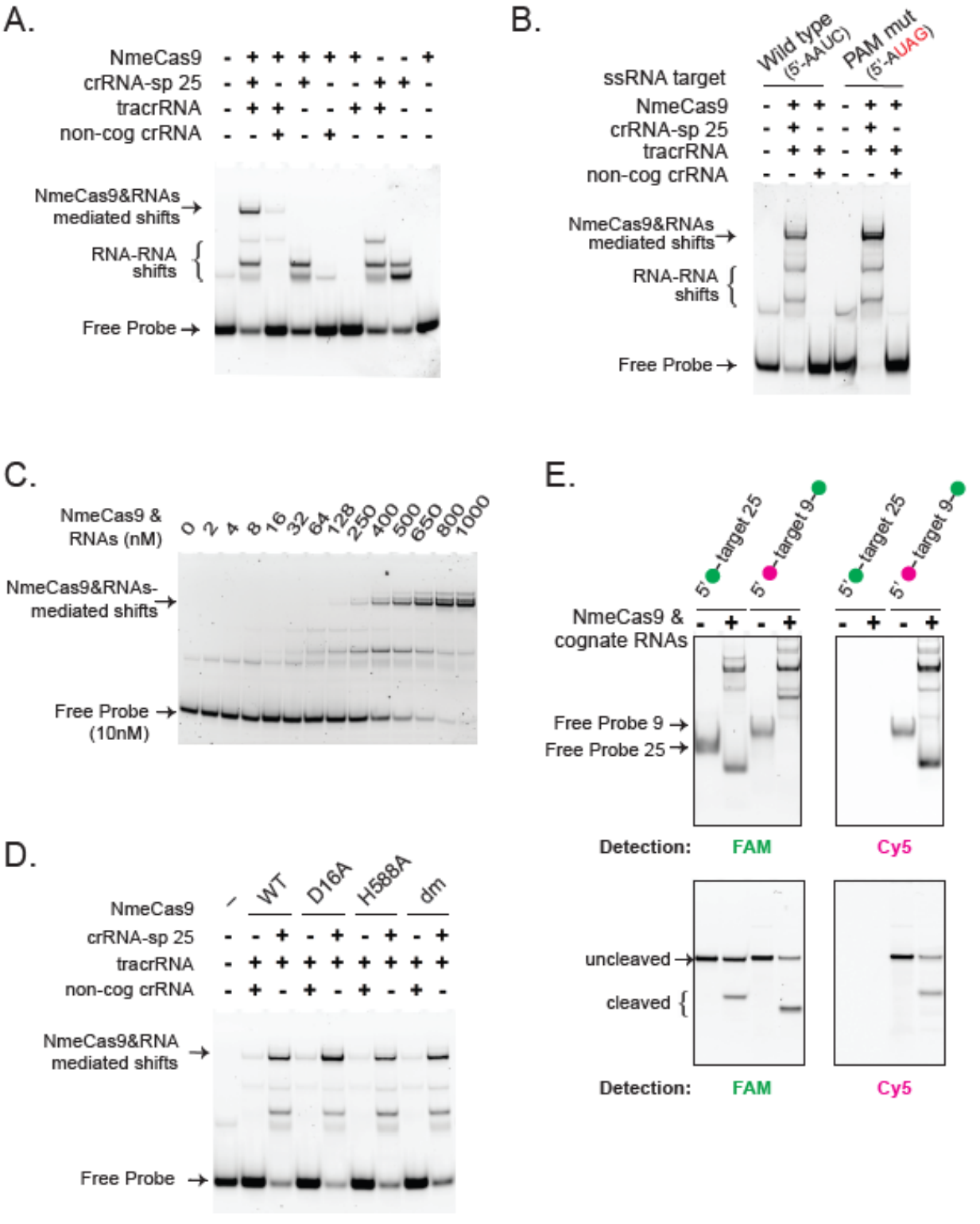
A.
NmeCas9
crRNA-sp 25
tracrRNA
non-cog crRNA
NmeCas9&RNAs mediated shifts
RNA-RNA shifts
Free Probe
B.
Wild type (5'-AAUC)
PAM mut (5'-AUAG)
ssRNA target
NmeCas9
crRNA-sp 25
tracrRNA
non-cog crRNA
NmeCas9&RNAs mediated shifts
RNA-RNA shifts
Free Probe
C.
NmeCas9 & RNAs (nM)
0 2 4 8 16 32 64 128 250 400 500 650 800 1000
NmeCas9&RNAs-mediated shifts
Free Probe (10nM)
D.
NmeCas9
WT
D16A
H588A
dm
crRNA-sp 25
tracrRNA
non-cog crRNA
NmeCas9&RNA mediated shifts
Free Probe
E.
5' target 25
5' target 9
NmeCas9 & cognate RNAs
Free Probe 9
Free Probe 25
Detection:
FAM
Cy5
uncleaved
cleaved
Detection:
FAM
Cy5

Figure 9 dNmeCas9 binds ssRNA target in vitro, in a crRNA-guided manner.
(A) CrRNA-guided, sequence-specific recognition of RNA target by NmeCas9. Electrophoretic mobility shift assays (EMSAs) were performed using the 5' FAM-labeled ssRNA target 25 (25 nM), NmeCas9 (500 nM), and various small RNAs (500 nM) as indicated. Binding was performed in cleavage reaction buffer (but with Mg^{2+} omitted) at 37˚C for half an hour. Non cog, non-cognate crRNA-sp23. (B) The association of NmeCas9 with RNA target is PAM-independent. The same PAM mutant RNA substrate used in Fig. 6E was analyzed here using EMSAs. Binding reactions were done as in (A). Red, mutated sequences in the PAM mutant substrate; non-cog, crRNA-sp23. (C) Dose-dependent binding to the RNA target by NmeCas9-RNP complex. Binding reactions were done as in (A) except that 10 nM substrate is used. The concentrations of NmeCas9 RNP used were indicated. (D) The wild type NmeCas9 and its nuclease domain active site mutants (D16A, H588A, and dm) were assayed for sequence-specific binding to the RNA target, as in (A). Dm, double mutant; non-cog, crRNA-sp23. (E) The 3' but not 5' product of RNA cleavage remained bound by the NmeCas9 RNP. NmeCas9-catalyzed RNA cleavage reactions (with Mg^{2+} added) were done as in Fig. 6C, for two separate fluorescent RNA substrates, target 25 and target 9. Half of the reactions were analysed by 6% native PAGE (upper panel), and the same native gel was visualized by FAM (left) and Cy5 (right) detection. The other half of the reactions were quenched by proteinase K treatment and formamide loading dye, then analyzed by 15% denaturing PAGE (lower panel), and the same denaturing gel was visualized by FAM (left) or Cy5 (right) detection. Schematics of RNA substrates used were at the top. Green, FAM label; magenta, Cy5 label. Note that the 5' cleavage product dissociates from the NmeCas9 (upper right panel).

The NmeCas9-mediated retarded migration in EMSAs occurred in a dose-dependent manner (Fig. 9C), and the binding was not affected by active site mutations in either the HNH or the RuvC domains (Fig. 9D), suggesting that the recognition of RNA substrate can happen independently of RNA scission. Since SpyCas9 still holds on to all four ends of the cleavage products after cleaving the DNA duplex (Sternberg et al. 2014; Richardson et al. 2016) we investigated whether the RNA cleavage products here are released from the NmeCas9 RNP. To this end, we performed standard *in vitro* RNA cleavage assay and analyzed the reactions using two parallel approaches. Half of reactions were quenched and analyzed by denaturing PAGE (Fig. 9E, lower), and the other half was analyzed by native PAGE (Fig. 9E, upper). While robust cleavage was observed on both ssRNA target 25 and target 9 (Fig. 9E, lower), the 5' cleavage products (FAM-labeled on target 25, Cy5-labeled on target 9) were released from the NmeCas9 RNP (Fig. 9E, upper). In contrast, the 3' cleavage products (FAM-labeled on target 9), which contain 20 nts of sequence complementary to the cognate crRNA, were present only in higher molecular weight shifts on native PAGE, suggesting that they were largely still bound by the NmeCas9 RNP (Fig. 9D, upper panel, left).

2.6 Expanded targeting capacities for CRISPR-Cas9 systems

Central to the CRISPR genome editing revolution are the Cas9 DNA endonucleases, which can be easily programmed to cut any dsDNA target of interest through PAM recognition and guide-target base pairing (Wright, Nuñez, and Doudna 2016). Our work here revealed that the Type II-C Cas9 from *N. meningitidis* is capable of programmable, RNA-guided, sequence-specific cleavage and recognition of ssRNA targets (Figs. 6, 7 and 9). Importantly, unlike the Type II-A SpyCas9 (O'Connell et al. 2014), NmeCas9's RNase activity is independent of PAMmer DNA oligo auxiliary factor (Figs. 6C-D). Therefore, to our knowledge, NmeCas9 represents the first native Cas9 endoribonuclease and expands the targeting capacities of CRISPR-Cas9.

The PAM- and PAMmer-independent nature of this ribonuclease activity implies that the selection of ssRNA substrate is achieved mainly through RNA-RNA pairing between the guide and the target, without a requirement for PAM recognition by NmeCas9. This feature may help enable the manipulation of cellular messenger RNAs while avoiding collateral cleavage or binding to the corresponding genomic sites lacking the PAMs. Moreover, the same HNH nuclease domain of NmeCas9 that cleaves the target DNA strand also cleaves the ssRNA substrate. This is not surprising given the existence of HNH motifs in homing endonucleases that can cleave both DNA and RNA molecules (Pommer et al. 2001). Future structural studies are needed to understand how the NmeCas9 and its HNH domain accommodate various kinds of nucleic acid targets. Finally, it is tempting to speculate that additional natural Cas9 RNase may exist in the divergent Cas9 family.

2.7 Potential RNA-targeting applications

RNA plays critical and diverse biological roles, and RNA-targeting tools have the potential to transform research and medicine. There are existing platforms such as RNA interference, antisense oligonucleotides and designer RNA-binding proteins (e.g. Pumilios) that can recognize specific transcripts or exogenous RNA tags. (Nelles et al. 2015) Recently, the CRISPR-Cas13 system that exclusively targets RNA was repurposed as a new tool to knockdown or edit specific mammalian transcripts (Abudayyeh et al. 2017; Cox et al. 2017), and dSpyCas9 can also help to remove repetitive RNAs in human cells (Batra et al. 2017). These programmable RNA-targeting tools will revolutionize how we modulate RNA metabolism in the cells.

NmeCas9 can potentially provide a unique single platform to simultaneously target both dsDNA and RNA. The PAMmer-independent nature of its RNase activity makes NmeCas9 a desired system to circumvent challenges that come with the delivery and toxicity of ssDNA PAMmer oligos. In addition, the fact that NmeCas9, SpyCas9 and Cas13s use orthogonal guide RNAs (Abudayyeh et al. 2016; Esvelt et al. 2013) can be exploited to achieve distinct but multiplexed RNA-targeting tasks. Moreover, the three families of Type II-C Acr proteins can provide off-switches to enable precise control of NmeCas9-based RNA cleavage applications.

2.8 Biological Implications of NmeCas9 RNA targeting

Species of the genus *Neisseria* rely on natural transformation for frequent genetic exchange (Hamilton and Dillard 2006). In CRISPR-containing *N. meningitidis* strains, dsDNA cleavage by Cas9 provides a barrier to genetic material transfer through natural transformation (Y. Zhang et al. 2013). NmeCas9, along with several other Type II-C Cas9s, are also found to be capable of robust ssDNA cleavage *in vitro*, a feature proposed to play an evolutionarily ancestral

role in restricting the ssDNA genome of filamentous bacteriophages or transforming ssDNAs (Ma et al. 2015; Y. Zhang et al. 2015).

As for the RNase activity of NmeCas9, the physiological relevance in *Neisseria* cells remains to be determined. One possibility is that RNA-targeting plays an auxiliary role in defense by helping the clearance of transcripts derived from transforming DNAs or bacteriophages. Another possibility is that NmeCas9 may target endogenous *Neisseria* RNAs in a crRNA-guided fashion. We bioinformatically examined potential matches between all the twenty-five spacers from native mature crRNAs and the self-chromosome of *N. meningitidis* strain 8013. BLASTN searches revealed no perfect or near-perfect hits (up to 4 nts mismatches allowed), consistent with an earlier finding that *Neisseria* CRISPR spacers primarily match to the genomes of other strains or species (Y. Zhang et al. 2013). The best matches for all but one crRNA have 7 or more nts of mismatches spread out in regions both proximal and distal to the presumed RNA cleavage site. This degree of mispairing, according to our *in vitro* mutagenesis study (Fig. 7C), would largely abolish NmeCas9-catalyzed RNA cleavage (Rousseau et al. 2018). Therefore, we cautiously speculate that robust cleavage of endogenous *Neisseria* transcripts by NmeCas9 and existing CRISPR spacers is not very likely, but could arise. Nonetheless, future work is needed to determine if crRNA-directed RNA binding or cleavage of endogenous transcripts by NmeCas9 occurs in *Neisseria*, and if so, what biological consequences could result.

Chapter 3 Development of Virus-Like Particles Derived from the Moloney Murine Leukemia Virus for the Delivery of Genome Editing Enzymes

This chapter consists of unpublished data. Beth Rousseau and David Turner planned experiments. Beth Rousseau conducted experiments. Huanqing Zhang assisted with confocal microscopy imaging; genotyping and collection of the mice for the mouse dermal fibroblasts; The mouse KRAS guide RNA was designed by Anne Vojtek; the KRAS sgRNA vector was constructed and validated by Huanqing Zhang.

3.1 Introduction

Virus-like particles (VLPs) are generated using the structural proteins from viruses, but without packaging the viral genome. They have been used for decades in vaccine development and as vehicles for the study of the viral life-cycle without the risks associated with a replication competent virus (Nooraei et al. 2021). More recently, VLPs have been used as vehicles for protein transduction (Mangeot et al. 2019; D.-T. Wu and Roth 2014; Bobis-Wozowicz et al. 2015). Proteins that act as genome editors, transcription modulators, or apoptosis inducers do not need to be expressed continuously for their effects to have lasting consequences (D.-T. Wu and Roth 2014). Their transient introduction at low concentration into cells can be sufficient for function, and transient protein delivery can have advantages over longer lasting nucleic acid mediated protein expression. In some cases, off-target effects from a protein may increase based on cumulative expression (e.g. for the genome editing enzyme Cas9 (Vakulskas and Behlke 2019)). Transient protein delivery via VLPs may reduce off-target effects relative to persistent expression from nucleic acid or viral vector-based delivery.

VLPs designed for transient protein delivery have been developed based on the Murine Leukemia Virus (Mangeot et al. 2019; D.-T. Wu and Roth 2014; Voelkel et al. 2010; Bobis-

Wozowicz et al. 2015). The MLV produces three proteins late in the viral life cycle, Gag, GagPol, and an envelope protein (Goff and Lobel 1987). Gag is the structural polyprotein of the MLV (Rein 2011). It is capable of forming particles that bud from the cell membrane when expressed alone. It is composed of four domains, matrix, P12, capsid, and nucleocapsid. Cargo proteins, such as transcription factors, fluorescent proteins, or genome editors such as Cre and Cas9 can be loaded into VLPs concomitantly via fusion to Gag. Gag and fusion proteins can be separated by proteolysis after the particles bud from the cell membrane if a viral protease site is included at the fusion junction and the viral protease naturally occurring in Pol is included in the design of the VLPs. GagPol is a polyprotein composed of Gag followed by the Pol domain consisting of protease, reverse transcriptase (RT), and integrase (IN). Gag and GagPol are translated from the same mRNA transcript. The stop TAG codon at the end of Gag is followed by a pseudoknot that encourages readthrough ~95% of the time so that there is 20-fold less GagPol in host cells than Gag, and this imbalance is also reflected in the composition of virions (Johnson et al. 2014). The ecotropic envelope (Eco Env) protein native to the MLV enables infection of rodent cells. For the purposes of engineered particles, expression of other envelope proteins (Mangeot et al. 2019) can effectively pseudotype MLV VLPs to allow infection of cells from humans or other non-rodent species.

Genome editing via *S. pyogenes* Cas9 (SpyCas9) has transitioned from discovery to ubiquitous commercial availability in under a decade (Jinek et al. 2012; 2013; Rees et al. 2021). The popularity of SpyCas9 is owed to the ease of reprograming its guide RNA to target any genomic locus adjacent to the motif 5’-NGG-3’ (Jinek et al. 2012). Cas9 catalyzes a double stranded break within the sequence specified by the guide RNA. Endogenous repair pathways lead to either imprecise or precise repair of the locus. A pathway called Non-Homologous End

Joining (NHEJ) sutures the broken ends of DNA together in a way that often causes insertions or deletions (indels) (Chang et al. 2017). Homology-dependent repair (HDR) is a more precise repair pathway that performs templated repair, often using a sister-chromosome as a template (H. Yang et al. 2020). Alternately, a programmed DNA repair template supplied in trans can be used by the HDR pathway to introduce a specific mutation or other sequence change, although the rate of precise repair is often far lower than the rate of Cas9 cleavage. Cell type and/or growth state can modulate the availability of the HDR or NHEJ repair mechanisms (Chang et al. 2017). In addition, a low rate of incorporation of the DNA repair template during HDR is partly explained by poor spatial and temporal co-localization of the repair template and the necessary repair enzymes at the site of Cas9 cleavage (H. Yang et al. 2020).

Prime editing was developed to overcome the challenge of localizing an HDR repair template to the site of Cas9 cleavage (Anzalone et al. 2019). A primer binding site and a template with a desired edit are encoded in a prime editing guide RNA (pegRNA). The pegRNA is bound by a Cas9 nickase (nCas9) and directed to a genomic locus. Cas9 binds, separates the DNA into two strands and cleaves only the strand of DNA not paired to the 5' end of the pegRNA. The nicked DNA binds to the primer binding site at the 3' end of the pegRNA. Tethered to Cas9 by a peptide linker is the reverse transcriptase (RT) enzyme that originates from the Moloney MLV (MMLV). RT polymerizes from the nicked DNA strand as directed by the template encoded in the pegRNA. This introduces an edit and a homology arm directly into the genome. Endogenous repair enzymes will incorporate or exclude the flap of homologous DNA containing the edit. If the newly synthesized strand is incorporated into the genome, the editing event is successful.

Here we describe VLPs, based on MMLV, that can deliver Cre or Cas9 to efficiently edit genes in mouse dermal fibroblasts and HEK293T cells. Our findings provide new insights into the formation of the VLPs and considerations for the design of future VLPs. The most important feature of the VLPs was codon optimization of the MMLV Gag protein to eliminate naturally occurring sequences detrimental to protein production in HEK293T cells. Comparisons of different protease sites separating Gag and Cas9 revealed that the protease site occurring naturally at the end of the Gag polyprotein yielded reasonably efficient VLPs. Finally, we found that truncation of the GagPol protein to eliminate integrase (IN) and point mutations to eliminate the function of reverse transcriptase (RT) also made VLPs nearly as effective as the wild-type GagPol without risking the unintended enzymatic activities of IN or RT. All of the optimizations on Cas9 VLPs were combined when designing prime editing VLPs, which we show can be used to introduce an epitope tag into a cellular protein in either HEK293T cells or NIH3T3 cells.

3.2 Results

3.2.1 Generation of VLPs for Cas9 Delivery

To generate VLPs based on MMLV that could deliver Cas9 or other genome editing enzymes, we initially used a design similar to the previously described “Nanoblades”, which are based on the related Friend MLV (Mangeot et al. 2019). We constructed a plasmid vector that expressed either Cre or the SpyCas9 protein, hereafter simply referred to as Cas9, fused to the C-terminus of the MMLV Gag protein. Between Gag and Cas9 we incorporated the native protease site at the end of the nucleocapsid protein and a short, flexible GASGAM peptide linker. There are three copies of the FLAG tag and an SV40NLS (nuclear localization signal) between the peptide linker and the beginning of the Cas9 protein. To produce Cas9 VLPs, HEK293T cells were transiently transfected with an expression vector for GagCas9 in combination with

expression vectors for an sgRNA, the MMLV Gag/GagPol proteins, and the VSV-g envelope protein (Fig. 10A). Alternately to produce Cre VLPs, the Gag-Cre vector was transfected in combination with expression vectors for the Gag/GagPol proteins, and the VSV-g expression vectors. The transfected cells released VLPs into their supernatants which were collected, in some cases concentrated, and applied to target cells (see Methods for additional information).

In order to facilitate functional testing of the VLPs, we generated an immortalized mouse dermal fibroblast (MDF) cell line derived from the Ai9:Ribotag mouse (See Methods)(Madisen et al. 2010; Sanz et al. 2009) that we call the Ai9 MDF cell line. The cell line harbors the Ai9 tdTomato fluorescent reporter in the *Rosa26* locus that can be activated either by Cre recombinase or by Cas9 with the Ai9 single guide RNA (sgRNA) (Fig. 10B) (Staahl et al. 2017). The CMV promoter is separated from tdTomato by two LoxP sites that flank a stop cassette comprised of three repeats of the SV40 polyadenylation site. Cre activates the reporter by removing the stop cassette entirely, while Cas9 with the Ai9 guide activates the reporter by deletion of two of the three SV40 polyadenylation sites, after cleaving at repeated recognition sites within the stop cassette (Staahl et al. 2017).

We transiently transfected HEK293T producer cells to generate Cas9 VLPs with the Ai9 sgRNA, which were then used to infect the Ai9 MDF cell line. Three days after infection, MDFs infected with Cas9 VLPs were fixed and imaged for tdTomato expression (Fig. 10C). VLPs containing Cas9 loaded with the Ai9 sgRNA were able to activate the tdTomato reporter. Cells with the activated reporter (i.e. tdTomato^{+}) also were quantified by flow cytometry (Fig. 10C). To determine the functional titers of VLPs expressing Cas9 with the Ai9 sgRNA or Cre, the Ai9 MDF cells were infected with VLPs, then the numbers of tdTomato expressing cells were divided by the volume of VLP supernatant applied to cells (infectious units per milliliter).

However, due to the need for two or three cuts to delete the stop cassette in the Ai9 reporter, as well as multiple possible repair outcomes, some of which do not lead to deletion of two copies of the SV40 polyA site, Cas9 with the Ai9 sgRNA cannot activate the reporter in every cell (Staahl et al. 2017). This limitation on reporter activation by Cas9, coupled with the fact that the MDF line is heterozygous for the Ai9 reporter, means that Cas9 VLP titers determined with the MDF cell line are conservative estimates. In addition to the Ai9 reporter, the Ai9 MDF cell line is also homozygous for the RiboTag reporter at the Rpl22 locus, which can be activated by Cre (Sanz et al. 2009).

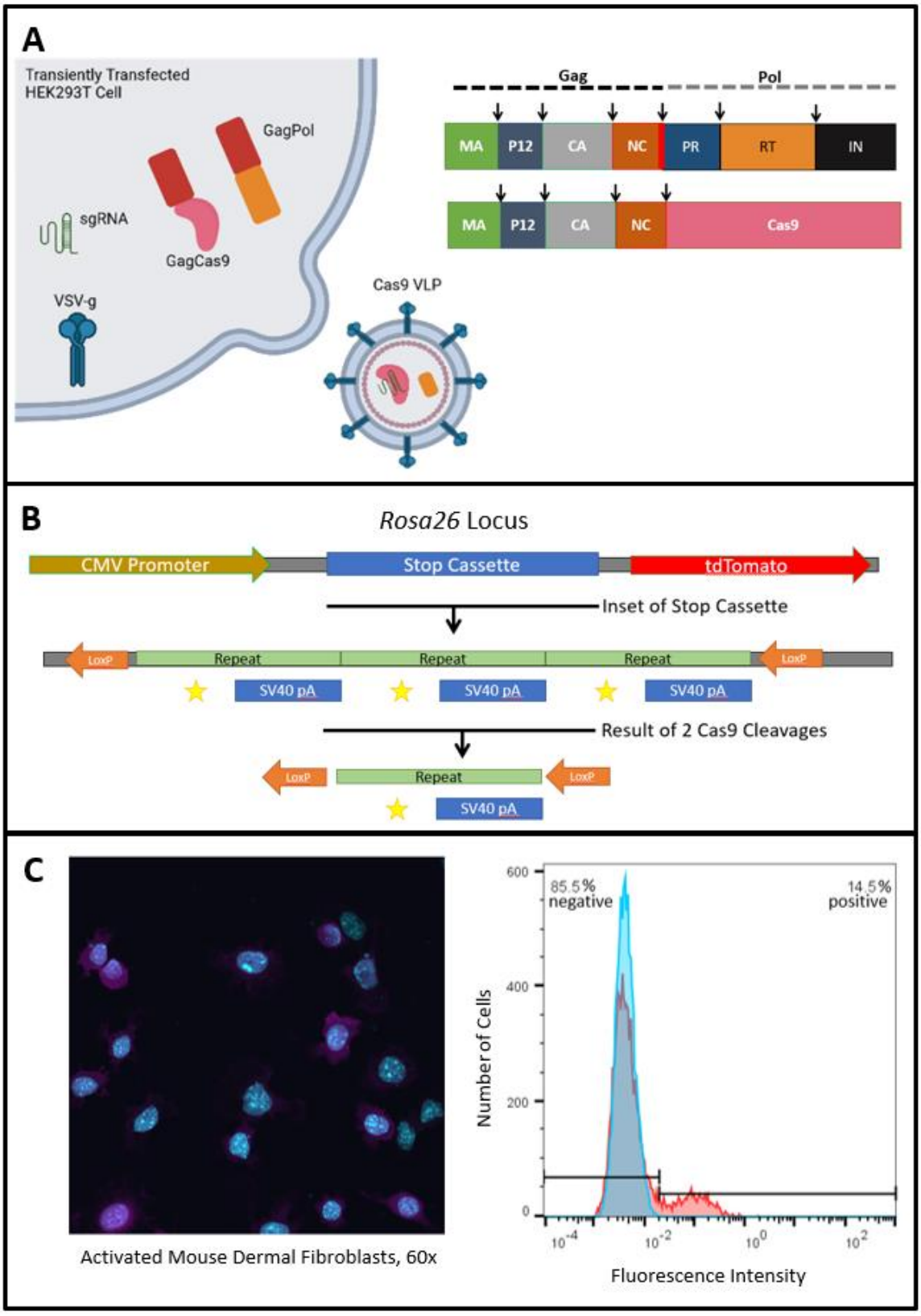
A
Transiently Transfected
HEK293T Cell
GagPol
sgRNA
GagCas9
VSV-g
Cas9 VLP
Gag
Pol
MA
P12
CA
NC
PR
RT
IN
MA
P12
CA
NC
Cas9
B
Rosa26 Locus
CMV Promoter
Stop Cassette
tdTomato
Inset of Stop Cassette
LoxP
Repeat
Repeat
Repeat
LoxP
SV40 pA
SV40 pA
SV40 pA
Result of 2 Cas9 Cleavages
LoxP
Repeat
LoxP
SV40 pA
C
Activated Mouse Dermal Fibroblasts, 60x
85.5%
negative
14.5%
positive
600
400
200
0
Number of Cells
10^-4
10^-2
10^0
10^2
Fluorescence Intensity

Figure 10 Cas9 VLPs activate a reporter in the Ai9 MDF cell line.
A) The Rosa26 locus harbors a gene that expresses the fluorescent protein tdTomato which can be activated by either Cre or Cas9 with a guide RNA. The cleavage sites for Cas9 are indicated by the stars. B) HEK293T cells are transiently transfected with plasmid DNA expressing GagCas9, GagPol, VSV-g, and a guide RNA. The proteins self-assemble into VLPs which bud into the supernatant. C) Right: Ai9 MDF cells express tdTomato when activated. tdTomato is pseudocolored magenta. The nucleus has been stained with DAPI which appears blue. Left: The blue peak represents MDFs that have not been exposed to Cas9. The red peaks are a population of Ai9 MDF cells that have been treated with Cas9 VLPs at a subsaturating concentration. The fluorescence intensity is measured in the 615/20 nm channel using the Macsquant VYB. 10A was created with Biorender.com.

3.2.2 Gag Codon Optimization Improves Cas9 VLP Titer

Cas9 VLPs generated with the MMLV GagCas9 fusion and the Ai9 sgRNA had titers of ~$1x10^3$ infectious units per mL. Of several optimizations tested, Gag codon optimization had the greatest effect on titer (Fig. 11A, Chapter 4.2.2, 4.2.5, 4.2.6, and data not shown). Prior published reports indicated that RNA sequences within the Gag coding region can inhibit Gag protein production (Bartels and Luban 2014). We therefore created a codon optimized Gag protein for fusion (GagCO), with an optimized Kozak consensus at the initiator methionine. Expressing the GagCas9 fusion protein from the vector containing GagCO increased VLP titer approximately ten-fold in comparison to the original vector. A western blot revealed a 2.5-fold increase in the amount of GagCas9 production (Fig. 11B). The calculated mass of GagCas9 is 226 kDa. Furthermore, the size of the predominant GagCas9 protein generated from the GagCO-Cas9 fusion is larger than the predominant protein produced from the original sequence. The protein may be longer due to increased translation starting from the first methionine in Gag, rather than the second methionine which would truncate the protein by 21 kDa. The codon-optimized Gag, GagCO, was used in the creation of all further constructs.

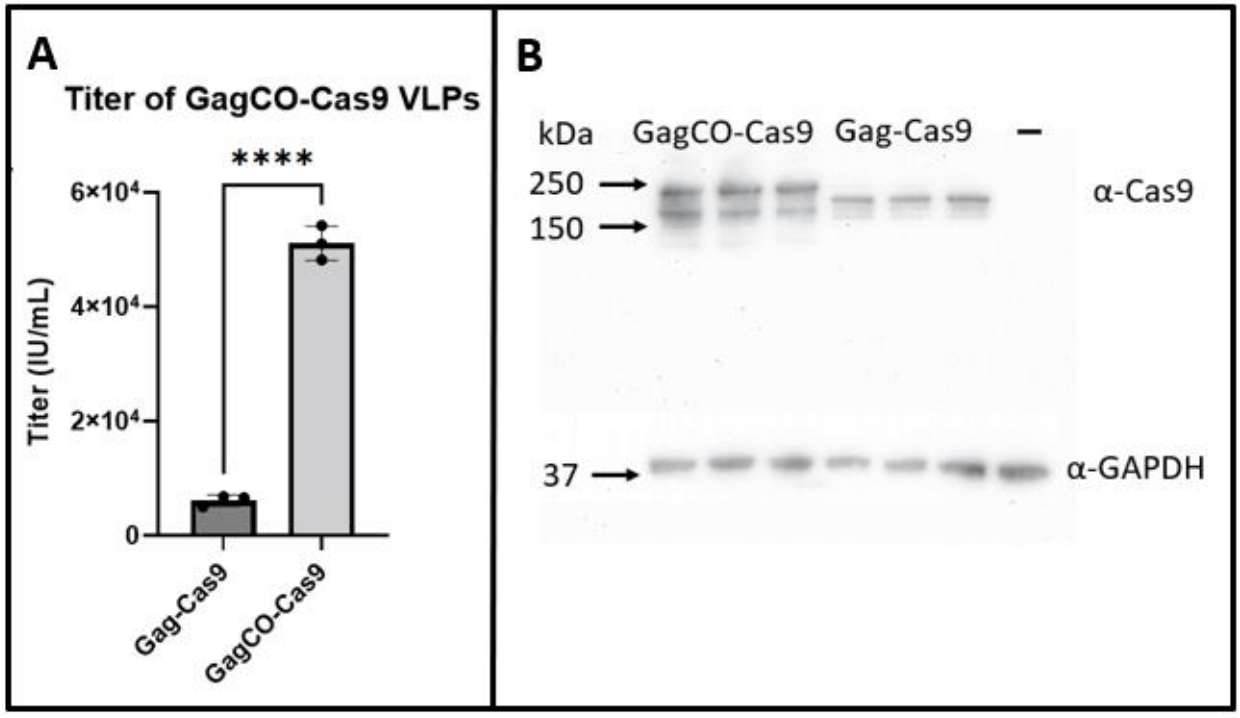

Figure 11 Codon optimization significantly improves Cas9 VLP titer.
A) Titers of Cas9 VLPs improved by approximately ten-fold when measured by activating the Ai9 reporter in MDFs. B) A western blot of the cell lysates of HEK293T cells transfected with the plasmids expressing the Ai9 sgRNA and either GagCO-Cas9 or GagCas9 shows that protein production increased by approximately 2.5-fold. The size of the protein increased, possibly due a change in start site. Untransfected cells were used as the negative control (-).

Both VSV-g and Eco Env are envelopes known to support VLP infection of mouse cells (Voelkel et al. 2010; Bobis-Wozowicz et al. 2015). We reasoned that it might be possible to improve VLP titer by combining both types of envelopes, as has been reported previously for other combinations of envelope proteins (Mangeot et al. 2019). VLP binding and fusion between target cells and VLPs mediated by cellular receptors is inherently limited by the number of receptors available on the cell surface. If one type of receptor becomes saturated, an envelope protein targeting a different receptor may still be able to increase the number of VLPs that bind/fuse to cells during the period of infection. We made our VLPs as stated above either without the addition of an envelope expressing vector, with each envelop expressing vector individually, or by adding both envelopes into the transfection. We found that the addition of

Eco Env to the VLPs did not improve the titer beyond the titer provided by VSV-g alone (Fig. 12A). All subsequent VLP experiments were performed with only the VSV-g envelope protein.

Due to the fact that the Ai9 MDF cell line is heterozygous for the reporter and needs multiple cleavage events in order to create a deletion that will activate the tdTomato reporter, it will underestimate the efficacy of Cas9 VLPs. We chose to make Cas9 VLPs targeting various endogenous mouse genes in order to better assess the efficiency of Cas9-mediated gene disruption in the Ai9 MDF cell line where there are presumably two copies of the endogenous gene. We designed a guide RNA against *Wtap*, which encodes Wilms' tumor 1-associating protein, a pre-mRNA-splicing regulator that is indispensable for embryonic development (Fukusumi, Naruse, and Asano 2008) and a guide RNA against *Vim*, encoding Vimentin, an intermediate filament protein expressed in many different cell types (Danielsson et al. 2018). We also tested a guide RNA against *Kras*, encoding K-RAS, a GTPase that is important in cell signaling and the cause of many cancers (Herdeis et al. 2021). Finally, we targeted the 3' end of histone H1C, encoding a histone that is widely expressed and could potentially be modified with a donor DNA template supplied in trans or by prime editing (Wang et al. 1997; X. Zhang et al. 2016). VLPs with each of the different guide RNAs were made as described above and used to infect Ai9 MDF cells. Cells were collected two days post-infection, and the targeted genomic locus was amplified by PCR. Gene disruption was measured by scoring changes in the genomic target sequence using Tracking of Indels by DEcomposition (TIDE) (Brinkman et al. 2014). This computational technique compares sequences at the target site from unedited genomic DNA to the Cas9 edited genomic DNA to determine the frequency that indels appear in the edited population. The efficiency of the GagCO-Cas9 VLPs in disrupting selected gene targets in MDFs as measured by TIDE varied between 55% and 83% (Fig. 12B). VLPs targeting *Vim* were

concentrated and used to infect MDFs, but VLP concentration did not improve *Vim* cleavage efficiency.

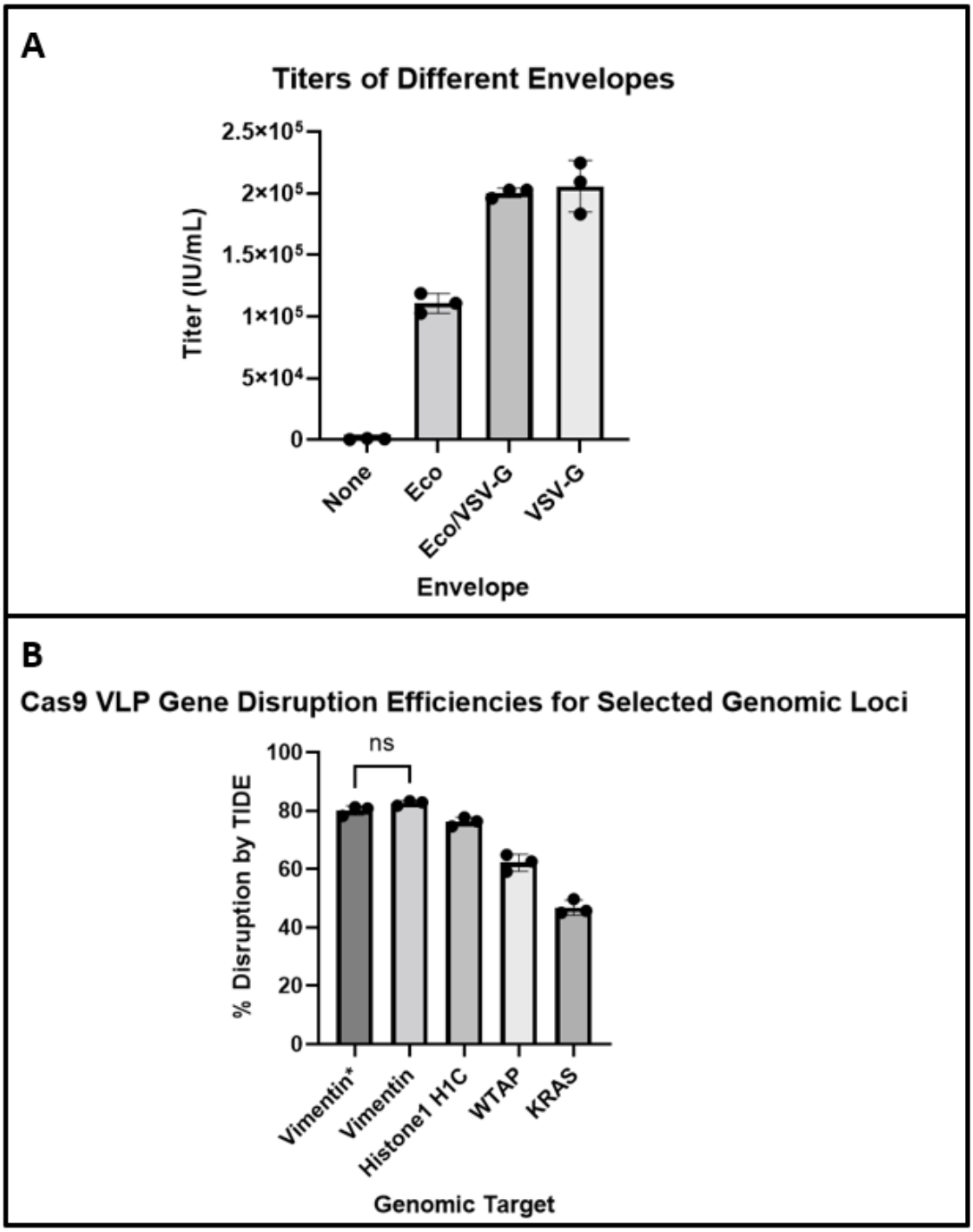

Figure 12 Cas9 VLPs are efficient.
A) Without an envelope, Cas9 VLPs have almost no ability to activate the Ai9 reporter. The Eco Env permits a Cas9 VLP titer of 1x10^5 IU/mL while the VSV-G envelope enhances the titer again by a factor of two. Pseudotyping with both envelopes does not appear to improve titer beyond that of VSV-g alone. B) As measured by TIDE, Cas9 efficiently edits four different genomic loci. Vimentin indicates that the Cas9 VLP supernatant was concentrated 3-fold, but it did not improve the editing efficiency beyond the unconcentrated supernatant.*

3.2.3 Integrase and Reverse Transcriptase Are Not Necessary for Cas9 VLP Function

The MMLV GagPol polyprotein provides the viral protease which can cleave the cargo protein from Gag. GagPol also contains reverse transcriptase (RT) and integrase (IN), two enzymes that are necessary to the lifecycle of the MMLV (Rein 2011). However, neither RT nor IN are necessary for viral particle formation, which Gag performs spontaneously (Johnson et al. 2014), suggesting that that these proteins may be dispensable in MMLV VLPs that deliver Cas9 or other proteins. We tested our hypothesis by creating Cas9 VLPs loaded with the Ai9 sgRNA using truncated GagPol that does not contain RT or IN (Fig. 13A) and infecting MDF cells. Initial truncated GagPol designs were based on GagCO, which is the codon optimized Gag. The GagCO-Prot vector expresses the codon optimized Gag with a glutamine in place of the naturally occurring TAG stop codon followed by the protease domain ending with a stop codon, allowing production of a GagCO-Prot fusion protein without RT or IN. The GagCO-Stop-Prot vector expresses the codon optimized Gag coding region, ending with the naturally occurring stop codon TAG, followed by the same protease coding region. The protease is not codon optimized which means that the pseudoknot contained within the protease coding region (Feng et al. 1992) is preserved in the RNA, and is expected to allow read-through approximately 5% of the time, as occurs in translation of the wild-type GagPol mRNA, although we did not confirm the read-through efficiency. We produced VLPs with the GagCO-Prot vector by replacing the GagPol vector in the HEK293T transfections with a mixture of the GagCO and GagCO-Prot vectors (20:1 molar ratio) and tested activation of the tdTomato reporter in the Ai9 MDF line. The ratio of the Gag expressing plasmids was intended to mimic the production of Gag-Pol in which the full-length protein containing the protease is translated at a rate 20-fold lower than Gag. We found that the Gag-protease VLPs did not recapitulate GagPol and the titer (~3.5x10^3) was

virtually indistinguishable from VLPs made with GagCO alone (~2.8x10^3) (Fig. 13B). We also tested the GagCO-Stop-Prot vector for VLP production similarly. As this vector was expected to more closely recapitulate the normal production of GagPol proteins, we used Gag-Stop-Prot by itself in the same amount as the GagPol vector is transfected. Although the GagCO-Stop-Prot vector generated VLPs considerably more efficiently than GagCO alone, GagCO-Stop-Prot VLPs titers still decreased by two-fold compared to Cas9 VLPs made with GagPol (Fig. 13C).

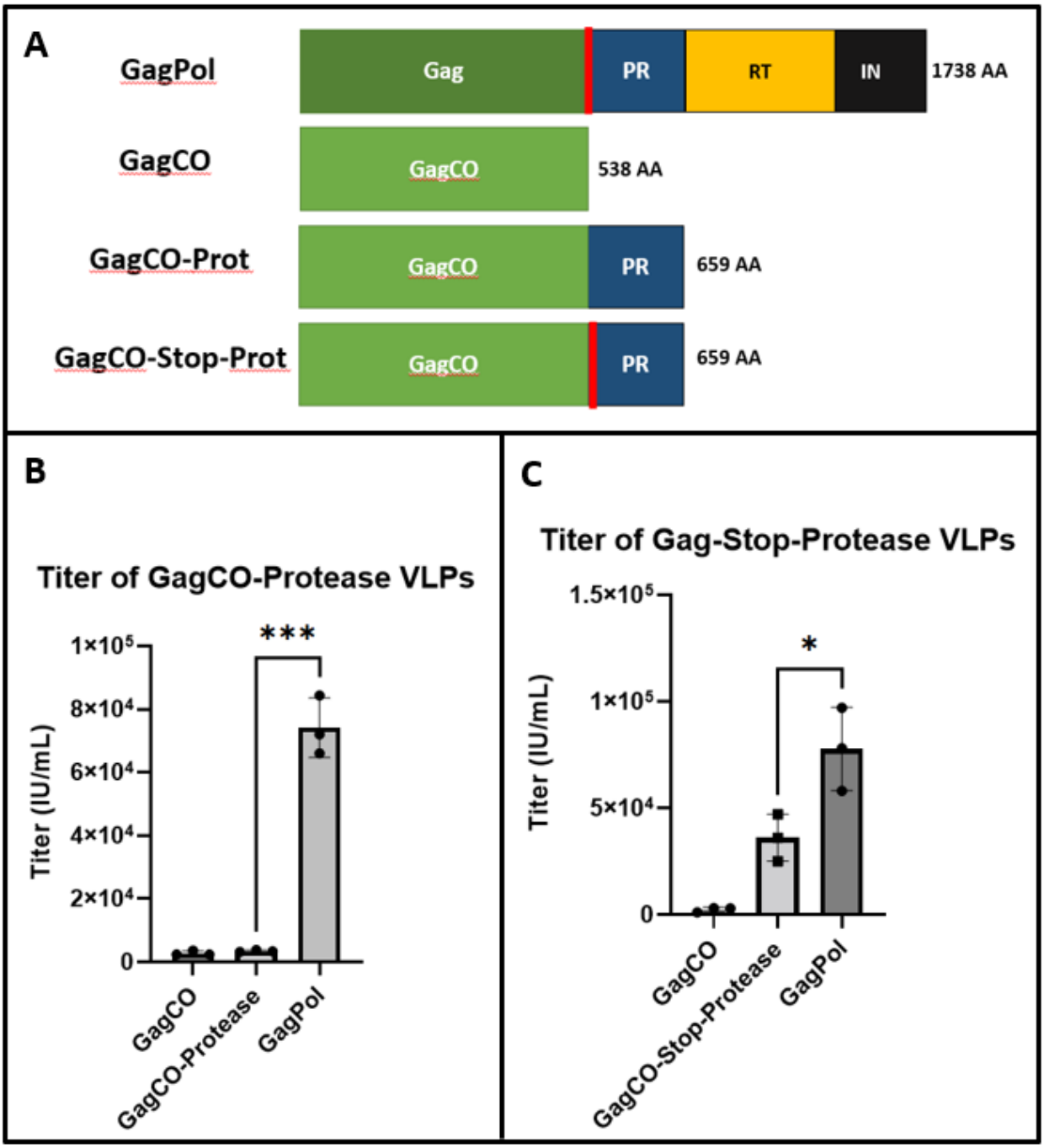

Figure 13 Testing Protease Only Cas9 VLPs.
A) Three different constructs were used as substitutes for GagPol, either alone or in combination. GagCO expressed only Gag. Gag-Prot expressed only Gag-Prot as it did not have a stop codon at the end of Gag. Gag-Stop-Prot retained the stop codon, indicated by the red line, and was intended to express Gag and Gag-Prot in a 95:5 ratio like GagPol. B) Cas9 VLPs created with a combination of GagCO and GagCO-Prot plasmids did not recapitulate VLPs created with GagPol. C) Cas9 VLPs created with the GagCO-Stop-Protease had titers approximately two-fold less than those created with GagPol.

We created two additional constructs derived directly from the original GagPol vector, GagPolIN$^-$ and GagPolRT$^-$IN$^-$, which contain the wild-type Gag coding region (not codon optimized). GagPolIN$^-$ bears an IN truncated to a small peptide, and part of the RNase H domain of RT is also deleted. GagPolRT$^-$IN$^-$ is the same as GagPolIN$^-$ except that the RT protein contains point mutations D224N and D225N (Telesnitsky and Goff 1993) that render it enzymatically inactive (Fig. 14A). The titer for GagRT$^-$IN$^-$ VLPs was similar to the titer for GagPol VLPs as determined by activation of the fluorescent reporter in the Ai9 MDF line (Fig. 14B). The titer for GagPolRT$^-$IN$^-$ VLPs is approximately two-fold lower than VLPs generated with the wild-type GagPol, as evaluated by reporter activation in Ai9 MDF cell line (Fig. 14C). We also infected MDF cells with Cas9 VLPs targeting the endogenous *Smad4* gene as a second method of evaluating the efficacy of the GagPolRT$^-$IN$^-$ VLPs. The cleavage efficiencies of *Smad4* measured by TIDE confirmed that there is an approximately 2.5-fold drop in efficacy for VLPs created with GagPolRT$^-$IN$^-$ relative to those created with GagPol (Fig. 14D).

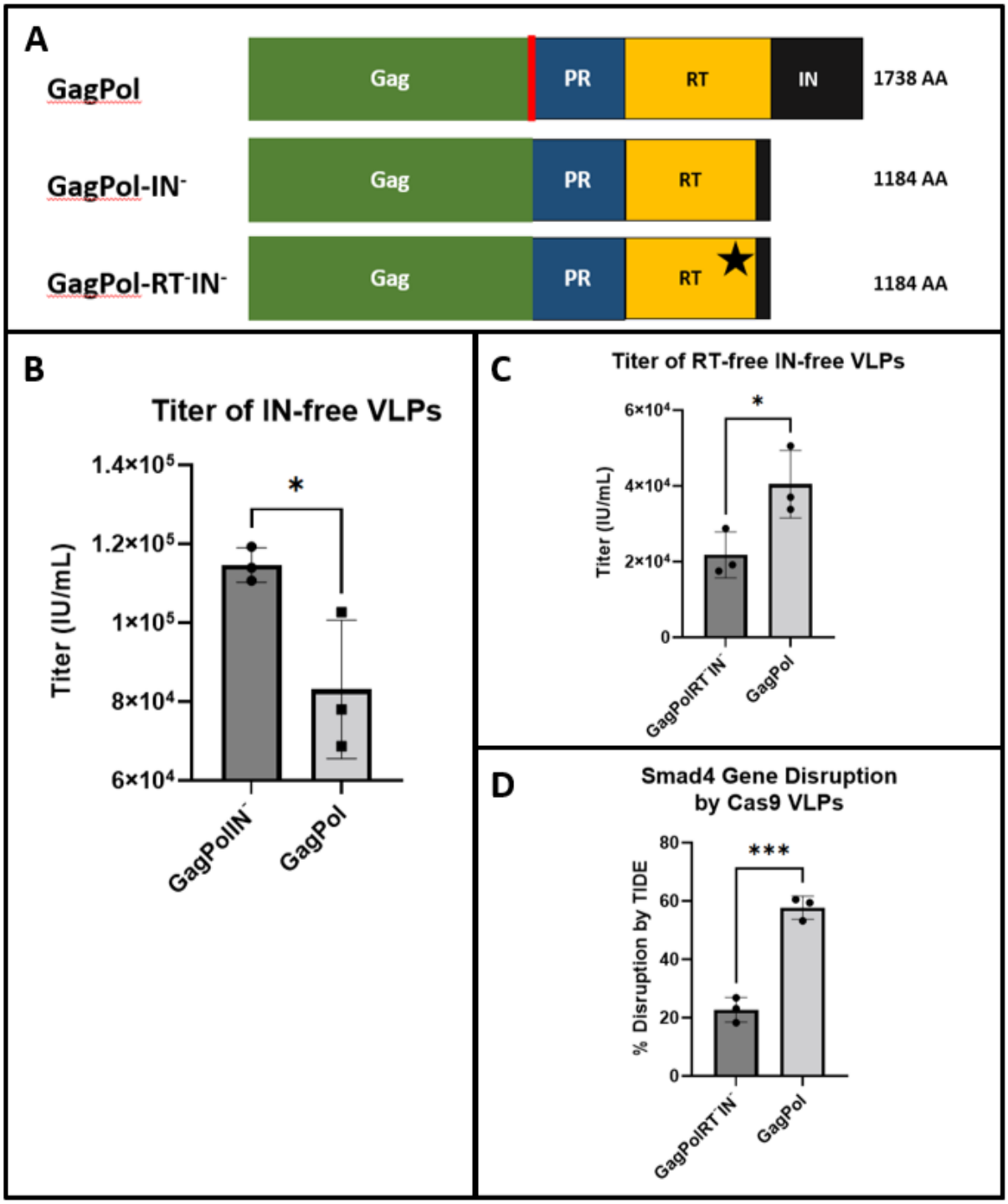

Figure 14 VLPs can deliver functional Cas9 with mutation of RT and deletion of IN.
A) Vectors expressing mutated GagPol proteins were derived from pCS7-GagPol. The length of the proteins is listed to the right. The star indicates the two point mutations in reverse transcriptase that abrogate its enzymatic activity. B) The titer of integrase deficient VLPs appear to be higher than those made with GagPol. C) The titer of Cas9 VLPs deficient in both integrase and reverse transcriptase appears to be two-fold lower than those made with GagPol. D) Cas9 VLPs loaded with sgRNA targeting Smad4 have an approximately 2.5-fold drop in titer when made with GagPolRT-IN- versus GagPol.

It is not an absolute requirement that Cas9 be separated from Gag by the viral protease in order for Cas9 to function in target cells, as shown by the modest reporter activation detected after infection with GagCO VLPs that do not contain the viral protease (Fig. 13BC). The titers of GagCO-Cas9 VLPs lacking any protease was approximately 2.4 x10^3 IU/mL. However, the 15-fold increased titer of VLPs produced with the GagCO-Stop-Protease vector relative to the Gag-only vector, suggests that protease processing increases the efficiency of Cas9 function in target cells. The constructs GagPolIN$^-$ and GagPolRT$^-$IN$^-$ both include the viral protease, but their titers are, respectively, significantly above and below Cas9 VLPs made with GagPol. This would not be an expected outcome if both of the constructs had viral proteases that function similarly and were loaded into VLPs at similar rates.

In order to examine protease processing, we created a western blot from the Cas9 VLP supernatants made from producer cells transfected with the GagPol, GagPolIN$^-$ or GagPolRT$^-$IN$^-$ vectors. We first probed the blot with an anti-Cas9 antibody (Fig. 15A) which revealed that all samples contained both the uncleaved GagCas9 fusion protein and the cleaved Cas9. The presence of uncleaved fusion proteins shows that none of the constructs, not even GagPol, can fully cleave Cas9 from Gag. However, the difference in the intensities of the GagCas9 and Cas9 bands suggests that that the efficiency of protease processing by each of the constructs is different. Cas9 VLPs made with GagPol have a band intensity ratio for Cas9:GagCas9 of 3.4. Cas9 VLPs made with GagPolIN$^-$ have a ratio of Cas9:GagCas9 of 4.7, while Cas9 VLPs made with GagPolRT$^-$IN$^-$ have a ratio of Cas9:GagCas9 of one. This result could be explained by either poor loading of the GagPolRT$^-$IN$^-$ protein into VLPs or poor activity of the viral protease in the GagPolRT$^-$IN$^-$ protein produced by that construct. We assessed the effects of viral protease processing on the Gag polyprotein by stripping the blot and reprobing with an anti-Gag antibody

(Fig. 15B). The results suggest that the Capsid domain is not efficiently cleaved by the viral protease from the GagPolRT$^-$IN$^-$ construct. We cannot be certain whether the lack of processing reflects poor VLP loading or poor protease activity because there is not a clear, quantifiable band for GagPolRT$^-$IN$^-$ itself, which has a mass of 133 kDa. Possibly too little of the protein loads into VLPs for it to be detectable on a western blot, but alternately the full-length protein may not be detectable because GagPolRT$^-$IN$^-$ processes itself into smaller proteins.

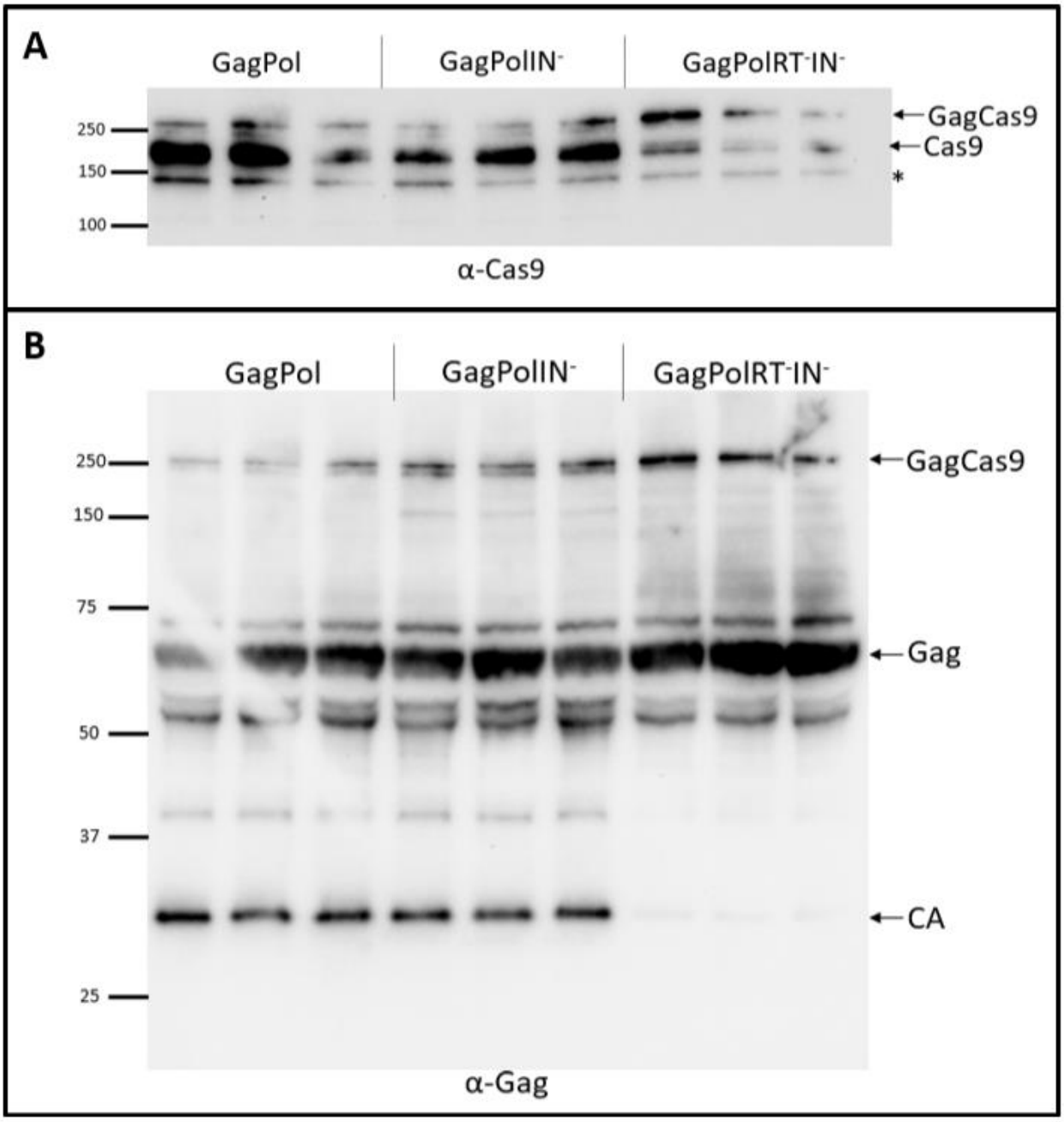

Figure 15 Further characterization of GagPolRT$^-$IN$^-$ VLPs.
A) Western blot to detect Cas9 proteins extracted from VLP supernatants collected in triplicate. Protease separation of Cas9 from Gag occurs in VLPs made with GagPol, GagPolIN$^-$, and GagPolRT$^-$IN$^-$ (triplicate lanes below the label). The top most band corresponds to GagCas9 (226 kDa). The middle band corresponds to Cas9 cleaved entirely from Gag (161 kDa). The lower band is likely a Cas9 degradation product caused by either cellular proteases or the viral protease. B) Western blot to detect Gag. This blot is the same as in A, stripped and reprobed with an anti-Gag antibody in order to show the cleavage products of Gag. The band that appears at ~30 kDa is the capsid (CA) domain of Gag. The lightness of the band in the lanes derived from Cas9 VLPs made with GagPolRT$^-$IN$^-$ indicates that CA is not cleaved as efficiently by the viral protease in these VLPs. A strong Gag band suggests that CA is not greatly reduced due to a dearth of Gag.

3.2.4 GagCO-PE Can Edit a Target Gene in Transiently Transfected HEK293 Cells

Before we attempted to deliver the prime editor by VLPs, we first made and tested constructs expressing the prime editor and prime editing guide RNAs. We obtained the original prime editing plasmid, pCMV-PE2 (Anzalone et al. 2019), for our tests and for constructing our plasmids. We used the pCS7 expression vector to express all of our Gag fusion and GagPol constructs, so we moved the prime editor into pCS7 to create pCS7-PE in order to accurately compare the efficacy of PE expressed from either vector as they have different promoters. We used a convenient restriction site within the Cas9 coding region to move the Cas9 nickase and RT into GagCO-Cas9, meaning that our prime editor construct contains the same GagCO sequence, the same viral protease cleavage site, and the same Cas9 sequence as the GagCO-Cas9 construct, apart from the H840A mutation that disrupts cleavage activity of the Cas9 HNH domain (nickase mutation). Although various pegRNAs have been described, we designed pegRNAs that would install an epitope tag into a highly expressed gene in order to facilitate assaying the results of prime editing by microscopy or flow cytometry. The two highly expressed genes that we chose to target were human lamin A (*LMNA)* and mouse beta-action (*Actb)*. For a test with a directly detectable fluorescent protein as a tag, we constructed a pegRNA designed to insert the entire eGFP coding region as a fusion at the N-terminus of lamin. We directly transfected HEK293T cells with pegRNA plasmids, and either pCMV-PE2, pCS7-PE, or pCS7-GagCO-PE. The cells were split at 24 hours post-transfection and fixed with PFA at 90 hours post-transfection and assayed by confocal microscopy (Fig. 16A) or by flow cytometry (Fig. 16B). The pCS7-PE and pCS7-GagCO-PE plasmids were both successful in installing eGFP in a small percentage of the cells. The eGFP positive cells observed by microscopy all appeared to have eGFP in a pattern that resembles the same nuclear localization that lamin is known to

occupy (Ranade et al. 2019). This staining pattern suggests that the eGFP protein was correctly inserted into *LMNA* and not inserted randomly into the genome. Flow cytometry quantified the frequency of eGFP insertion as approximately 1 in every 250 cells. The pCS7-PE vector was more effective than the original pCMV-PE vector. In addition, although the viral protease was not present, the GagCO-PE fusion protein appeared to be similarly effective or slightly more effective than the PE expressed without Gag from the same vector, suggesting that cleavage of the Cas9-RT prime editor protein from Gag is not essential for PE function. Although eGFP is easy to score, the rate of insertion was likely too low for our initial efforts in creating prime editing VLPs, probably due to the large size of the eGFP insertion. We tested the HA-lamin-pegRNA which introduces the much smaller HA (hemagglutinin) tag at the same locus. The highest rate of editing that we observed by flow cytometry was 27% of the cells positive for the HA tag (data not shown). We therefore decided to move forward with creating prime editing VLPs with the HA-lamin-pegRNA.

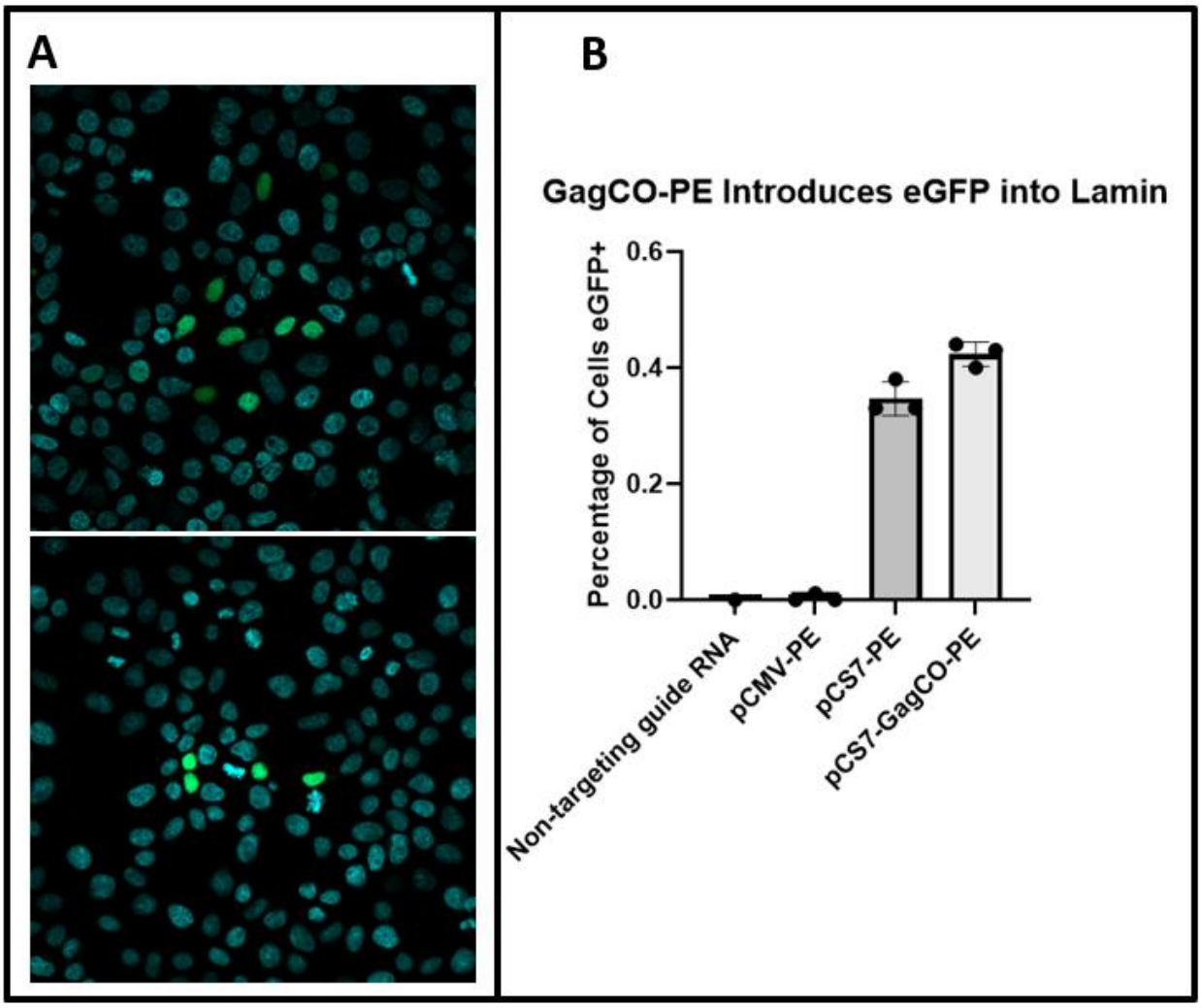

Figure 16 GagCO-PE functions when directly transfected into HEK293T cells.
A) Two representative images obtained from confocal microscopy of HEK293T cells fixed on a glass slide 90 hours after transfection with GagCO-PE and the eGFP-lamin-pegRNA. The nuclei have been stained blue with DAPI. The green signal obscuring the DAPI stained nucleus of some of the cells is eGFP. The magnification is 60x. B) Cells were fixed in suspension and detected by flow cytometry for eGFP expression 90 hours after they were co-transfected with the plasmid encoding eGFP-lamin-pegRNA and the indicated plasmid under each bar of the graph. The co-transfection of pCMV-PE2 and a non-targeting guide RNA served as the negative control.

3.2.5 MLV VLPs Deliver Functional Prime Editing Enzymes

Prime editing can be useful for introducing specific edits into cells, but delivery of the prime editor, Cas9-RT, and a pegRNA can still be challenging, particularly in difficult to transfect cell types. In order to overcome this challenge, we created GagCO-PE VLPs (Fig. 17A). We first tested these prime editing (PE) VLPs using pegRNAs that would install a hemagglutinin (HA) epitope tag in the reading frame of *LMNA*. The PE VLPs were made similarly to the Cas9 VLPs with the exception of transfecting a larger amount of the GagCO-PE plasmid (see Methods). PE VLPs were made using either GagPol or GagPolIN$^-$RT$^-$ as the source of Gag and the viral protease. PE VLPs were concentrated 4-fold before application to HEK293T cells. VLP supernatant was exchanged for fresh media at 24 hours post-infection. Cells were split into two wells 48 hours after infection to allow parallel assays of the same infected cells by immunohistochemistry in monolayer and either flow cytometry or Next-Generation Sequencing.

We used commercially available antibodies to visualize the expression of the HA epitope tag by indirect immunofluorescence (see Methods). Cells treated with PE VLPs showed subcellular localization of the HA tag in a pattern characteristic of the lamin protein (Fig. 17B) consistent with integration of the HA tag into *LMNA*. By microscopy, we estimated the number of cells positive for the HA tag to be about 1 in 200. In order to better quantify the efficacy of prime editing, we fixed and stained cells in suspension and used flow cytometry (Fig. 17C) (See methods). The percentage of HA positive cells was similar, ~0.45%, whether the PE VLP was made with GagPol or GagPolRT$^-$IN$^-$. We also used next-generation sequencing of PCR amplicons from the targeted genomic locus to assess both integration frequency and accuracy. We repeated the PE VLP experiment and collected the genomic DNA from infected cells 72 hours after infection, then amplified the target site by PCR. The targeted *LMNA* locus would

generate create an amplicon of 220 base pairs if the locus is unedited or 250 base pairs if the HA tag is correctly inserted. The NextGen sequencing showed that approximately 0.5% of the sequenced alleles contained the HA insert, and the majority of HA inserts were inserted with perfect fidelity to the pegRNA (Fig. 17C). The sequencing results were consistent with our observations of HA insertion frequency by flow cytometry, although it remains possible that the sequencing may undercount insertions because of the size preference in PCR for amplifying shorter sequences.

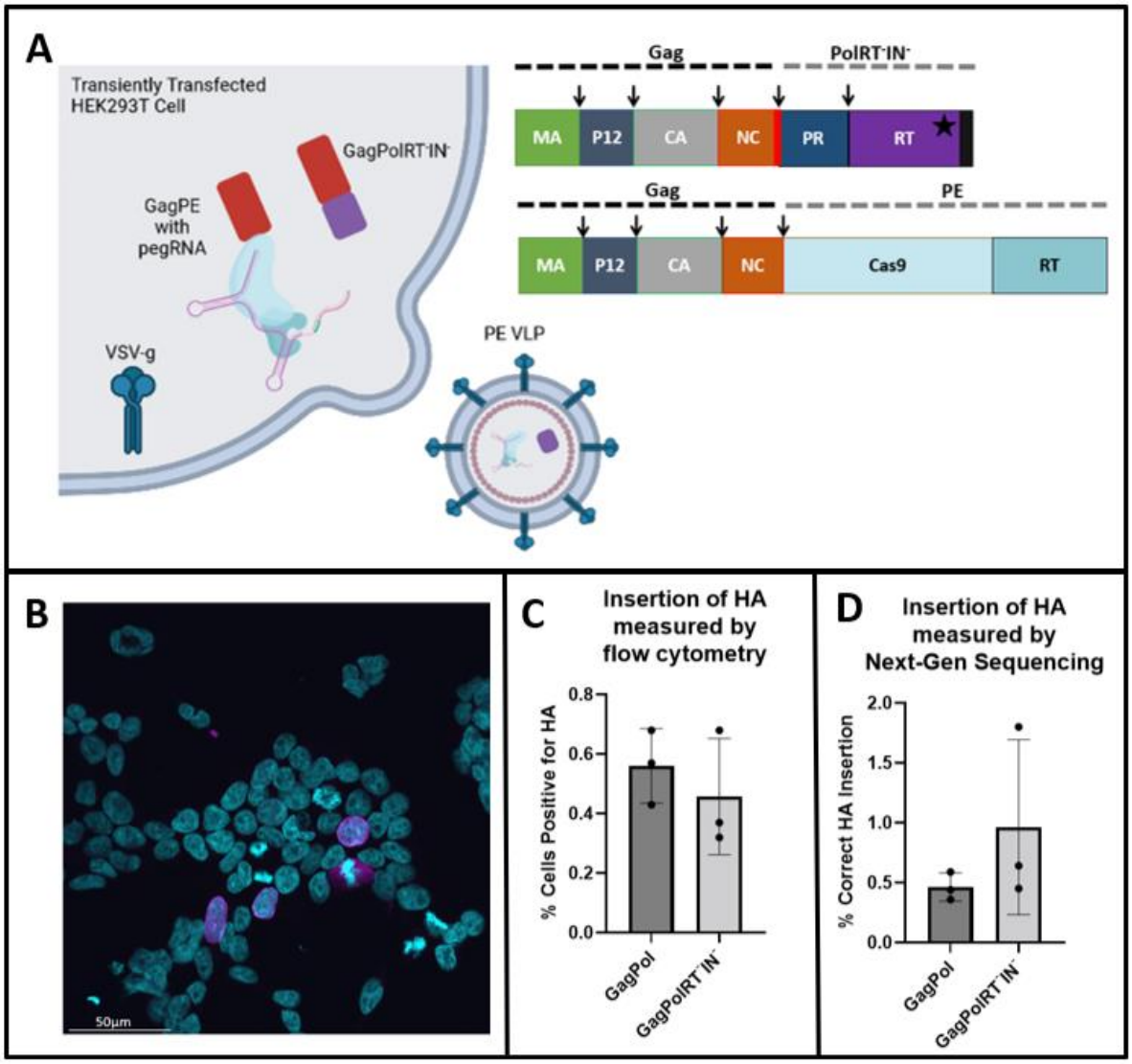

Figure 17 Prime editing VLPs introduce the HA epitope tag into lamin in HEK293T cells.
A) Schematics of the GagPolRT-IN- and Gag-PE proteins. Vectors expressing these proteins, together with VSV-g and a pegRNA were transiently transfected in HEK293T cells to produce PE VLPs. The GagPol vector was also used to create PE VLPs as indicated by "GagPol" in the subsequent graphs. B) HEK293T cells were co-transfected with plasmids expressing GagPE and the HA-lamin-pegRNA. Cells were fixed 72 hours later and stained for the HA tag which shows an expression pattern around the nucleus which is characteristic of lamin's localization in the cell. Cells co-transfected with GagPE and a non-targeting guide were used as a control (data not shown). C) PE VLPs loaded with HA-lamin-pegRNA were used to infect HEK293T cells. The cells were fixed and stained with an anti-HA antibody conjugated to Alexa Fluor 594. Cells infected with PE VLPs loaded with a non-targeting guide RNA were stained as the negative control population. D) PE VLPs loaded with the HA-lamin-pegRNA were used to infect HEK293T cells. Genomic DNA from infected cells was collected 72 hours later and next-generation sequencing revealed that the HA tag was installed at the N-terminus of lamin in almost 0.5% of alleles for VLPs made with GagPol. PE VLPs made with GagPolRT-IN- averaged a rate of slightly over 0.5% for correctly installing the HA epitope tag. 17A was created with BioRender.com

The efficiencies of the PE VLPs appeared to be substantially lower than expected based on prior experiments with Cas9 VLPs. Cas9 VLPs could achieve a saturating efficacy in terms of their ability to disrupt a genomic locus without concentration (Fig. 12B). In contrast, PE VLPs were concentrated and still had an efficacy far less than the editing efficiency of GagPE and the HA-lamin-pegRNA when directly transfected into HEK293T cells, which was as high as 27% (data not shown). Possible explanations could be that GagPE is not produced as efficiently as GagCas9 in producer cells, or that it is degraded in producer cells. We transfected the vectors encoding GagCO-Cas9 or GagCO-PE with their respective guide RNAs into HEK293T cells, harvested the cell lysates, and analyzed expression of the Gag fusions on a western blot probed with an anti-Cas9 antibody (Fig. 18A). GAPDH was used as an internal control to normalize the amount of protein. It appeared that the amount of GagPE (304 kDa) produced was not less than the amount of GagCas9 (226 kDa) produced. Furthermore, there were no smaller bands in the GagPE lanes that would indicate possible degradation by endogenous proteases. Therefore, poor production or stability of GagPE in producer cells is unlikely to explain the low efficacy of PE VLPs. An alternate possibility is that Gag is separated poorly from PE by proteolytic processing in VLPs. In order to examine viral protease processing of GagPE, we harvested PE VLP supernatants from producer cells and detected their products on a western blot, using both anti-Cas9 (Fig. 18B) and anti-Gag (Fig. 18C) antibodies. The anti-Cas9 probed blot suggests that PE VLPs made with GagPol are partially processed. The ratio of PE (244 kDa) to GagPE (304 kDa) is 0.37. In contrast, the western blot of Cas9 VLP supernatants (Fig. 15A) showed that GagPol processing of GagCas9 produced a greater amount of free Cas9 protein than the GagCas9 fusion protein in VLPs. Therefore, GagPol processing of GagPE is less efficient than of GagCas9. The blot in Figure 18B was stripped and probed again with an anti-Gag antibody (Fig. 18C). This

revealed that there is still GagPE present in VLP supernatants made with either GagPol or GagPolRT$^-$IN$^-$. As seen previously (Fig 15B), the amount of capsid cleaved from Gag is much lower in VLPs made with GagPolRT$^-$IN$^-$. However, our experiments with PE VLPs made with GagPol or GagPolRT$^-$IN$^-$ do not show a significant difference between their efficacies. While better cleavage efficiency between Gag and PE would likely improve the efficacies of the PE VLPs, there are likely to be additional factors that contribute to the low efficacy of PE VLPs.

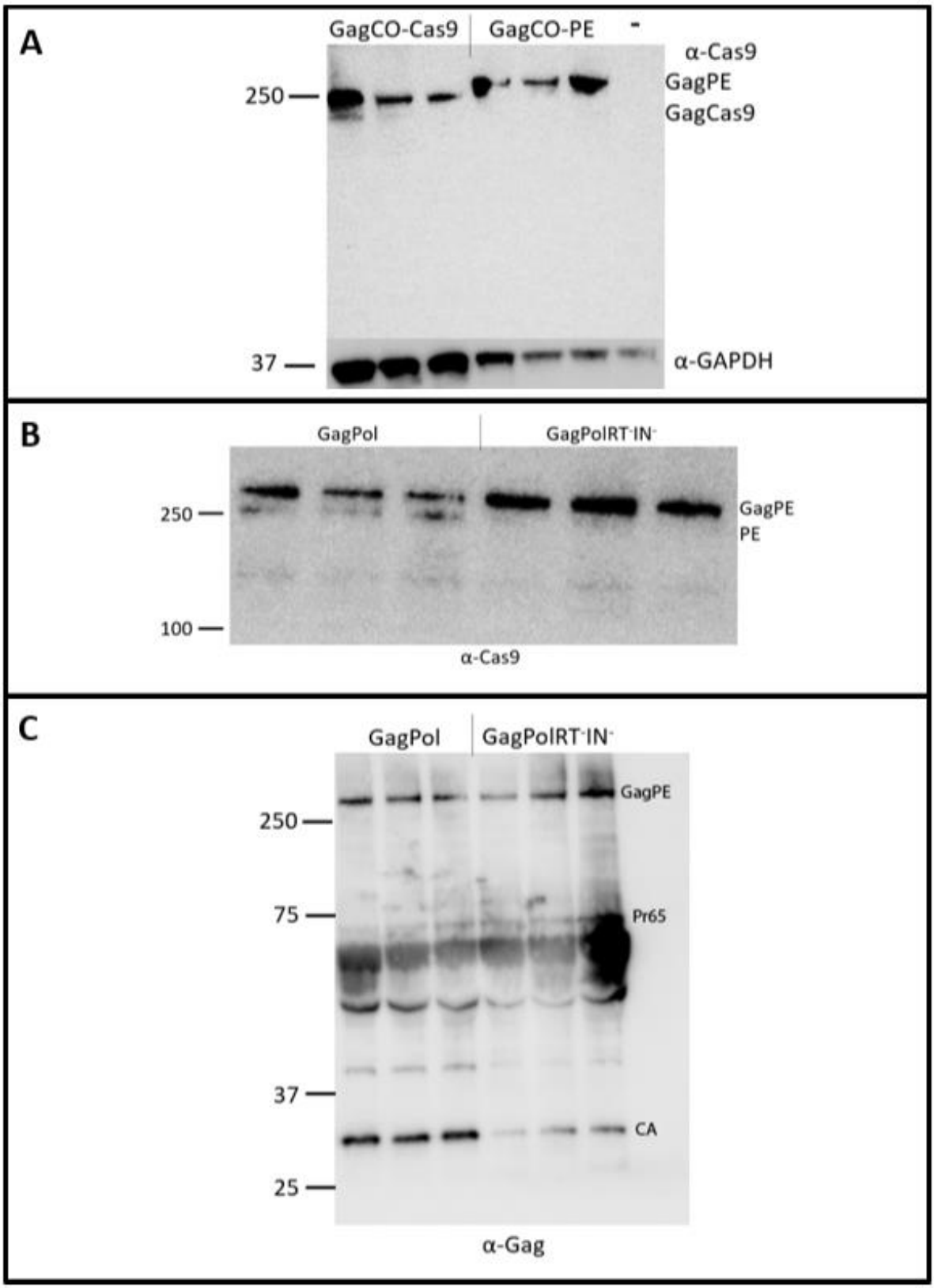
A
GagCO-Cas9
GagCO-PE
-
α-Cas9
GagPE
GagCas9
250
37
α-GAPDH
B
GagPol
GagPolRT-IN-
GagPE
PE
250
100
α-Cas9
C
GagPol
GagPolRT-IN-
GagPE
Pr65
CA
250
75
37
25
α-Gag

Figure 18 Western blots show GagPE production in HEK293T cells and processing by the viral protease in PE VLP supernatants.
A) Plasmids encoding GagCO-Cas9 and GagCO-PE as well as their respective guide RNAs were transfected into HEK293T cells. The cell lysates were harvested 24 hours later. The blot was probed with an anti-Cas9 antibody and an anti-GAPDH antibody. The amount of GagPE produced is similar to the amount of GagCas9 produced. Lanes loaded with GagPE do not have many bands smaller than GagPE indicating that the protein is largely intact in producer cells. The lane labeled "–" was loaded with the lysate from non-transfected HEK293T cells. B) PE VLPs were made as described (see Methods). Lysates from the concentrated supernatant from PE VLPs made with either GagPol or GagPolRT$^-$IN$^-$ were used for a western blot and probed with an anti-Cas9 antibody. The GagPol containing VLPs show two bands, one for GagPE and one for PE. This indicates processing by the viral protease. The GagPolRT$^-$IN$^-$containing VLPs show no clear signs of cleavage of GagPE. C) The blot in B was stripped and reprobed with an anti-Gag antibody. The reduction in the intensity of the capsid band in the lanes for PE VLPs made with GagPolRT$^-$IN$^-$ suggests that cleavage of Gag is not as efficient as with GagPol.

In order to test whether PE VLPs could infect and perform high fidelity editing in mouse cells, we made a pegRNA to install the HA epitope tag at the 3' end of the mouse *Actb* gene encoding β-actin, a cytoplasmic component of the cytoskeleton (Bunnell et al. 2011). We generated PE VLPs loaded with the HA-ACTB-pegRNA the same way as for the HA-lamin-pegRNA and infected mouse NIH 3T3 cells. We assayed the 3T3 cells at 96 hours by fixing and immunostaining them for the presence of the HA tag. Confocal microscopy shows that the HA tag appears in the pattern typically expected for β-actin (Fig. 19A). Our estimate of the frequency of HA positive cells was only around 1 in 1,000, at most. However, the signal for the HA tag as presented by β-actin was stronger than lamin, meaning that cells could be counted by eye using a light microscope with an appropriate filter circumventing the need to capture many photographs in order to count the number of cells positive for the HA tag. Approximately 80,000 cells were in a 24 well plate when they were fixed with paraformaldehyde and then viewed by microscopy. The number of cells in each well positive for the HA tag was counted (Fig. 19B). We found that PE VLPs generated under two conditions, either with GagPol or with GagPolRT$^-$IN$^-$ successfully generated cells with HA expression, but the frequency of positive cells was low, substantially less than one in 1000. The concentrated VLPs yielded ~3-fold more positive cells, indicating dose dependency (Fig. 19B).

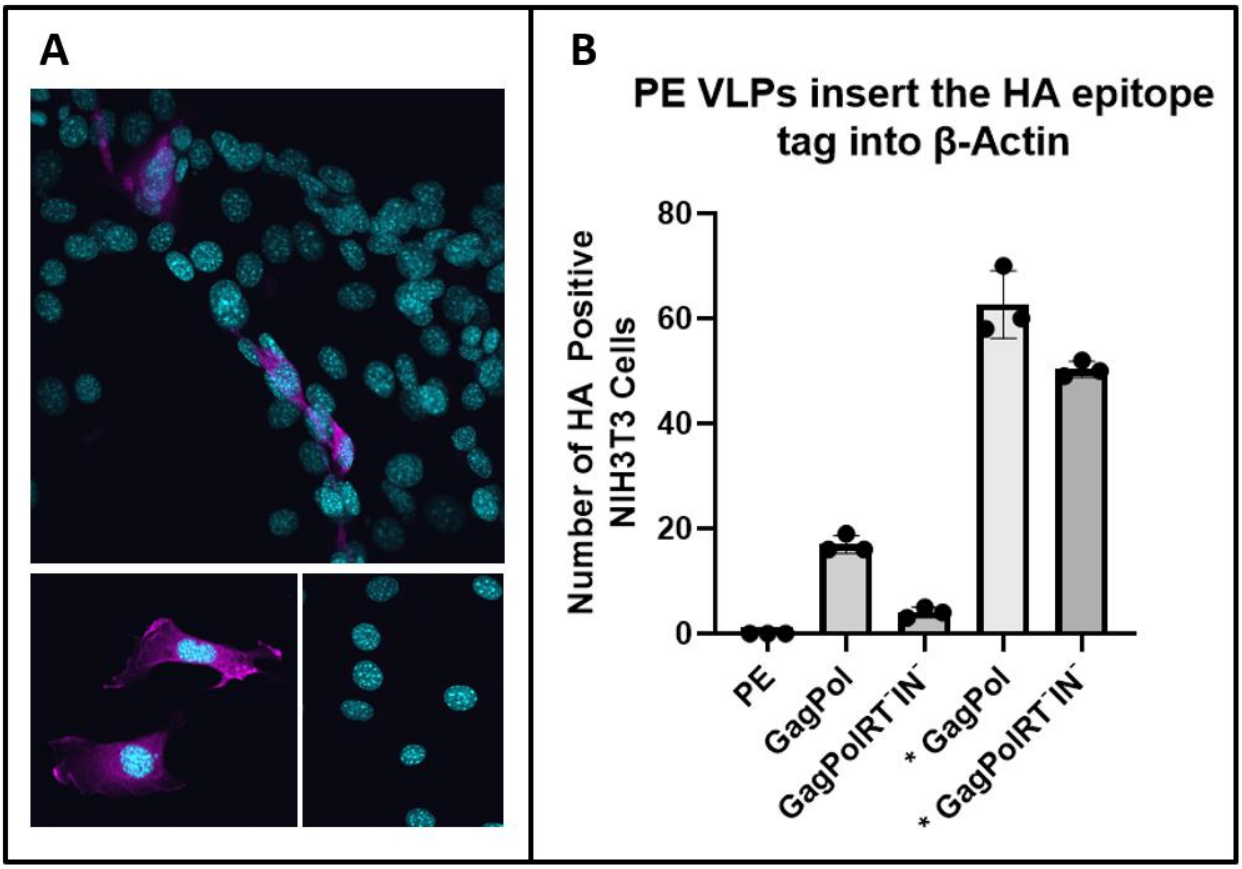

Figure 19 PE VLPs introduce the HA tag into β-actin.
A) Confocal microscopy of NIH3T3 cells fixed and stained 96 hours after infection with PE VLPs. DAPI was used to stain the nuclei. An anti-HA primary antibody and a secondary antibody conjugated to Alexa Fluor 594 were used to detect HA, the signal has been pseudocolored magenta and appears in the cell cytoplasm. B) Cells treated with PE VLPs were assayed 96 hours after infection by immunostaining in monolayer and subsequent counting by microscopy of the total number of HA positive cells in each well, with approximately 80,000 cells per well. PE represents supernatant taken and concentrated from cells transfected only with pCS7-PE and the HA-ACTB-pegRNA, which acted as negative control. GagPol and GagPolRT-IN- are PE VLPs made as previously described with either construct. The replicates with asterisks represent the PE VLPs that were concentrated 4-fold before their application to cells.

3.3 Discussion

VLPs provide an efficient transient delivery method for genome editing proteins. Here we describe the development of new VLPs, based on MMLV, that can deliver Cas9, Cre, or prime editing enzymes. We demonstrated that these VLPs can deliver the editing proteins into mouse fibroblasts, a cell-type that is difficult to transfect efficiently, as well as human HEK293T cells. The delivery of proteins for prime editing by VLPs should allow the application of prime editing in a wider range of cells. We also showed that MLV VLPs can be produced without active viral reverse transcriptase or integrase proteins, which should reduce the potential for unintended genome modifications in target cells. Our VLP design is similar to that used for the "Nanoblades", previously described MLV VLPs for Cas9 or Cre delivery (Mangeot et al. 2019), although the Nanoblades were based on the related Friend MLV. Similar to Mangeot et al., we used the VSV-g envelope protein for pseudotyping the VLPs, allowing for infection of a wide variety of cells (Cronin, Zhang, and Reiser 2005). We also found that the MMLV Eco Env protein could be used for VLP production, providing an additional envelope protein option for rodent cells.

To obtain high VLP titers for GagCas9 and other Gag fusion proteins, we found that it was necessary to recode the MMLV Gag coding region with codons optimized for human cells. Since MMLV is a mammalian virus, we did not anticipate needing to codon optimize, and prior MLV VLP designs did not alter the viral Gag codon usage. However, it has been reported that RNA sequences within the MMLV Gag protein coding region reduce Gag synthesis, while that additional RNA elements in the viral Pol coding region can compensate to increase Gag production (Bartels and Luban 2014). In addition, Bartels and Luban observed that codon optimization of the Gag coding region in the viral RNA could obviate the reduction in Gag in the

absence of the Pol RNA sequences. Likely, RNA regulatory sequences are present within the Gag coding region that impede either mRNA export from the nucleus or protein translation by the ribosome, either of which would effectively reduce the amount of protein produced. The sequence changes (third position substitutions, etc.) during codon optimization likely disrupt the embedded regulatory elements. Prior to constructing codon optimized Gag, we also constructed a GagCas9 vector with 890 nt of sequence from Pol mRNA placed out of frame in between the stop codon of GagCas9 and the SV40 polyA site. This construct improved Cas9 VLP titers by at least two-fold relative to GagCas9 without the Pol sequences (data not shown), which indicated that Gag codon optimization might be useful in improving Cas9 VLP titers. Previous MLV VLP designs did not codon optimize Gag. Previous VLPs based on MMLV included the Pol gene as part of the Gag-cargo protein transcript (Voelkel et al. 2010; Bobis-Wozowicz et al. 2015), likely allowing higher Gag production. The Nanoblades VLPs used the Gag coding region from Friend MLV. Friend MLV Gag and MMLV Gag are similar at the protein level (89.4% identity), but the mRNA sequences are more divergent, such that 27% of codons have changes in the third position. Potentially this divergence also has disrupted the Gag regulatory elements in the Friend MLV Gag coding region.

In addition to improving protein synthesis, the codon optimized version of MMLV Gag also generated more full-length protein. In Eukaryotes, the initiation of protein translation is improved by the Kozak consensus sequence (GCC)GCCRCCAUGG, the start codon is highlighted in red and R is a purine (Dhar et al. 2014). In the original GagCas9 construct, the first AUG sequence for initiating translation was AUCCGUAUCAUGG, whereas the codon optimized sequence was changed to GCCGCCAACAUGG. If translational initiation does not happen at the first AUG, the ribosome can still begin translation at an AUG further downstream.

This alternative translation is common in biology and one mechanism in which alternate forms of proteins can be produced. If the second AUG in Gag becomes the initiator methionine, the protein is truncated by 193 amino acids which amount to ~21 kDa. This N-terminal truncation would result in the absence of the Matrix domain and half of the P12 domain, both of which are important to VLP formation (see Chapter 1.6). The Matrix domain is the most N-terminal of Gag's domains. Gag is myristylated at the N-terminus, which is the chemical signal directing the Gag protein to the plasma membrane. If the Gag fusion protein does not efficiently reach the plasma membrane it cannot load into a particle that can potentially be released from the cell membrane. Also, the Matrix domain is known to be involved in the Gag-Gag interactions that are important for multimerization and particle release from the cell membrane. If Gag cannot multimerize, it cannot form particle. If it cannot release, it will not be able to infect other cells. Adjacent to the Matrix domain is the P12 domain. The N-terminal truncated Gag protein would be missing half of the P12 domain, specifically the half that includes the PPPY sequence which may impair cleavage of Gag and Gag-fusion proteins. In sum, the expression of Gag beginning from the second methionine could potentially result in impaired viral particle formation, release, and proteolytic processing of Gag. We conclude that the poor titer of VLPs generated with the wild-type GagCas9 construct is likely to be a result of both the reduction in the amount of protein produced and the alteration in the length of GagCas9. However, we did not investigate these properties independently, so we cannot assess their relative contributions to the observed change in titer.

The Nanoblade GagCas9 VLPs described by Mangeot et. al. were made using a single GagPol vector as the source of both Gag Pr65 and viral protease (from GagPol) for the creation of VLPs. We found that GagCas9 VLPs created with the coexpression of only MMLV Gag Pr65

were low, at ~2500 IU/mL (Fig. 13BC). The inclusion of the viral protease from the construct Gag-Stop-Prot increased VLP titer ~10-fold suggesting that the separation of Gag from Cas9 is important for improving Cas9's function in a cell, possibly by improving Cas9's ability to localize to a targeted genomic locus once inside the cell. We also made VLPs with GagPolRT$^-$IN$^-$. The GagCas9 and Cas9 bands on the western blot (Fig. 15B) indicate that the GagPolRT$^-$IN$^-$ protein still produces a functional viral protease, albeit it is less efficient than the viral protease from GagPol. This could be because of either poor loading of GagPolRT$^-$IN$^-$ into VLPs or due to a lower activity of the protease. Although our interpretations of the western blots assume that Gag cleavage is due to the viral protease as previously observed for MLV, it is formally possible that the cleavage products we observed were cleaved by a cellular protease rather than the viral protease, as we did not test GagPol or GagPolRT$^-$IN$^-$ with a disruption of the protease. An alternate explanation for why constructs containing the protease improve GagCas9 VLP titers could be that they improve the polymerization of GagCas9 into particles and therefore more Cas9 is loaded into VLPs and subsequently delivered to target cells. Nonetheless, our data indicates that neither functional reverse transcriptase nor integrase are required to generate VLPs, and the absence of these factors has a minimal effect on VLP titer. GagPolRT$^-$IN$^-$ VLP titers were about two-fold lower than Cas9 VLPs made with GagPol, and the reduced cleavage of GagCas9 (Fig. 15B) provides a likely explanation for that discrepancy. It might still be possible to make a construct lacking reverse transcriptase and integrase while still preserving the functionality of the protease. Reverse transcriptase and integrase are powerful enzymes with the potential to cause off-target changes to the genome which could confound the results of assays meant to only perturb the function of a gene targeted by Cas9. Furthermore, if VLPs are to be considered as a potential vehicle for the therapeutic delivery of Cas9, the omission of these

enzymes' presence removes the need to consider their contributions to off-target effects, which likely would be worth a two-fold drop in efficacy when making GagPolRT$^-$IN$^-$ VLPs. We do not know why there is a small drop in VLP titer from the loss of these two enzymes, particularly since point mutations were used to disrupt RT activity. It is possible that the folding of the truncated/mutated proteins makes them less amenable to forming productive Gag-Gag interactions and/or reduce loading into particles. In that case, introducing additional mutations to the truncated GagPol might be able to correct for this defect or even improve VLP titer beyond what can be achieved with GagPol VLPs.

To our knowledge, delivery of prime editing enzymes with VLPs has not been reported previously. We have shown that not only can the prime editor be delivered by VLPs, but also that it can accurately install an epitope tag in an infected target cell. The longer sequence (~30 nt) necessary to insert a tag at the nicking site is known to be more challenging to efficiently insert with prime editing than a point mutation (Anzalone et al. 2019). We have shown that prime editing installs the HA tag into the human *LMNA* gene by the pattern of immunostaining for the tagged lamin protein and by next-generation sequencing of the insertion sites. Furthermore, we have shown that PE VLPs can be made using the GagPolRT$^-$IN$^-$ (GagPol without functional reverse transcriptase and integrase). The RT and IN enzymes are not essential to the formation of VLPs. For prime editing, the prime editing enzyme is a fusion between Cas9 and MMLV RT. In this situation it may be possible for viral RT derived from GagPol to interfere with prime editing. The use of GagPol RT$^-$IN$^-$ obviates the need to test whether there is any interference. Although PE VLPs are not as efficient as Cas9 VLPs, our data suggests that there may be a way to improve PE VLPs in the future: cleavage of GagPE is less efficient than

cleavage of GagCas9 by the viral protease from GagPol (Fig. 18B) so improving the efficiency of separation between Gag and PE could improve PE functionality in target cells.

Prime editing VLPs are an important proof of concept that expand the range of gene editing enzymes that can be delivered with VLPs. The use of prime editing is likely to be an improvement in how genome editing can be achieved with minimal risk of off-target effects. Mangeot et. al. demonstrated the first use of Cas9 VLPs to introduce a genomic edit with the assistance of a repair template. They combined their Nanoblades with single-stranded DNA donor templates in order to introduce a FLAG tag at the human *DDX3* locus in HEK293T cells. The efficacy of this approach is likely better than the current, first-generation prime editing VLPs. However, prime editing still has the advantage of the use of a nicking Cas9 which avoids double-stranded breaks and the higher risk of mutations at off-target loci, and makes investment in this strategy potentially more beneficial in the long-term. It is clear that the prime editing VLPs we have tested to date do not approach the same efficacy as Cas9 VLPs. When the prime editor was directly transfected into HEK293T cells with the HA-lamin pegRNA, the highest observed number of cells positive for the HA tag by flow cytometry was 27% (data not shown). The prime editor successfully installed an HA tag ~20-30-fold less frequently when delivered by VLP. In comparison, Cas9 can disrupt genes with similar or higher efficiency when delivered by VLP (Fig. 12B) or by the transient transfection of plasmids into HEK293T cells. There are several possible explanations for this discrepancy. It is unknown whether there is a hard size limit to the proteins that can be delivered with VLPs. Mangeot et. al. delivered Cas9-derived transcriptional activator (SP-dCas9-VPR), a 224 kDa protein, successfully with their virus-like particles. The prime editor is even larger at 243 kDa, which could possibly limit loading into particles. Another possibility is that the structure of the Gag-prime editor fusion sterically

hinders Gag from polymerizing into particles. Although we could not make a fair comparison between Cas9 and prime editing VLPs, we note that Cas9 VLPs were visible on a western blot probed with the anti-Cas9 antibody without concentration whereas the PE VLPs required a 20-fold concentration in order to be visible on a western blot (data not shown). This observation suggests that less Gag-prime editor is loaded into VLPs than GagCas9. There also is the question of delivery to the nucleus of target cells. Our version of the prime editor has an NLS near the N-terminus of nCas9 and another near the C-terminus of reverse transcriptase. However, there is a viral protease site right before the second NLS sequence. It is possible that the viral protease cleaves this site before the prime editor is delivered to target cells, thus decreasing the ability of the protein to localize to the nucleus in the target cell. Furthermore, the pegRNA associated with the prime editor is longer than a standard guide RNA. It may be that the guide RNA is partially degraded before the prime editor reaches the site to be edited, or that the longer RNA reduces Gag-prime editor incorporation into VLPs. Despite the initial challenges to delivering the prime editor by VLP, this technology does function at a detectable level which allows for further optimizations to be made. In addition, for cell culture applications with difficult to transfect cells (e.g. primary cells), a low percentage of prime editing via VLP delivery may still be more practical than transfections, if identification of the modified cells is practical. Future optimizations may lead to VLPs capable of performing prime editing more efficiently, allowing general applications in cell types that are difficult to transfect, or even in vivo for therapeutic purposes.

The most recent example of Cas9 delivery with VLPs was published mere weeks ago in Cell (Banskota et al. 2022). Banskota et. al. described engineered virus-like particles (eVLPs) derived from the Friend MLV that deliver base editors. Base editors consist of Cas9 fused to an

enzyme capable of catalyzing the interconversion of DNA bases. Banskota et. al. made several improvements to VLPs that are relevant to our work. The first is the variation in cleavage site. This is compared to our findings and discussed at length in Chapter 4.2.5. The second improvement Banskota et. al. made was to increase Gag-fusion protein localization to the cytoplasm. The bipartite nuclear localization signals on their constructs, also found on our GagCO-Cas9 and GagCO-PE vectors, encourage protein trafficking to the nucleus. This is at odds with multimerization into particles that will bud from the cell membrane. Banskota et. al. added a nuclear export signal (NES) to the end of Gag before the protease cleavage site or at the C-terminus of the Gag-fusion protein separated by a peptide substrate for the viral protease. The NES increases Gag-fusion protein concentration in the cytoplasm, and is separable from the base editor after it is delivered to target cells where it is no longer needed. The addition of the NES after Gag, but not at the end the Gag-fusion protein substantially increased eVLP efficacy. We also suspected that localization to the nucleus was a problem with Cas9 VLPs. We added a NES to the end of the GagCas9 protein and placed a viral cleavage site in between it and Cas9. We observed no increase in titer of VLPs created with the GagCas9-NES sequence (data not shown). This aligns with the data gathered by Banskota et. al. where they also found that placing the NES at the end of the Gag-fusion did not improve eVLP efficacy. We chose to pursue other optimizations of Cas9 VLPs instead of further NES focused investigations, but we would highly recommend the inclusion of a NES at the end of Gag in the creation of future Cas9 or PE VLPs.

In the MLV, the ratio of Gag to GagPol has been optimized by evolution. By high-jacking the viral machinery to deliver Gag-fusion proteins, we are necessarily introducing a protein that might have a different optimal stoichiometry with Gag Pr65 for its multimerization into particles. Electron microscopy has revealed that Gag and GagPol have fairly rigid structures

in particles made with a mutated protease that will not cleave them into their constitutive components (Kol et al. 2006). A Gag-fusion protein may not inherently possess the rigidity or flexibility needed for efficiency loading in VLPs. Figure 13C shows that the Gag-Stop-Prot construct is only two-fold different from the GagPol construct. We hypothesized that the Gag-Stop-Prot, which presumably produces Gag Pr65 and Gag-Prot, might have different requirements for the stoichiometry of loading GagCO-Cas9. We varied the ratio of GagCO-Cas9 to Gag-Stop-Prot by varying the amounts of each plasmid transfected into producer cells in order to test whether the ratio of GagCO-Cas9 to Gag-Prot might be optimized, but our initial efforts were unsuccessful (data not shown, but I actually have this and could add it in). Banskota et. al. also attempted to optimize the stoichiometry of Gag/GagPol to the Gag-fusion protein. They found that decreasing the ratio of the transfected Gag-fusion plasmid to the GagPol plasmid improved their eVLP efficiency. Further optimization of the ratios between GagPol, GagPolRT-IN-, or Gag-Stop-Prot may also be a source of improvement for our Cas9 and PE VLPs.

3.4 Methods

Plasmids:

pCS7 expression vectors were built by standard methods. The pCS7 expression vector is based on pCS2p+(Turner and Weintraub 1994; Rupp, Snider, and Weintraub 1994) but contains the human UBC intron 1 inserted into the transcription unit, between the simian CMV promoter and the multiple cloning sites. For pCS7-Gag-Pol, the Moloney MLV Gag from pUMVC3-gag-pol (University of Michigan Vector Core) was inserted into pCS7. For pCS7-Gag, Gag was PCR amplified from pUMVC3-gag-pol with primers GCAGGATCCGTATCATGGGCCAGACTGTTACCACT and GAATCTAGAGGCTAGAATTCCCCCATGGCCCCGCTAGCTCCCTGGTCATCTAGGGTC

AGGAGGGAGGTCTGAGGCCTTGGTCCCCG and inserted into pCS7. pCS7-Gag-Cre was made by inserting SV40NLS-Cre into pCS7-Gag. pCS7-GagCas9 was made by inserting 3x-Flag-Cas9 from pUBC4580p-sp6-Cas9-SV40. The codon optimized Gag in pCS7-GagCO, was designed using Genscript's codon optimization tool GenSmart and synthesized as a gBlock (IDT) that was inserted into pCS7. Cre and Cas9 were inserted similarly into pCS7-GagCO to create pCS7-GagCO-Cas9 and pCS7-GagCO-Cre. pCS7-GagCO-Cre-Cas9 was created by inserting the protease site and Cas9 from pCS7-GagCO-Cas9 after pCS7-GagCO-Cre. pCS7-Gag-CAG-Pol was created by amplification of pCS7-Gag-Pol with primers TGACCAGGGAGGTCAGGGT and TCTAGAGTCAGGAGGGAGGTCT. pCS7-GagCO-Prot was created by amplifying pCS7-Gag-CAG-Pol with primers CCGCAGACCTCCCTCCTGACCCTAGATGACCAGGGAGGTCA and TGAGCTAGCGCTCCCAGGCCTATACTCATCTTCTATATTTAGGGTCAACACTTGCAGGGG followed by insertion into pCS7-GagCO. pCS7-GagCO-Stop-protease was created by amplifying pCS7-Gag-Pol with primers CCCCAGACCTCCCTCCTG and GCTAGAATTCCTACAACACTTGCAGGGGCTG and insertion into pCS7-GagCO. pCS7-Gag-Pol-IN$^-$ was created by cutting pCS7-Gag-Pol with EcoRV and SmaI and recircularizing. pCS7-Gag-Pol-RT-IN$^-$ was created by PCR with primers CTTGATCCTGCTACAGTACGTGAACAACTTACTGCTGGCCGCCACTTC and GAAGTGGCGGCCAGCAGTAAGTTGTTCACGTACTGTAGCAGGATCAAG. pCMV-VSV-G (plasmid #8454) and pHCMV-EcoEnv (plasmid #15802) (ref) were from AddGene. All guide RNAs were cloned into a vector containing the U6 promoter. The SV40 Large T Antigen was cut from pSP64-TAg (a gift from the Imperiale lab) and subcloned into pUS6 (Zhang et al. 2022) to create pUS6-TAg. GagCO-PE was created by moving part of Cas9-RT from pCMV-

PE2 into pCS7-GagCO-Cas9 using SacI and EcoRI to create pCS7-GagCO-PE-short. The missing 10 amino acids from the end of PE were introduced by ligating annealed oligos AATTCGAGCCCAAGAAGAAGAGGAAAGTCTAAT and CTAGATTAGACTTTCCTCTTCTTCTTGGGCTCG into pCS7-GagCO-PE-short digested with EcoRI and XbaI to create pCS7-GagCO-PE.

Cell Culture and Mice:

HEK293T (Gesicles) cells (Clontech 632617) were grown in DMEM (Gibco 11995065) with 10% Fetal Bovine Serum (FBS) and penicillin/streptomycin (Gibco 15140122). All mouse (Mus musculus) experiments were approved by the Institutional Animal Care & Use Committee at the University of Michigan. The Mouse Dermal Fibroblast (MDF) Ai9/RiboTag line was derived from the skin of a postnatal day 4 C57Bl6 Ai9/RiboTag (Durkin et al. 2013; Kheirollah et al. 2016). The mouse was euthanized, skin was removed from both forelimbs in PBS, and small pieces of skin were transferred to a petri dish containing 0.25% trypsin. The skin was minced with a scalpel. The dish was transferred to a 37°C incubator with 5% carbon dioxide for ten minutes. Aliquots of the digested cells were then transferred to conical tubes and the trypsin was quenched by the addition of 10 parts DMEM/10% FBS/Pen/Strep. The cells were centrifuged for five minutes at 500 x g. The supernatant was aspirated and the cell pellets were resuspended in DMEM/10% FBS/Pen/Strep. The primary cells were allowed to adhere to the bottom of 10 cm dishes pre-treated with 0.1% gelatin. The cells were frozen or transfected for immortalization within 5 passages. pUS2-Tol2 (Gupta et al. 2018) and pUS6-TAg were transfected into the primary MDFs using TransIT-LT1 (Mirus). Individual cells were plated to individual 96-wells and a proliferating clone was selected from the transfected cells for expansion. MDFs were

grown in DMEM with 10% FBS and penicillin/streptomycin. HEK293T and MDF Ai9/RiboTag cells tested negative for mycoplasma.

Virus-Like Particles:

Cre and Cas9 VLPs

HEK293T cells were plated 24 hours before transfection in 15.6mm dishes (24-well plates). Their density at the time of transfection was approximately 85%. They were transfected with *Trans*IT-VirusGen (Mirus 6703) according to the manufacturer's instructions. The plasmids and their amounts varied depending on the experiment. Unless otherwise specified, the amount of plasmid transfected totaled 500 ng of which 25 ng was pCMV-VSV-G, 130 ng was pCS7-Gag-Pol, 85 ng was pCS7-GagCas9, and 250 ng was the pU6 vector expressing a guide RNA. Figure 12A had 25ng of pHCMV-EcoEnv transfected alone, or in addition to 25ng of pCMV-VSV-g. At 24 hours post-transfection, the media was removed and 1mL of pre-warmed media was added to the cells. At 48 hours post-transfection, the supernatant was harvested and centrifuged at 6,000 × *g* for 20 seconds and 0.95mL was moved to a new tube. Polybrene (EMD Millipore TR-1003-G) was added to a final concentration of 8 µg/mL. Cells were plated in 15.6 mm dishes 24 hours prior to infection, with approximately 20,000 cells/well at the time of infection. Cells were infected with 300µL of viral supernatant. If less than 300µL supernatant was used, the total volume was increased to 300µL by the addition of fresh media. At 24 hours post-infection, 300µL of fresh media was added to the cells. VLP titers are calculated by dividing the numbers of cells that express tdTomato by the volume of VLP supernatant applied to cells (infectious units per milliliter).

Prime Editing VLPs

HEK293T cells were plated 24 hours before transfection in 15.6mm dishes (24-well plates). Their density at the time of transfection was approximately 85%. They were transfected with *Trans*IT-VirusGen (Mirus 6703) according to the manufacturer's instructions. The amount of plasmid transfected was 25ng of pCMV-VSV-G, 130 ng of pCS7-Gag-Pol, 200 ng of pCS7-GagCO-PE, and 250 ng was the pU6 vector expressing a pegRNA. Two wells were transfected for each condition in order to create enough viral supernatant for one tube to be subsequently centrifuged. At 24 hours post-transfection, the media was removed and 1mL of pre-warmed media was added to the cells. At 48 hours post-transfection, the supernatant from two wells was harvested and filtered through a 0.45 micron surfactant-free cellulose acetate filter (Thermo Scientific 723-2545) and 1.2 mL was moved to a new tube. The supernatant was centrifuged at 16,000 × *g* for 98 minutes at 4°C. The 1.15 mL of supernatant was removed slowly to avoid disturbing the white pellet at the bottom of the tube. 250µL of fresh media spiked with 10 µg/mL of polybrene (EMD Millipore TR-1003-G) was added and the pellet was resuspended in approximately 300 µL total volume. Do not introduce bubbles during resuspension. Cells were plated in 15.6 mm dishes 24 hours prior to infection, with approximately 20,000 cells/well at the time of infection. The applied VLP supernatant was removed at 24 hours and fresh media was applied to cells. Cells were passaged at 48 hours. Cells were assayed 96 hours after infection.

TIDE Assays:

Cells were harvested 48 hours after infection. Genomic DNA was isolated with a Zymo Quick-DNA miniprep plus kit. PCR was done with the appropriate primers (see Table 5) and Sanger sequencing was performed by Eurofins. TIDE analysis came from (Brinkman et al. 2014).

Next-Generation Sequencing:

Cells were harvested 72 hours after infection. Genomic DNA was isolated with a Zymo Quick-DNA miniprep plus kit. PCR was done with the appropriate primers (see Table 5). PCR products were purified with a Zymo DNA Clean and Concentrator-5 kit. PCR products were analyzed using Illumina sequencing by the Massachusetts General Hospital CCIB DNA Core. The online version of CRISPResso2 (Clement et al., 2019) was used to analyze sequence reads for insertions and/or indels.

Flow Cytomtery:

MDFs infected with Cas9 VLPs are harvested 48 to 72 hours post-infection, fixed with 4% PFA, and quantified with a flow cytometer. Both the MACSquant VYB and the Biorad Ze5 flow cytometers were used for quantification. The flow cytometry results were analyzed using FlowJo™ v10.8 Software (BD Life Sciences).

Indirect immunofluorescent detection for cells in monolayer:

For detection of HA, cells were fixed with 4% PFA in PBS for ten minutes and then washed three times with PBS. Cells were incubated with a blocking solution of 0.2% Triton and 30% goat serum in PBS for ten minutes. The anti-HA antibody (Cell Signaling Technology C29F4) was diluted 1:1600 in 0.05% Triton and 10% goat serum in PBS. Incubation was for 1 hour at room temperature, followed by three washes with PBS. Alexa Fluor® 594 AffiniPure Donkey Anti-Rabbit IgG (H+L) (Jackson Immunoresearch 711-585-152) was diluted 1:1000 in 0.05% Triton and 10% goat serum.

Indirect immunofluorescent detection for cells in suspension:

Cells were fixed and stained using the eBioscience™ Foxp3/Transcription Factor Staining Buffer Set (Invitrogen) according to the manufacturer's instructions. Rabbit anti-HA was diluted 1:1600

in perm buffer and 200µL were applied to each tube of cells. Incubation was 1 hour at room temperature. Cells were centrifuged and washed once with perm buffer before addition of the secondary antibody. Anti-rabbit Alexa 594 was diluted 1:1000 in permeabilization buffer and 200µL was applied to each tube and incubated for 30 minutes. Cells were washed once in permeabilization buffer, then twice in PBS.

Western Blotting

Cell Lysates:

Cells were harvested 24 hours post-transfection in RIPA buffer (Sigma Aldrich) with protease inhibitor cocktail (Sigma Aldrich P8340). Lysis proceeded for twenty minutes on ice. Lysate was centrifuged at 16,000 × g for five minutes at 4°C. Supernatant was combined with 4x Laemmeli sample buffer (Biorad #1610747) and boiled for five minutes. All samples were loaded onto a NuPAGE 10% bis-tris gel (Invitrogen NP0301BOX) with Precision Plus Protein Kaleidoscope Prestained Standard (Biorad #1610375). After electrophoresis, wet transfer was achieved using an X-Cell II blot module (Invitrogen EI9051) onto a PVDF membrane (EMD-Millipore IPVH00010) for 60 minutes at 30V. The membrane was blocked with 3% milk in Tris Buffered Saline with 0.1% Tween-20 (TBST) for thirty minutes. The Cas9 antibody (Cell Signaling Technology 7A9-343) was diluted 1:1000 in 3% milk in TBST and incubated at room temperature for 1 hour. The blot was washed three times with TBST. The secondary antibody, goat anti-mouse HRP (Biorad) was diluted 1:3000 in 3% milk TBST and incubated for 45 minutes. The blot was washed three times in TBST before the addition of the Clarity Max Western ECL substrate (Biorad #1705062). After imaging, the staining process was repeated with the GAPDH antibody (Cell Signaling Technology 14C10) and the goat anti-rabbit HRP (Biorad 172-1019).

Cas9 Virus-Like Particles:

VLPs were made as indicated above. VLP supernatant was combined with RIPA buffer (Sigma Aldrich) with protease inhibitor cocktail (Sigma Aldrich). Lysis proceeded for twenty minutes on ice. Lysate was combined with 4x Laemmeli sample buffer and boiled for five minutes. Samples were briefly centrifuged. All samples were loaded onto a NuPAGE 10% bis-tris gel with Precision Plus Protein Kaleidoscope Prestained Standard. After electrophoresis, wet transfer was achieved onto a PVDF membrane for 60 minutes at 30V. The membrane was blocked with 3% milk in TBST for 30 minutes. The anti-Cas9 antibody (Cell Signaling Technology 7A9-343) was diluted 1:1000 in 3% milk in TBST and incubated at room temperature for 1 hour. The blot was washed three times with TBST. The secondary antibody, goat anti-mouse HRP (BioRad) was diluted 1:3000 in 3% milk TBST and incubated for 45 minutes. The blot was washed three times in TBST before the addition of the Clarity Max Western ECL substrate (Biorad #1705062). The blot was stripped with Restore^{TM} Western Blot Stripping Buffer (Thermo Scientific) according to the manufacturer's instructions. The blot was then blocked with 3% milk in TBST for 15 minutes. The rabbit polyclonal anti-Gag antibody (Abcam ab100970) was diluted 1:1000 in 3% milk in TBST and incubated overnight at 4°C. The blot was washed three times with TBST. The secondary antibody goat anti-rabbit HRP (Biorad 172-1019) was diluted 1:3000 in 3% milk in TBST and incubated for 45 minutes at room temperature. The blot was washed three times in TBST before the addition of the Clarity Max Western ECL substrate (Biorad #1705062).

Prime Editing Virus-Like Particles:

Prime Editing VLPs were made as described above. The PE VLP concentrate was combined with RIPA buffer (Sigma Aldrich) with protease inhibitor cocktail (Sigma Aldrich). Lysis

proceeded for twenty minutes on ice. The rest of the protocol is identical to the one described above for Cas9 virus-like particles.

3.5 Nucleic Acid Sequences

Table 2 The sequences of guide RNAs and prime editing guide RNAs

Guide RNA Sequences	
Ai9	AAGTAAAACCTCTACAAATG
Hist1 H1C1	GGTTGCAGCCAAGAAAAAGT
Smad4	TGTATGGTGACACACTTGCT
WTAP	TTTACAGAAGAAATATAGTG
Rpl22	TGCTTGCCTTGTTAGAGCAC
KRAS	AAGAGGAGTACAGTGCAATG
Vimentin	TGCCGAGGACCGGGTCACAT
Human HA-lamin-pegRNA	CCATGGAGACCCCGTCCCAGGTTTCAGAGCTATGCTGG AAACAGCATAGCAAGTTGAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTGCCCCGC TGCGGGTGGCGCGCCGCTGAGCGTAATCTGGAACATCG TATGGGTAGGACGGGGTCTCCTTTTT
Human eGFP-lamin-pegRNA	gCCATGGAGACCCCGTCCCAGGTTTCAGAGCTATGCTG GAAACAGCATAGCAAGTTGAAATAAGGCTAGTCCGTT ATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTATGCA TGCCCCGCTGCGGGTGGCGCGCCGCTGCTTGTACAGCT CGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAA CTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGT CTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTG TCGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGT TCTGCTGGTAGTGGTCGGCGAGCTGCACGCTGCCGTCC TCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCC GTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCT GTTGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCC GTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTT CACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGG TCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGCGC TCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAA GTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGC ACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGG CCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAAC TTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCC CTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGC CGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAA CAGCTCCTCGCCCTTGCTCATGGACGGGGTCTCCAGAT CTTTTTTT

Mouse HA-ACTB-pegRNA	CGTGCACCGCAAGTGCTTCTGTTTCAGAGCTATGCTGGAAACAGCATAGCAAGTTGAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTGGGTGTAAAACGCAGCTCAGTAACAGTCCGCCTAAGCGTAATCTGGAACATCGTATGGGTAGAAGCACTTGCGGTGCATTTTTT

Table 3 The sequences of oligonucleotides used to amplify genomic DNA for TIDE sequencing.

TIDE and Next-Generation Sequencing Primers	
Hist1 Forw	CAAGGGCATCCTGGTGCAAA
Hist1 Rev	GTCTAGGTGGGACACGCAAC
Smad4 Forw	CAGCATCTGGGAATGCTCTCTT
Smad4 Rev	CTTTCTTAGGTCATCCTGCTCAC
WTAP Forw	ACACCTTAAACAAAACATCACACA
WTAP Rev	TTCTGTACGTGTTGATGCTCCA
KRAS Forw	TGTGGTAGTTGGAGCTGGTG
KRAS Rev	TCCTGAGCCTGTTTTGTGTC
Vimentin Forw	CTACCGCAGGATGTTCGGTG
Vimentin Rev	CTGCTATCCTCCAGACAACCG
HA-lamin Forw	CCCTTTCCGGGACCCCT
HA-lamin Rev	TTCAGACTCGGTGATGCGAA
Mouse Actin Forw	AGGGTGTAAAACGCAGCTCA
Mouse Actin Rev	TCACCATCTTGTCTTGCTTTCTT

Chapter 4 Dual Purpose Virus-Like Particles Derived from the Moloney Murine Leukemia Virus

This chapter consists of unpublished data. Beth Rousseau and David Turner planned experiments. Beth Rousseau conducted experiments. Huanqing Zhang assisted with mouse experiments.

4.1 Introduction

The data in Chapter 3 demonstrated that Cas9 VLPs derived from the Moloney Murine Leukemia Virus can be used to deliver enzymes to efficiently edit the genome in cultured cells. In Chapter 3, genomic editing was performed only on single locus after VLP infection. However, there are biological questions that are not addressable by making edits at only one genomic locus, or by delivering only the Cas9 enzyme. For example, a biological effect may not be observed unless multiple redundant genes are disrupted, or paired Cas9 sites can be used to delete a noncoding genetic element. The delivery of multiplexed Cas9 VLPs that could target two or more genes or sites simultaneously could help to solve this problem. Furthermore, there are transgenic mouse lines or cell lines that include Cre-inducible markers, such as our Ai9:RiboTag MDF line (see Chapter 3), where the delivery of Cas9 and Cre simultaneously could overcome the challenge of identifying the Cas9 edited cells within a population. Identification of CRISPR edited cells will likely continue to be an ongoing challenge, as Cas9 cleavage is never complete at any genomic locus. Reasons for cells in a population remaining unedited include variations in chromatin accessibility, the efficacy of RNA-DNA heteroduplex formation with the guide RNA which is a property of the sequence of the guide RNA, and the

efficiency of conformational changes in the HNH and RuvC domains that ensure concerted cleavage of both strands (Jensen et al. 2017; Sternberg et al. 2015). The potential advantages of delivering multiple guide RNAs and/or delivering multiple enzymes simultaneously indicates that there is a need to develop VLPs that can deliver two guide RNAs simultaneously.

4.2 Results

4.2.1 Dual Guide RNA Delivery with Cas9 VLPs

There are several different strategies for using the Cas9 RNP to disrupt gene function. The simplest method is to direct Cas9 to a specific genomic locus so that resulting double-stranded breaks are repaired imperfectly by NHEJ, leading to indels that disrupt protein coding genes through frameshifts. Another method is to employ two different guide RNAs against a pair of target sites flanking a region of interest. The segment of DNA in between the two cut sites can be lost during NHEJ when the DNA on either side is ligated together (Canver et al. 2014). In some cases, a single Cas9 guide against a repeated sequence can be used to create a deletion, reducing the number of guides required to create two or more cuts (for an example, see the Ai9 reporter guide used in Chapter 3). We used this strategy to generate a guide RNA to activate the RiboTag reporter, by targeting the duplicated intron 7 present at the *Rpl22* locus in the Ribotag mouse (Sanz et al. 2009). The Ribotag mouse was created as a tool to help study the transcriptome. *Rpl22* encodes ribosomal protein 22, a small, non-essential protein that is part of the ribosome (O'Leary et al. 2013). When Rpl22 bears an HA-tag, anti-HA immunoprecipitation can be used to collect ribosomes and purify their bound transcripts. The Ribotag can be selectively activated in cells exposed to Cre, or by using Cas9 with what we will call the *Rpl22* sgRNA. To create the RiboTag, the HA tag was placed at the 3' end of the *Rpl22* gene encoded

in an alternate exon 7 and placed after the endogenous *Rpl22* exon 7 which is flanked by duplicate intronic sequences that include LoxP sites. When the native, non-tagged exon 7 is removed by Cre or Cas9, the alternate exon is expressed which contains the HA tag. The Ai9 MDF line is homozygous for the Ribotag at *Rpl22*, and we have been able to activate RiboTag expression with the *Rpl22* sgRNA. If Cas9 and two different guide RNAs can be efficiently delivered to MDFs, it should be possible to investigate the effects of gene disruption on the transcriptome for any gene of interest by using the Ribotag to isolate mRNA from cells containing the Ribotag.

The ability to deliver multiplexed Cas9 VLPs could enable applications such as the simultaneous knockout of two genes, genomic deletions, lineage tracing of edited cells, or the study of transcriptome changes to a population of cells (Hsieh-Feng and Yang 2020). There are two obvious strategies for the simultaneous delivery of Cas9 VLPs with two different guide RNAs. One method would be to prepare two sets of VLPs through separate transfections and then infect cells simultaneously or sequentially. Alternately, a single set of VLPs with both guides loaded onto Cas9 in the same particles could be used. The latter approach could potentially improve co-delivery of both guides to individual cells. The exact number of particles that infect an individual cell is unknown, as is the number of Cas9 proteins per VLP. However, we assume that there are multiple molecules of Cas9 per particle, and the Cas9 proteins should load guide RNAs independently, allowing the generation of VLPs with two different GagCas9/guide RNAs incorporated. One of the two guides we chose for dual guide RNA delivery was against *Wtap*, which encodes Wilms' tumor 1-associated protein, a pre-mRNA-splicing regulator that is indispensable for embryonic development (Fukusumi, Naruse, and Asano 2008). Genes required for development are of particular interest for Cas9 mediated knock-

out in somatic cells because only conditional knockout mouse lines can be created for the study of such genes. The other guide RNA that we chose was the *Rpl22* sgRNA that activates the Ribotag in the Ai9 MDF cell line. The VLPs were made as previously described with the exception of substituting 125 ng of each guide RNA plasmid for the usual 250 ng of one guide RNA plasmid in the transfection.

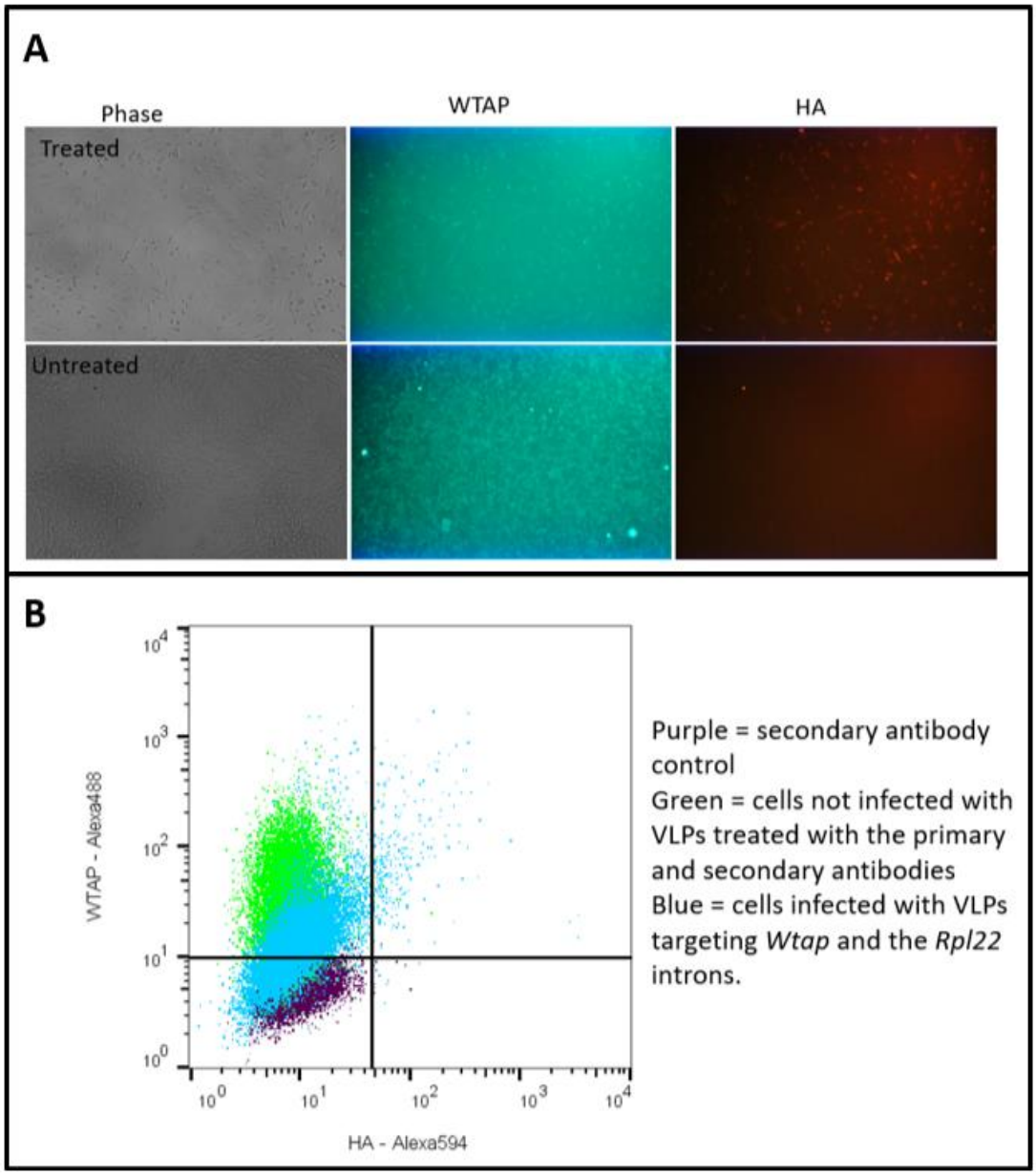

Figure 20 Cas9 VLPs can target two genomic loci simultaneously.
A) MDF cells were infected with Cas9 VLPs programmed to target both WTAP and Rpl22 (top) or cells were untreated (bottom). The cells were fixed in monolayer and stained for WTAP using anti-WTAP and a secondary antibody conjugated to Alexa Fluor 488. The WTAP signal was reduced in a subset of the infected cells compared to the untreated cells (middle image). The cells were also stained for the HA epitope tag using an anti-HA antibody and a secondary antibody conjugated to Alexa Fluor 594. The activated cells appear red (right, top image) whereas the untreated cells do not show any HA signal (lower right image). B) This is an example of data acquired by flow cytometry. MDF cells were infected with Cas9 VLPs programmed to target WTAP and Rpl22, fixed, and stained in suspension. The fluorescence intensities of the cells using channels for Alexa Fluor 488, which is used to detect the anti-WTAP antibody, and for Alexa Fluor 594, which is used to detect the anti-HA antibody, quantified the number of cells that were positive for each antibody. The fluorescence intensity for the indication of WTAP is on the y-axis in arbitrary units and it is plotted against the fluorescence intensity that indicates expression of the HA tag. The population of cells indicated in purple is a control that was analyzed with secondary antibody only. The population in green is a control that was not infected with VLPs but was analyzed with both antibodies. The population in blue was treated with the multiplexed VLPs and was analyzed with both antibodies.

Ai9 MDF cells were treated with multiplexed Cas9 VLPs or left untreated then fixed in monolayer at 72 hours post-infection for assay by immunohistochemistry. Signals from the HA and WTAP antibodies and their associated secondary antibodies revealed cells with reduced WTAP signal or cells exhibiting HA expression, but only a very small fraction of cells that exhibited both WTAP signal loss and HA signal gain (Fig. 20A). Cells were also fixed and stained in suspension for analysis by flow cytometry (Fig. 20B). The same antibodies were used for both monolayer immunohistochemistry and flow cytometry. The anti-WTAP primary antibody was visualized by exposure to an Alexa Fluor 488 conjugated secondary antibody. The HA primary antibody was visualized by exposure to an Alexa Fluor 594 conjugated secondary antibody. Background fluorescence was determined by comparison with cells exposed to both secondary antibodies but not the primary antibodies.

For comparison, VLPs targeting either only *Wtap* or only the *Rpl22* intron were also used to infect Ai9 MDFs in parallel (Fig. 21A, the two leftmost bars). Earlier experiments where MDF cells were infected with GagCO-Cas9 VLPs targeting *Wtap* showed a gene disruption efficiency as detected by TIDE of approximately 60% (Fig. 12B). Flow cytometry reveals a decrease in fluorescence intensity for Wtap in only approximately 31.5% of the cells treated with *Wtap* targeting VLPs. This discrepancy could be due to the relatively high level of background that the anti-Wtap antibody displays, which means that there is an unavoidable one percent overlap between the negative and positive populations. The amount of time it takes for residual Wtap to degrade after gene disruption also may be longer than the seventy-two hours between infection and the time the cells were fixed, leading to an underestimate of disruption at the protein level (since TIDE directly measures the change in DNA sequence in the target). In addition, it is possible that heterozygous gene disruption could allow Wtap protein levels to be maintained at

levels close to normal. Some combination of these factors is likely to explain the smaller fraction of cells that have reduced Wtap protein expression after Cas9 disruption, when detection is by immunohistochemistry/flow cytometry. We could not perform a similar comparison between staining efficiency and TIDE for *Rpl22* because the RiboTag activating deletion is too large to be analyzed by TIDE.

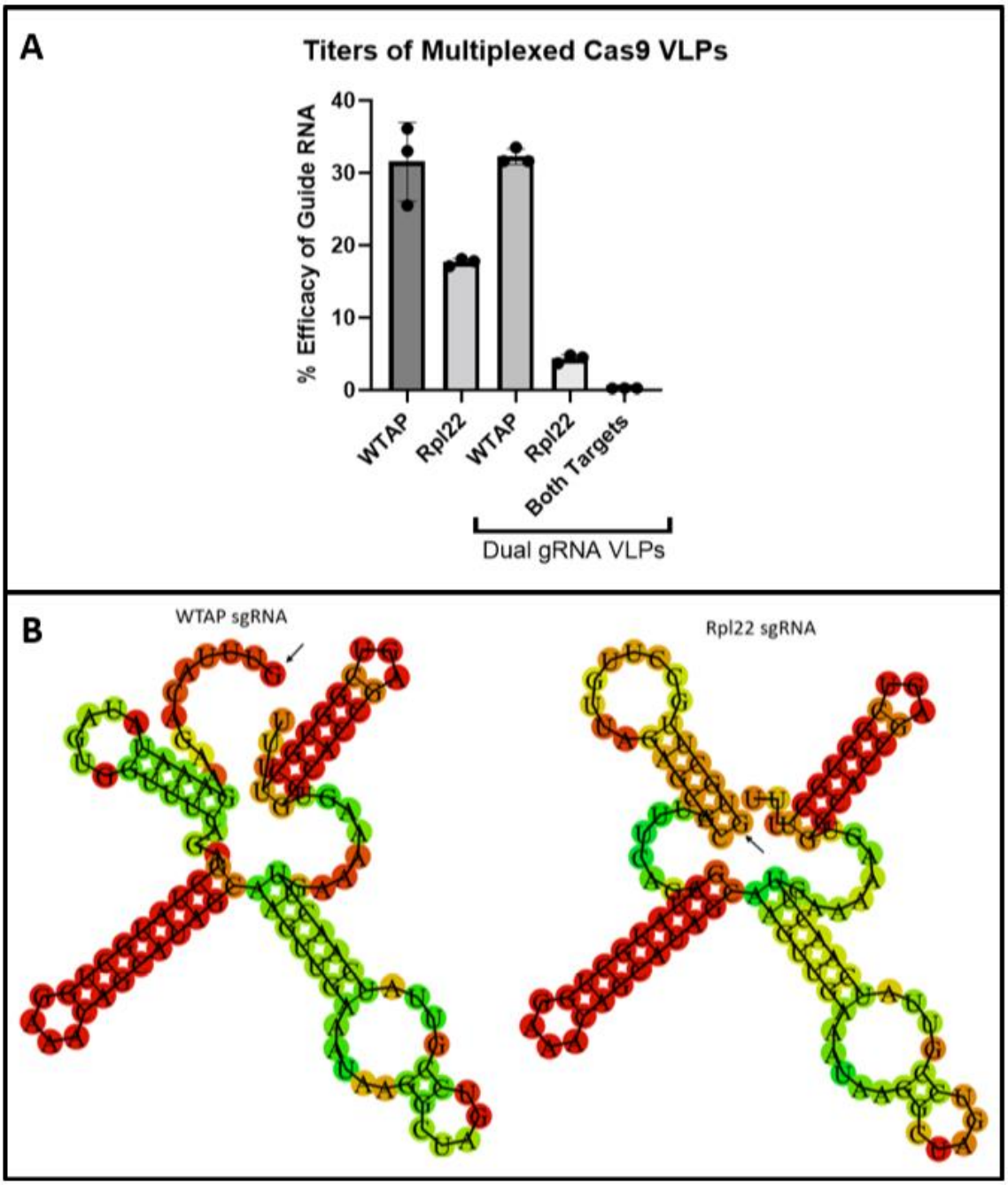

Figure 21 Dual delivery of Rpl22 and Wtap sgRNA is less efficient than the independent delivery of either guide. A) Flow cytometry reveals that Cas9 VLPs can deliver either the Wtap or Rpl22 sgRNAs with reasonable efficiencies to MDF cells. The rightmost three bars indicate the percentage of cells of that exhibit either loss of the WTAP protein, activation of expression of the Rpl22-HA tagged protein, or both events after treatment with VLPs made with both guide RNAs. B) RNAfold predicts the WTAP sgRNA (left) adopts a typical sgRNA structure with the 5' end of the guide (indicated by arrow) free. However, RNAfold predicts that the Rpl22 sgRNA (right) adopts a structure that is dissimilar, with the 5' end guide region (indicated by the arrow) base-paired with structural elements of the sgRNA. The alternate structure of the Rpl22 sgRNA alters the structure for regions of the sgRNA bound by Cas9, and may reduce loading onto Cas9 relative to other sgRNAs, including the Wtap sgRNA.

The results of flow cytometry suggest that the ability of multiplexed Cas9 VLPs to genome edit a target is lower than in cells treated with VLPs containing only one guide RNA (Fig. 21A). About 17.7% of the MEFs treated with Cas9 VLPs containing only the *Rpl22* gRNA displayed HA activation (Fig. 21A, Rpl22). About 31.5% of the MEFs treated with Cas9 VLPs containing only the *Wtap* gRNA displayed a decrease in WTAP staining (Fig. 21A, WTAP). When both guide RNAs were combined into one transfection to make VLPs, the *Wtap* gRNA showed no sign of decrease in activity. However, the number of cells edited by the *Rpl22* gRNA appeared to decrease to only 4.3%. Furthermore, the number of cells that appeared to be negative for WTAP and positive for Rpl22 was very small at about 0.23% of the cells. One possible explanation for this result is that there is competition between the sgRNAs for loading onto Cas9 in the VLP producer cells, with the Wtap sgRNA strongly favored to load onto Cas9. The RNAfold webserver (Gruber et al. 2008) predicts the *Wtap* sgRNA to fold as is typically expected of a guide RNA with the guide portion at the 5' end of the guide RNA being unpaired and accessible for target recognition (Jensen et al. 2017). In contrast, RNAfold predicts that the *Rpl22* sgRNA preferentially folds into a form in which the guide sequences at the 5' end of the sgRNA pair with internal sequences in the sgRNA framework, sequences that normally interact with Cas9, leading to an alternate secondary structure (Fig. 21B). The alternate folding likely interferes with loading of the Rpl22 sgRNA onto Cas9, allowing the Wtap sgRNA to outcompete the Rpl22 sgRNA for Cas9 loading.

4.2.2 Cas9 VLP Efficacy Is Not Improved by the Addition of the ET-208 Nuclear Export Sequence to the Guide RNA

An early observation from our experiments infecting Ai9 MDF cells with Cas9 VLPs loaded with the Ai9 sgRNA was that the amount of guide RNA plasmid transfected into producer cells can be a limiting factor for Cas9 VLP titer. In our plasmids, the transcription of sgRNA for Cas9 genome editing is driven by the constitutively active RNA polymerase III U6 promoter (Cong et al. 2013). Cas9 mRNA expression is driven by the simian CMV (sCMV) promoter; the mRNA includes a human ubiquitin C intron to improve nuclear export/expression, and utilizes the SV40 late polyadenylation site to add a polyA tail (Radici et al. 2013; Proudfoot 2011; G.-Y. Kim et al. 2011). Ideally, there should be enough sgRNA to load all Cas9 proteins before they are incorporated into VLPs because Cas9 proteins without an sgRNA will not be directed to cleave a specific genomic locus in VLP infected cells and also may compete with sgRNA/Cas9 complexes for VLP loading. We constructed additional plasmids to investigate the effect of changing the ratio of U6 expression cassette to the sCMV driven Cas9 gene on VLP efficacy. pCS7-Ai9-GagCas9 was made by inserting the U6 cassette driving the expression of the Ai9 sgRNA upstream of the sCMV promoter and oriented in the same direction. This means that VLPs made solely with that plasmid will have a one-to-one ratio of the U6 driven sgRNA and sCMV driven Cas9 gene, and that there is no need for cotransfection with a U6 sgRNA plasmid to generate VLPs that can activate tdTomato expression in Ai9 MDF cells. We made Cas9 VLPs (see Methods) with pCS7-Ai9-GagCas9 cotransfected with either the U6-Ai9-sgRNA plasmid or with an equal amount of noncoding plasmid DNA. When adjusted for plasmid sizes, these transfections included the plasmid templates at a molar ratio of either one U6 sgRNA template to one sCMV Cas9 template, or 11.4 U6 sgRNA templates to one sCMV Cas9 template,

respectively. The VLPs were used to infect Ai9 MDF cells as previously described (See Chapter 3). The Cas9 VLPs made with a ratio of one U6 sgRNA template to one sCMV Cas9 template had a titer of only 380 IU/mL whereas Cas9 VLPs made with a ratio of 11.4 U6 sgRNA templates to one sCMV Cas9 template had a titer of 1,333 IU/mL (Fig. 22A). Thus, when the molar ratio of the U6 sgRNA template to sCMV Cas9 template was increased by a factor of ~11, the titer increased by a factor of 3.5. This result suggests that the amount of sgRNA being produced did not saturate the loading of the sgRNA onto Cas9 at the one-to-one template ratio. Furthermore, the nonlinear increase in the titer compared to the increase in the ratio of the promoters suggests that 11.4 U6 to 1 sCMV template likely reached or exceeded a point of saturation where all or nearly all Cas9 proteins in VLPs are bound to sgRNAs. This is significant because although this amount of sgRNA would appear to saturate Cas9, the vector expressing the guide RNA occupies 50% of the total mass of DNA transfected into cells in order to make VLPs, with 89% of that mass consisting of extraneous vector backbone. If the amount of guide RNA vector that needed to be transfected into cells was less because the guide RNA was more loaded onto Cas9, then it would be possible to replace that mass with a vector that expresses GagCas9, GagPol, or VSV-g, which may improve the stoichiometry of those vectors and increase production of Cas9 VLPs. Due to the limit on the total amount of DNA that can be forced into cells, it seems that one possible solution to sgRNA underloading may be to improve the efficiency with which Cas9 and sgRNA meet in the cell, prior to VLP formation in order to reduce the overall amount of guide RNA vector that must be co-transfected.

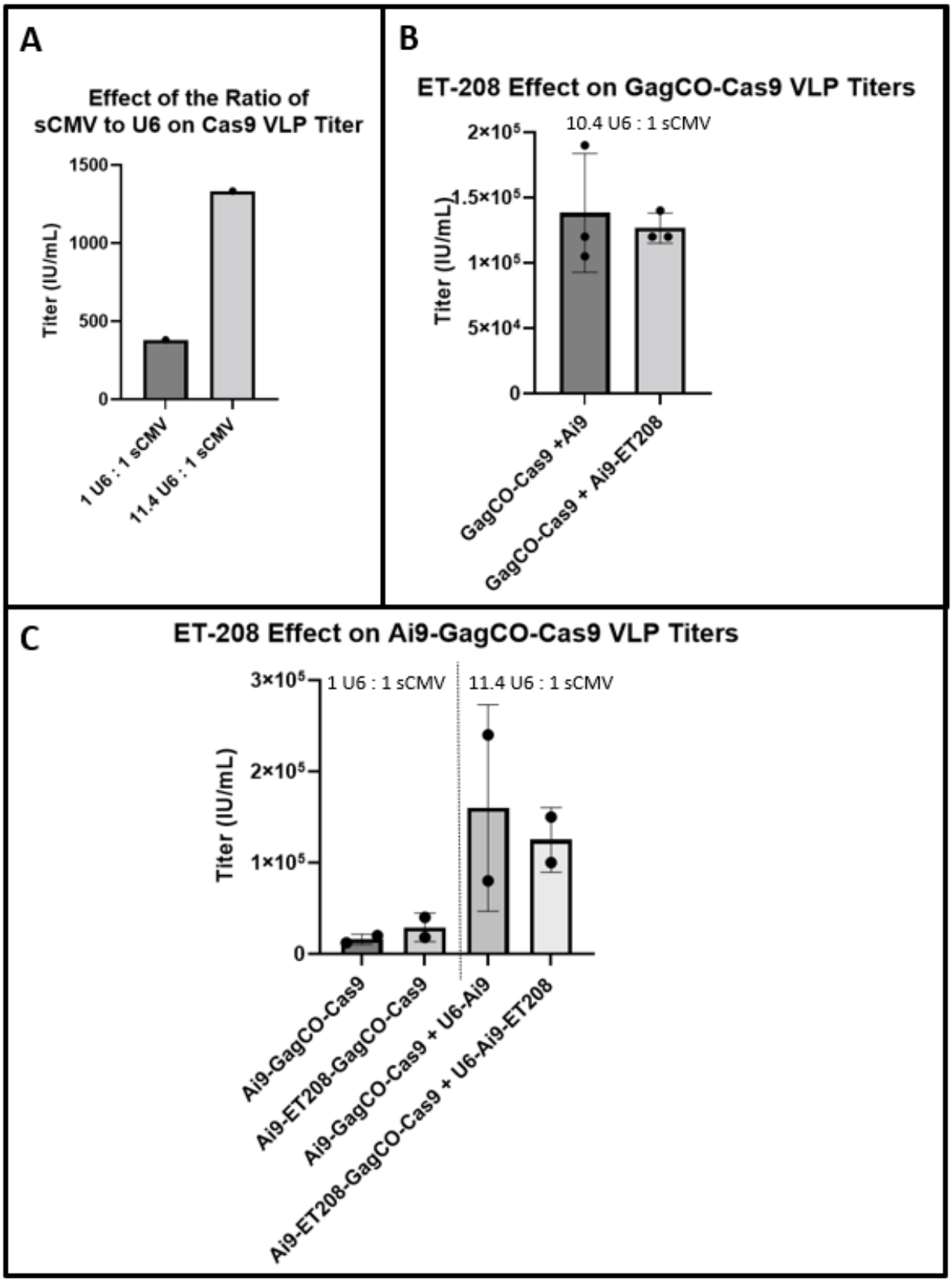
A
Effect of the Ratio of sCMV to U6 on Cas9 VLP Titer
Titer (IU/mL)
1500
1000
500
0
1 U6 : 1 sCMV
11.4 U6 : 1 sCMV
B
ET-208 Effect on GagCO-Cas9 VLP Titers
10.4 U6 : 1 sCMV
Titer (IU/mL)
2×10⁵
1.5×10⁵
1×10⁵
5×10⁴
0
GagCO-Cas9 +Ai9
GagCO-Cas9 + Ai9-ET208
C
ET-208 Effect on Ai9-GagCO-Cas9 VLP Titers
1 U6 : 1 sCMV
11.4 U6 : 1 sCMV
Titer (IU/mL)
3×10⁵
2×10⁵
1×10⁵
0
Ai9-GagCO-Cas9
Ai9-ET208-GagCO-Cas9
Ai9-GagCO-Cas9 + U6-Ai9
Ai9-ET208-GagCO-Cas9 + U6-Ai9-ET208

Figure 22 Addition of the ET208 aptamer nuclear export sequence to the sgRNA does not alter the efficacy of Cas9 VLPs.
A) Cas9 VLPs were made using pCS7-Ai9-GagCas9 that has a U6 cassette driving the expression of the Ai9 sgRNA on the same plasmid as GagCas9 which is driven by the sCMV promoter. The cotransfection of pCS7-Ai9-GagCas9 with noncoding DNA produced VLPs with a 1:1 ratio of U6 guide RNA to sCMV Cas9 templates. Cotransfection of pCS7-Ai9-GagCas9 with a U6 vector encoding the Ai9 sgRNA increased ratio of U6 guide RNA template to sCMV Cas9 templates from 1 to 11.4. Increasing the ratio of U6 guide RNA template to sCMV Cas9 templates increased the titer of the Cas9 VLPs by a factor of 3.5 when Cas9 VLPs were used to infect MDF cells. B) VLPs made with pCS7-GagCO-Cas9 activated the tdTomato reporter in infected MDFs with a similar titer when made with a U6 vector expressing the Ai9 sgRNA or expressing the Ai9 sgRNA fused to ET-208. C) The Ai9 sgRNA, with or without the ET-208 aptamer, was added on to pCS7-GagCO-Cas9. This allowed fine tuning of the ratio between the U6 and sCMV promoters by creating VLPs with or without the addition of a U6 plasmid encoding the coordinating Ai9 sgRNA. The ET-208 aptamer has no effect on Cas9 VLP titer when it is expressed at levels that are either subsaturating (left two bars) or saturating (right two bars) for the Cas9 protein.

The U6 promoter generates sgRNA in the nucleus, while the GagCas9 fusion protein is translated in the cytoplasm and then localized to the plasma membrane for export. It is possible that GagCas9 fusion proteins enter and exit different compartments of the cell after translation. We know that at least some of the GagCas9 fusion protein enters the nucleus of a transfected cell, as gene editing can be detected in both HEK293 cells and MDF cells transfected with plasmids encoding GagCas9 and sgRNAs (data not shown). However, since GagCas9 is translated in the cytoplasm, it is possible that protein may be loaded with sgRNA in the cytoplasm before trafficking to the membrane for VLP production or to the nucleus. If so, the sgRNA must leave the nucleus after it is transcribed to reach Cas9. RNA trafficking out of the nucleus can be increased by specific RNA sequences (Czaplinski 2014). One such RNA sequence is a 77 nt synthetic aptamer called ET-208 (Lund and Dahlberg 2006; Grimm, Lund, and Dahlberg 1997). The aptamer binds Exportin5 and is efficiently exported from the nucleus of *Xenopus* oocytes. Exportin5 mediates nuclear export of various small RNAs, including during miRNA biogenesis by binding pre-miRNAs and facilitating their transport through the nuclear pore complex into the cytoplasm (K. Wu et al. 2018). ET-208 does not need free 5’ or 3’ ends to be exported, a factor which lead us to believe that it could likely be appended to a sgRNA and

still function. We assessed the titers of Cas9 VLPs made with the Ai9 sgRNA, either with or without ET-208 appended to the 3' end, using the MDF reporter line. We found no significant difference in the titers (Fig. 22B), indicating that ET-208 likely did not increase sgRNA loading or Cas9 activity. However, our initial tests suggested that Cas9 VLPs made with the usual ratios of 11.4 U6 to 1 sCMV are saturated with sgRNA, which means that ET-208 could still have an effect on Cas9 VLP titer, an effect that would only be seen at subsaturating levels of sgRNA loading onto Cas9. To ensure that the result of the experiment in Figure 22B was not an artifact of sgRNA saturating conditions, we also cloned the U6 cassettes onto the GagCO-Cas9 vector to allow varying the ratio of the U6 sgRNA expression cassette to the sCMV driven Cas9 (Fig. 22C). The addition of ET-208 did not increase Cas9 function as assessed at either the 1:1 ratio (right bars) or 11:1 ratio (left bars). Nonetheless, these experiments indicate that the sgRNA can be elongated on its 3' end and still be successfully loaded into VLPs and function. Prime editing guide RNAs are longer on their 3' ends than standard sgRNAs and evidence that a guide RNA with a similar length 3' extension to a pegRNA could be loaded into a VLP supported our decision to pursue prime editing delivery via VLP (see Chapter 3.2.3).

4.2.3 Cre-Cas9 VLPs Can Deliver Two Enzymes Simultaneously

The Ai9:Ribotag mouse suggested an opportunity for the lineage tracing of cells edited by Cas9. Cre could be used to mark cells by activating the tdTomato and RiboTag reporters, while Cas9 with an appropriate sgRNA could be used simultaneously to knock-out a specific target gene in the marked cells. The Ai9 MDF cell line provided a system to test this idea, as it had proven useful for initial tests of Cre and Cas9 VLPs (see Chapter 3.2.1).

We initially envisioned using VLPs containing both Cre and Cas9. These could potentially be made by co-transfection of Gag-Cre and GagCas9 plasmids to produce VLPs.

However, this approach did not work because of a large imbalance in the efficacies between these two enzymes (data not shown). Cre activation of the tdTomato reporter was highly efficient even when only 1% by mass of the DNA transfected into producer cells consisted of pCS7-GagCO-Cre. Cas9 is an inherently less efficient enzyme because it is limited by the efficacy of the sgRNA and cleavage events that are repaired in frame without a stop codon will not result in gene knockout. High efficiency of tdTomato activation and low activity of Cas9 led to a low overlap between cells that were edited in culture, and suggested that efficient in vivo editing in marked cells was not practical. Another possible reason for the disparity in Cre and Cas9 efficacies may arise from Gag-Cre's ability to form particles without the presence of Gag (see Section 4.2.7). If Gag-Cre can polymerize independently whereas GagCas9 needs some amount of Gag in order to create a geometry that supports particle formation, then GagCas9 might even be preferentially excluded from particles that contain Gag-Cre. In order to mitigate this problem, we created a chimeric polyprotein of Gag, Cre, and Cas9 (Fig. 23A).

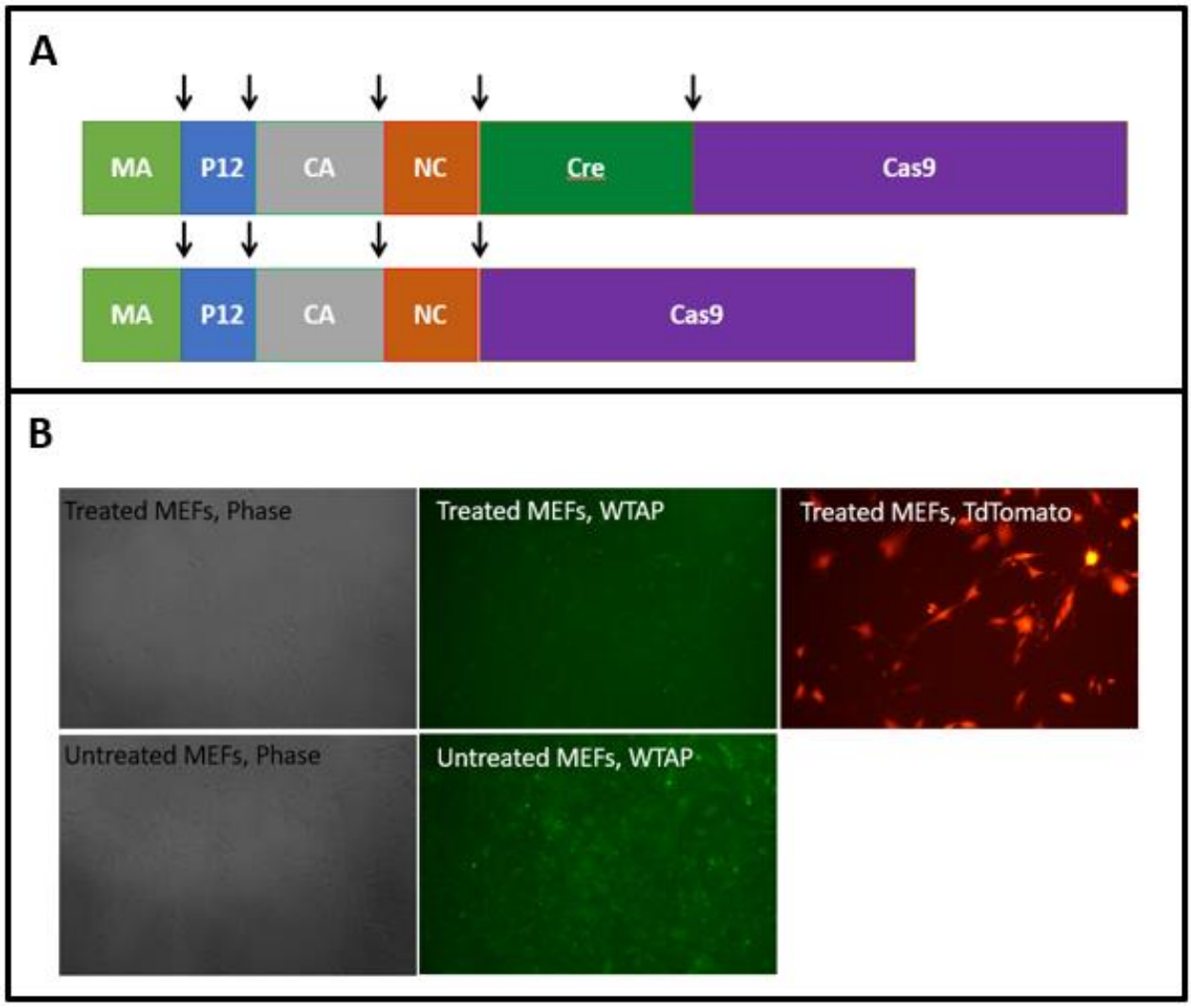

Figure 23 Gag-Cre-Cas9 VLPs can deliver both enzymes to MDF cells.
A) Schematic of the proteins for Gag-Cre-Cas9 and GagCas9. Cleavage sites are indicated by arrows. B) MDFs infected with Gag-Cre-Cas9 VLPs were fixed in monolayer 72 hours after infection. tdTomato was activated by Cre in a subset of cells. Cells were stained for WTAP which decreased in signal in a subset of the cells.

The inclusion of Cre and Cas9 on the same polyprotein reduced Cre's activity to closer to the activity of Cas9. We do not know if this is because Gag-Cre-Cas9 is less efficiently loaded into particles or because Cre function is reduced when delivered in this form. There are protease cleavage sites that could potentially separate Cre from both Gag and Cas9, but we did not assay the VLP supernatants to determine the level of processing of Cre. If the amount of Cre in the VLP supernatants was indeed low, assaying processing may not be feasible because the protein

would not be detectable on a western blot. Furthermore, we do not know if the protein containing Cre that is being loaded into VLPs is full length Gag-Cre-Cas9 or if it is a truncation that leaves Cre intact but may omit all or part of Cas9. Nonetheless, we could assess the effects of VLPs produced with the Gag-Cre-Cas9 fusion protein in combination with the Wtap sgRNA by scoring the Ai9 reporter and expression of Wtap protein in infected Ai9 MDF cells.

We infected Ai9 MDFs with VLPs that contained the Gag-Cre-Cas9 fusion and an sgRNA against Wtap, then assessed disruption of Wtap by Cas9 and activation of the tdTomato reporter by Cre. We counted approximately 500 cells from different fields of view in monolayer to find that over 50% of the cells expressing tdTomato also were negative for WTAP, but the overlap was still imperfect (Fig. 23B). We attempted to assay the cells by flow cytometry, but found that non-specific staining of the primary WTAP antibody was so high that it was not possible to differentiate signal from noise. Our results indicate that CreCas9 VLPs can deliver both enzymes during the same infection. However, the incomplete overlap between reporter and gene disruption, as well as the difficulty in balancing the activity between the two enzymes suggests that these VLP designs would likely prove to be inadequate for obtaining high overlap of Cre cell marking with Cas9 editing.

4.2.4 Mutating the Cleavage Site Between Gag and Cas9 Reduces Cas9 VLP Titer

In considering how to improve VLP delivery of Cas9, either by itself, in multiplexed VLPs or for co-delivery of Cre and Cas9, we considered whether it might be possible to improve Cas9 delivery to the nucleus of target cells. Little is known about exactly what happens to Cas9 VLPs after they are endocytosed by target cells. It is possible that improved nuclear delivery of Cas9 could improve Cas9 editing. Risco et. al. found that immunolabeled NC proteins (the most C-terminal protein in the Gag polyprotein) entered the nucleus in non-dividing cells, possibly

because NC itself has nuclear localization signals (Risco et al. 1995). Cas9 is fused after NC in GagCas9, with the two proteins separated by a viral protease cleavage site (see Chapter 3). We wondered if mutating the protease cleavage site between NC and Cas9 (QTSLL↓TLDDQ), leaving Cas9 fused to NC but separated from the rest of the Gag polyproteins could increase Cas9's ability to enter the nucleus of target cells. Although there were no mutations known to specifically inhibit viral protease cleavage at the site separating NC and Cas9, there were mutations at other viral protease target sites that were known to inhibit cleavage (Oshima et al. 2004). We used those as a guide in choosing to make the mutant site QTSLN↓TLDDQ, L534N, which we called Gag-Asn-Cas9. VLPs were made with the GagCO-Cas9 vector either with or without the Asn mutation and loaded with the Ai9 sgRNA. VLPs were used to infect MDF cells in order to activate the tdTomato reporter. Counting of tdTomato positive cells indicated that the titer of the mutant Gag-Asn-Cas9 was about two-fold lower than the original GagCas9 protein (Fig. 24). This reduction in titer suggests that separation of Cas9 from Gag does improve VLP titer slightly, although it is not necessary for Cas9 delivery and function. However, we did not investigate whether the protease site mutation functioned as predicted, or whether equal amounts of GagCas9 were incorporated into VLPs. It is possible that protease cleavage at the NC-Cas9 junction is not completely abrogated, but only reduced. It is also possible that the folding of the protein is changed and the new structure alters the loading of Gag-Asn-Cas9 into VLPs. Since our goal was to improve Cas9 delivery we did not investigate these possibilities further.

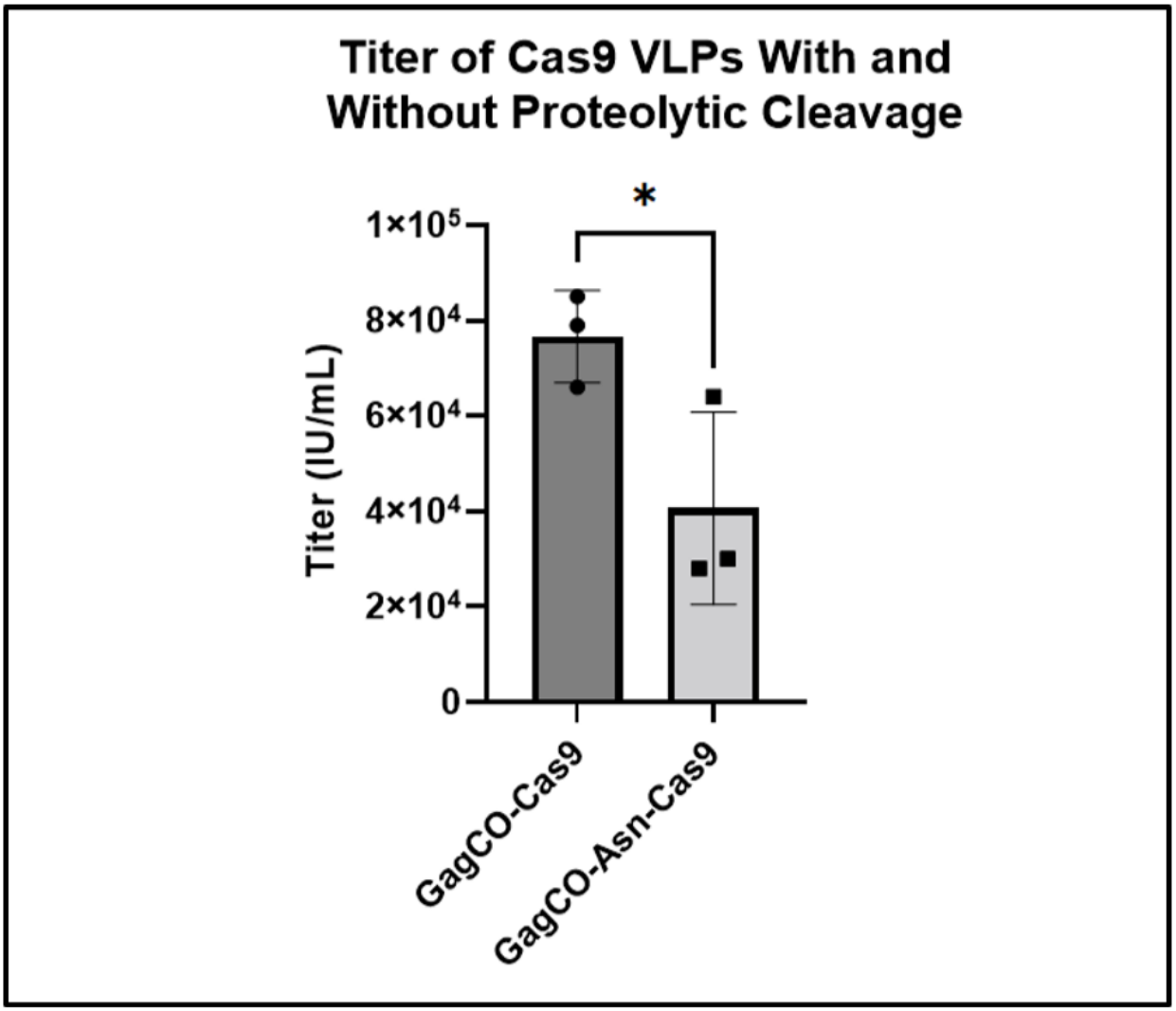

Figure 24 Mutating the cleavage site between Gag and Cas9 reduces Cas9 VLP titer.
We created a GagCas9 construct with point mutation L334N which we called Gag-Asn-Cas9. We made VLPs loaded with the Ai9 sgRNA with a construct expressing either the original GagCas9 or expressing Gag-Asn-Cas9 and applied these to MDF cells. Quantification of tdTomato positive cells indicated that the titer of the Gag-Asn-Cas9 VLPs was slightly reduced compared to GagCas9 VLPs.

4.2.5 Comparison of Three Different Protease Cleavage Sites in GagCO-Cas9

As described in the discussion of Chapter 3, Banskota et. al. recently reported the creation of engineered VLPs that deliver the base editor. They tested a few different modifications to the design and formulation of VLPs in order to improve their efficacy. They found that the efficiency of cleavage between Gag and the base editor impacted the efficiency of base editing. A higher cleavage efficiency correlated to increased base editing. Similarly (Fig. 13BC, Fig. 24), we have observed that cleavage of Cas9 from Gag appears important in improving Cas9 functionality. This may be true for the prime editor as well. We considered that is possible that Gag sterically hinders binding or enzymatic activity if it is not cleaved efficiently from the fusion protein by the viral protease. Banskota et. al. (2022) tested four different cleavage sites and found that the site that permitted the second highest cleavage efficiency as observed by western blot and the highest level of base editing in HEK293T cells was the site TSTLL/MENSS. The slash indicates the site of cleavage. That site is not a perfect match for either Friend or Moloney cleavage sites, but it most closely corresponds to the site in between RT and IN in the Moloney Murine Leukemia Virus which is TSTLL/IENSS. Banskota et. al. surmised that a balance between cleavage efficiency and avoiding premature cleavage of Gag and the fusion protein is the reason why the second-best cleavage site yielded eVLPs with the greatest activity. It is also possible that the three-dimensional geometry of the cleavage site impacts the ability of the Gag fusion proteins to be incorporated into VLPs.

The data presented in Figures 13BC as well as Figure 24 led us to test different protease sites. We built two more constructs in addition to the original GagCO-Cas9 construct that contained the native NC/Prot cleavage site. The three cleavage sites that we tested were the native site NC/Prot, the MA/P12 site (modified so that it would exactly match the site in the

Nanoblades construct (Mangeot et al. 2019)), and the Prot/RT site (Fig. 25A). Part of our motivation for testing these sites was based on their reported cleavage efficiencies determined in vitro on peptides using a recombinant viral MLV protease (Fehér et al. 2006). We prepared VLPs as described in Chapter 3 Methods and infected Ai9 MDF cells in order to calculate a titer for each construct (Fig. 25B). We found that the titer for the construct containing the MA/P12 site appeared to be the highest. This is different from what Banskota et. al. (2022) found, but we do not know how codon optimization of the Gag-fusion protein affects translation, or what other subtle differences, such as the presence of a specific linker or the identity of the fusion protein may alter the end result of the measured efficiency of VLP delivery.

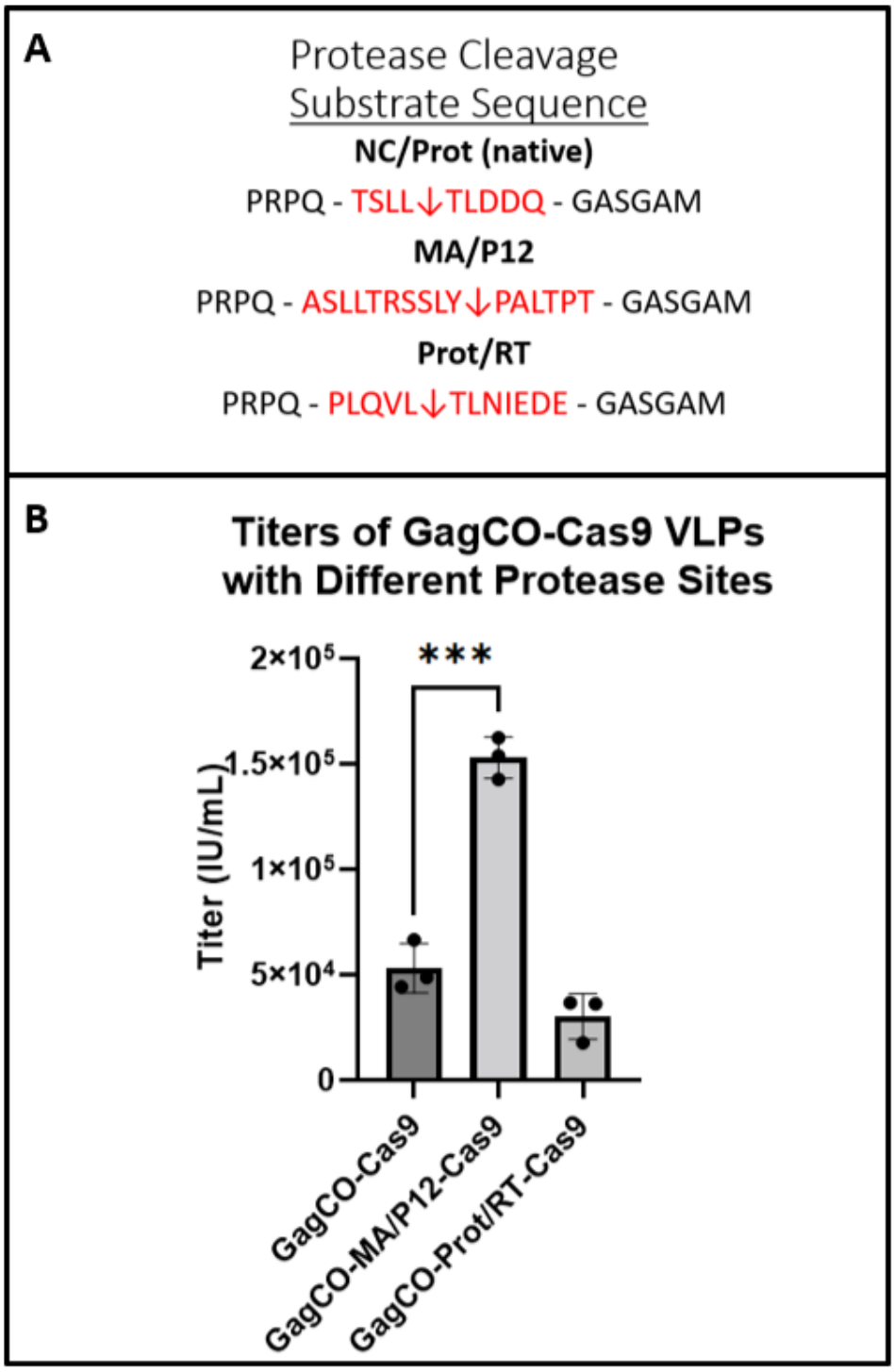

Figure 25 Using different viral protease sites between Gag and Cas9 alters the titer of Cas9 VLPs.
A) The exact sequences changed for each of the three protease sites are highlighted in red. The flanking sequences are part of GagPr65 (left) and the short, flexible linker preceding the Cas9 fusion protein (right). B) Cas9 VLPs loaded with the Ai9 sgRNA were made as described in Chapter 3 and used to infect Ai9 MDF cells where they activated the tdTomato reporter. Titers were calculated as previously described in Chapter 3.

4.2.6 The Effect of Altering the Stoichiometry in Cas9 VLPs Made with Gag-Stop-Prot

The stoichiometry of MLV Gag to GagPol is observed to be 95:5 at both the protein expression level and level of protein incorporation into virions (Johnson et al. 2014). While the composition of Cas9 VLPs has to include some ratio of Gag to GagPol to GagCas9, the optimal ratio of these may only be determined by empirical methods. We found that we could substitute Gag-Stop-Prot for GagPol and achieve a Cas9 VLP titer only two-fold lower than that of Cas9 VLPs made with GagPol (see Chapter 3.2.2). However, the mass of the Gag-stop-Prot vector transfected was the same as GagPol, despite the fact that the Gag-Stop-Prot construct is only two-thirds the size of the GagPol construct, yielding a different molar ratio of the plasmids. We made Cas9 VLPs again using the same constructs except we adjusted their molar ratios as depicted in Figure 26A. The molar ratio of the GagCas9 expression vector to Gag-Stop-Prot expression vector was reduced to 1:1.6 which is closer to the original GagCas9 to GagPol vector ratio which was 1:1.75. In order to change the ratios and still keep the total amount of DNA transfected into HEK293T cells the same, we had to increase the amount of the GagCas9 vector from 85 ng to 175 ng and the accessory vectors from 140 ng to 50 ng (see Methods). The vectors expressing the Ai9 sgRNA and VSV-g were kept at their normal levels. VLPs were made as described previously and used to infect Ai9 MDF cells and assayed by flow cytometry at 72 hours as previously described in Chapter 3. The Cas9 VLPs made with GagPol still retained an approximately two-fold increase in efficacy compared to those generated with Gag-Stop-Prot VLPs. However, the overall efficacy of the VLPs diminished. In the original experiment (See Fig. 13C), Cas9 VLPs made with GagPol had a titer of ~$7.8x10^4$ IU/mL, whereas the new formulation brought the titer to ~$6.3x10^4$ IU/mL (Fig. 26B). Comparisons between titers of VLPs done in different experiments on different days can vary by up to two-fold or occasionally even

more because of the exact density of the producer cells when they are transfected, the quality of the plasmid DNA preparation (e.g. free of *E. coli* genomic DNA and RNA, endotoxins, etc.), and possibly other subtle factors that remain unidentified. As a result, precise comparisons of titers between experiments can be difficult. Nonetheless, this pilot experiment to determine whether adjusting the stoichiometry between GagCas9 and the other expression vectors improved VLP titer did not yield encouraging results. It is possible that adjusting the vector ratios differently or the overall mass of each vector differently could still produce a better outcome, as shown possible by Banskota et. al. (2022) with their base editing eVLPs.

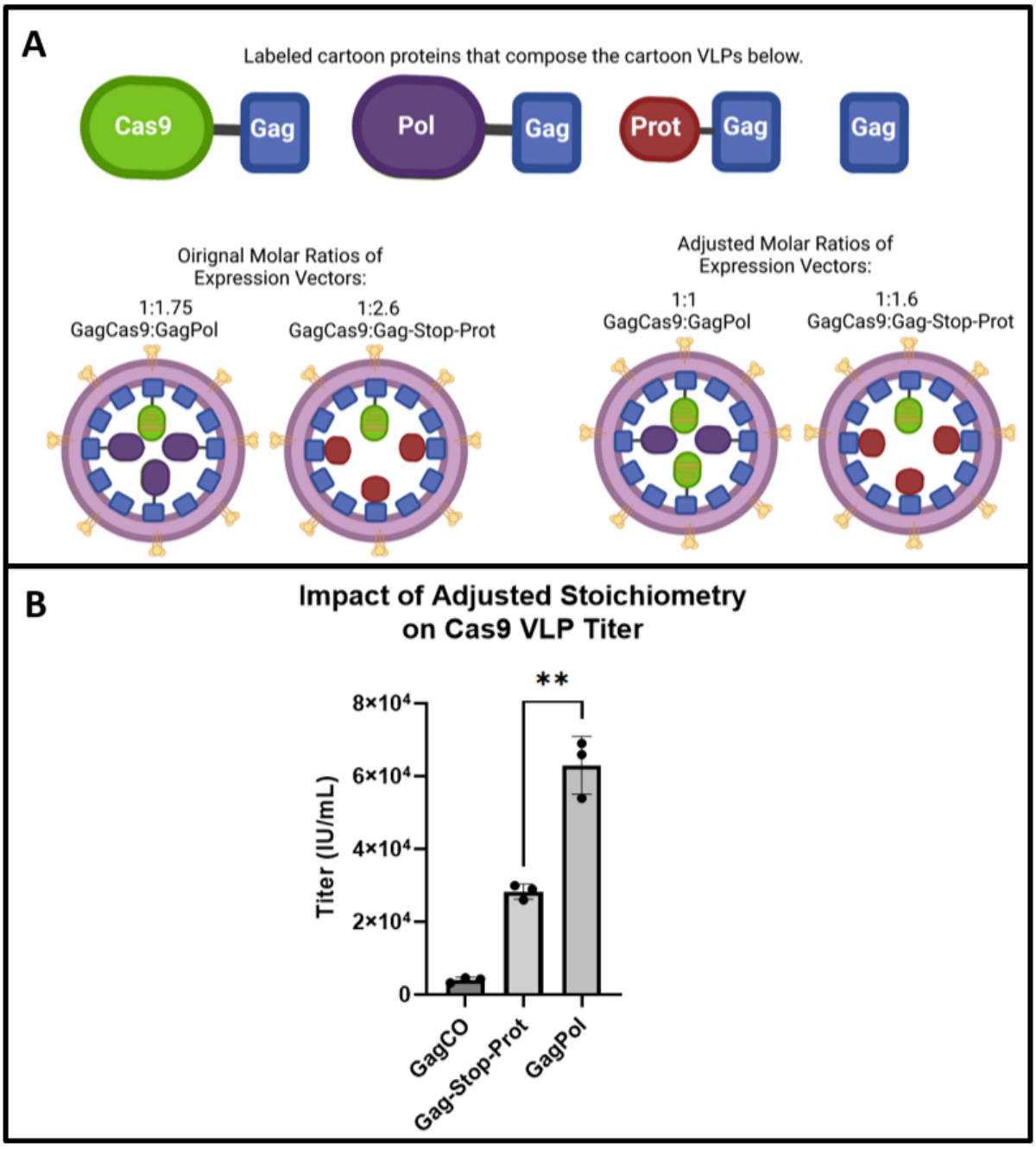

Figure 26 Adjusting the stoichiometry of VLP components did not improve Cas9 VLP titer.
A) Schematic of Cas9 VLPs made with GagPol or Gag-Stop-Pol with the stoichiometry used in the original experiment shown in Figure 13C indicated in comparison with the adjusted stoichiometry used in B. Although the number of molecules shown in the schematic is too small to accurately express the ratio of one molecule to another, we believe that it may be possible to affect the composition of GagCas9 to GagPol or other loaded proteins by varying the relative amounts of expression vectors that are transfected into HEK293T cells (see text). B) Cas9 VLPs were made using the adjusted ratios of the indicated vectors. The GagCO vector is very similar in size to the Gag-Stop-Prot vector and it can be assumed that it also was transfected in a 1:1.6 ratio with the GagCas9 vector. The titers of Cas9 VLPs made with Gag-Stop-Prot is still about two-fold lower than the titer for Cas9 VLPs made using GagPol. 26A was created with Biorender.com.

4.2.7 MMLV VLPs Deliver Cre recombinase Efficiently

Cre recombinase is an enzyme capable of excising DNA in between two LoxP sites (Madisen et al. 2010, 9). This can be used for many applications including activation tdTomato in the Ai9 MDF cell line. We fused NLS-Cre after the non-codon optimized Gag with the native protease site to enable their separation (Fig. 27A). We made VLPs using only VSV-g and Gag-Cre vectors or in conjunction with the vector expressing Gag/GagPol (see Methods). We found that the VLPs made without the vector expressing Gag/GagPol yielded lower titer VLPs than the VLPs made only with Gag-Cre (Fig. 27B). This was surprising as GagCas9 expression with VSV-g and guide RNA yielded very low titers (data not shown). This suggests that the geometry of Gag likely does not sterically hinder the multimerization of Gag-Cre, enabling it to form particles in the absence of Gag Pr65. Also, this implies that Cre must not require cleavage by the viral protease in order to function.

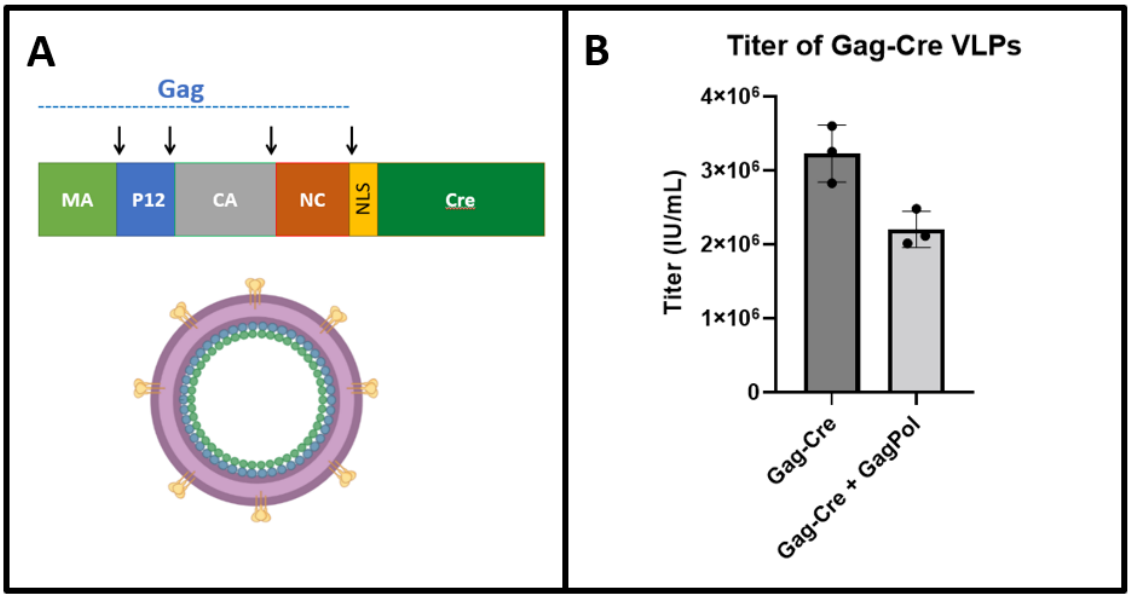

Figure 27 VLPs efficiently deliver Cre into Ai9 MDF cells.
A) Schematic of the Gag-Cre construct as well as a cartoon of a VLP made solely of Gag-Cre molecules with Gag (blue circles) and Cre (green circles) able to multimerize side-by-side with little to no steric hindrance. B) Cre VLPs were made with or without the cotransfection of pCS7-GagPol. The VLPs made without GagPol appeared to have a higher titer than those made with GagPol. 27A was created with BioRender.com.

4.2.8 MMLV VLP Delivery into Embryonic Murine Retinas In Vitro

One of our goals was to use VLPs to deliver Cas9 into murine retinas to disrupt genes involved in retinal development. In particular, we were interested in *Foxn3* which encodes the forkhead box protein N3 (Foxn3), a transcriptional repressor (Scott and Plon 2005). *Foxn3* functions in interneuron formation in the developing mouse retina (H. Zhang et al. 2022), and is embryonically lethal when deleted (Samaan et al. 2010). There is no conditional mouse model for the study of *Foxn3* disruption in retinal progenitor cells. In order to knock-out *Foxn3* in retinal progenitor cells, Zhang et. al. electroporated plasmid vectors expressing Cas9 and a guide RNA into embryonic murine retinas. They found that disruption of *Foxn3* by CRISPR in embryonic murine retinas leads to an increase in the number of amacrine cells. Conversely, ectopic overexpression of *Foxn3* in retinal explants decreases the number of amacrine cells generated and reduces cell cycle exit of retinal progenitor cells. Furthermore, Zhang et. al. identified an upstream regulator of *Foxn3*, miR-216a/b-5p. However, the genomic loci that Foxn3 binds in order to prevent amacrine cell differentiation and maintain proliferation have not been identified. Our concept for elucidating the targets of Foxn3 was to pursue VLP co-delivery of Cre and Cas9 into embryonic retinas harvested from Ai9:Ribotag mice. Cre would be used to activate the tdTomato reporter and the Ribotag. Cas9 would be used to disrupt *Foxn3* with a guide RNA previously validated by our lab (H. Zhang et al. 2022) or loaded with a non-targeting guide RNA. The Ribotag would allow purification of mRNA from specifically from the electroporated cells, allowing a comparison between the transcriptomes of cells with and without *Foxn3* disruption. Transcripts of the genes targeted by Foxn3 would be derepressed and increased in cells treated with the Cre-Cas9 VLPs targeting *Foxn3* compared to the control cells. If the efficiency of the Cre-Cas9 delivery by VLP were low, cells could be sorted using the

tdTomato fluorescent reporter prior to the collection of the transcriptome using the Ribotag in order to enrich for the number of cells that have been edited by Cre and Cas9. This could still be a viable method of studying *Foxn3*, but due to the less than complete overlap between Cre and Cas9 using our Cre-Cas9 VLPs (see Chapter 4.2.3), we decided not to attempt the aforementioned experiment with the currently available VLPs.

Another way in which VLPs could aid us in the study of *Foxn3* would be through delivery of the prime editor. The 5' end of the *Foxn3* coding region could be targeted by a guide RNA designed to install an epitope tag at the start codon of the Foxn3 coding region, since there are no poly-T tracks which could potentially terminate RNA polymerase type III polymerization of a prime editing guide. This means that a prime editing guide RNA could potentially be used to install an epitope tag at the N-terminus of Foxn3. Currently, there are no available antibodies suitable for immunodetection of mouse FoxN3. If prime editing could be used to install an epitope tag in the endogenous protein, it could allow for the immunodetection of Foxn3 expression patterns or CUT&RUN assays (Skene, Henikoff, and Henikoff 2018) in order to isolate Foxn3 genomic binding sites in retinal progenitor cells.

As an initial test of the feasibility and efficiency of protein delivery by VLPs into the cells of embryonic mouse retinas, we generated Cre VLPs which were concentrated 75-fold (see Methods). We harvested retinas from E17 Ai9:RiboTag transgenic mouse embryos. At E17, new retinal cells are still forming, differentiating, and migrating in the murine retina and these processes continue *in vitro* in explant cultures (H. Zhang et al. 2022). The newly collected retinas were incubated with concentrated VLP supernatant for 10 hours before being transferred to a support membrane for *in vitro* explant culture. After three days, many cells in the whole mount were expressing tdTomato in a widespread pattern (Fig. 28A). After whole mount imaging,

retinas were fixed and sectioned. The activation of the Ai9 tdTomato from the Ai9 reporter was detected by fluorescence in numerous retinal cells (Fig. 28B). These experiments provide proof of concept that the MMLV VLP platform can be used to deliver a genome editing enzyme into the developing murine retina.

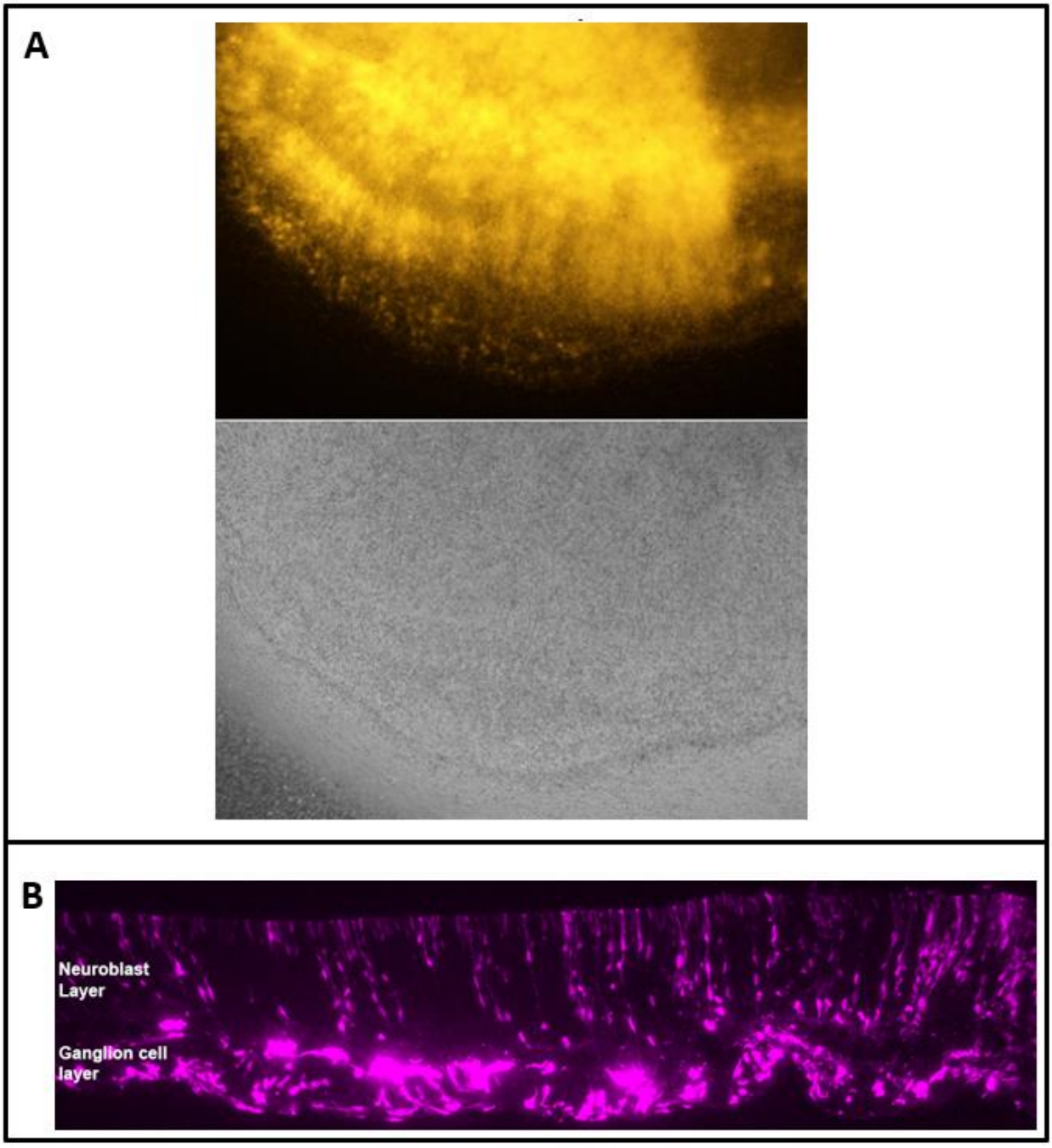

Figure 28 Cre VLPs can activate the tdTomato reporter in mouse retinal cells.
A) Whole mount of an embryonic retina which has been infected with Cre VLPs. The bottom phase image shows the edge of the retina as it lays almost flat on a supportive membrane. The top image taken with a fluorescent microscope shows widespread infection of the retina as evidenced by the cells expressing tdTomato. B) retina isolated from an Ai9:RiboTag mouse infected with concentrated Cre VLPs. After three days in explant culture, the retina was fixed, sectioned, and imaged for tdTomato expression (magenta pseudocolor). Numerous retinal cells express tdTomato indicating infection by the Cre VLPs. Some tdTomato positive cells migrated to the inner retina, consistent with neuronal differentiation. The neuroblast and ganglion cell layers are indicated.

4.3 Discussion:

VLPs are an efficient vehicle for the delivery of Cas9 loaded with a sgRNA. However, there are countless examples of situations where questions in biology cannot be answered unless Cas9 RNPs loaded with multiple sgRNAs can be delivered simultaneously. We have shown that this can be done with at least two different guide RNAs, but the overlap in cells edited at both targeted genomic loci was poor. However, that does not mean that there is no way to overcome this problem. Similarly, although the overlap between cells in a population edited by both Cre and Cas9 was low, there may be further manipulations of VLPs design and generation that improve editing enzyme co-delivery. The potential uses of dual purpose VLPs, either in terms of delivering multiplexed Cas9 or multiple enzymes, is great enough that our investment in them was useful for identifying the nature of the challenges surrounding this problem and narrowing down the list of improvements that could be made to VLPs in order to improve their capabilities.

4.3.1 Challenges and Potential Solutions to Multiplexed Cas9 VLPs

We have empirically identified two challenges in the delivery Cas9 VLPs with multiple guide RNAs. The first is that the amount of sgRNA coexpressed in producer cells can be subsaturating for loading onto Cas9. VLP delivery of Cas9 is only useful if it is loaded with sgRNA. Cas9 may not load with sgRNA if there are fewer sgRNA molecules than Cas9 protein molecules expressed. One method of increasing the amount of sgRNA coexpressed with Cas9 protein is to place the U6 cassette driving the sgRNA and the sCMV promoter driving GagCas9 mRNA expression on the same vector, specifically in tandem as there are indications that divergent placement may lead to transcriptional interference (Nie et al. 2010). The larger combined vector does slightly reduce the molar amount of GagCas9-expressing vector that is

transfected into cells when the mass of the vector is kept constant, but titers of VLPs made with these combined constructs are not substantially different than those made with vectors where GagCas9 and the sgRNA are expressed only from separate vectors (compare Fig. 22B to 22C). One method to increase the amount of sgRNA loaded onto Cas9 would be to put multiple U6 cassettes in tandem onto the same vector, which we have found can deliver two sgRNAs simultaneously (data not shown). Another method may be to use a stronger promoter such as U6 combined with the CMV enhancer which has been shown to increase production of short hairpin RNAs (Xia 2003). It also may be possible to improve the colocalization of sgRNA and Cas9 in producer cells. We unsuccessfully attempted to do this with the ET-208 synthetic aptamer. We do not know why ET-208 did not appear to increase Cas9 VLP titers. It may have been because the sgRNA was not being exported from the nucleus as expected, or because the sgRNA is only limiting due to the absolute number of sgRNA molecules in the cell and not subcellular localization.

The second challenge that we identified is that the loading efficiencies of sgRNAs may not be equal when they are coexpressed in producer cells. The results in Figure 21A suggest that the Wtap sgRNA loads into Cas9 VLPs more efficiently than the Rpl22 sgRNA. If this is due to incorrect folding of the Rpl22 sgRNA as we propose, the Rpl22 sgRNA may need to be redesigned in such a way that reduces misfolding, or a different guide RNA may be needed entirely.

4.3.2 Challenges and Potential Solutions to Delivering Multiple Functional Enzymes

Our experiments revealed that Cre and Cas9 have very different levels of efficacy when delivered by VLP. This issue may be inherent and unavoidable for any specific set of co-delivered proteins. While we were able to improve the overlap between Cas9 and Cre editing by

creating a chimera, that still was not a perfect solution. We made our VLPs in a way that ensured that only a small percentage, estimated to be 5%, of cells in the population were edited by Cre, but most of the cells in the total population were edited by Cas9. If a small number of Cre edited marked cells is acceptable, while ensuring that most or nearly all of the Cre marked cells are edited by Cas9, this method may be useful, especially if combined with a mechanism of isolating the Cre marked and edited cells such as flow cytometry. A higher overlap between cells edited both by Cre and by Cas9 would require further optimization of the manner in which the plasmids encoding the molecules that compose VLPs are transfected, or more likely, creating new constructs that further modulate the enzymes' efficiencies or possibly make the activity of one protein (i.e. Cre) dependent on the activity of the other protein (i.e. Cas9).

4.4 Methods

Cells were made and used as described in the Methods for Chapter 3.

Plasmids:

All plasmids were constructed as in Chapter 3 with the addition of the following plasmids. pCS7-Ai9-GagCas9 was made by inserting the U6 cassette driving the expression of the Ai9 sgRNA upstream of the sCMV promoter and oriented in the same direction. pCS7-Ai9-GagCO-Cas9 and pCS7-Ai9ET208-GagCO-Cas9 were cloned similarly using the pCS7-GagCO-Cas9 plasmid. pCS7-GagCO-Asn-Cas9 was made by digest of pCS7-GagCO-Cas9 and ligation of the annealed oligos CCTCAGACCAGCCTGAACACCCTGGACGATCAGGGAGCTAGCGGAGC and CATGGCTCCGCTAGCTCCCTGATCGTCCAGGGTGTTCAGGCTGGTCTGAGG.

Virus-Like Particles:

Figure 22A: 1 U6 : 1 sCMV VLPs were made by transfecting 5ng of pCMV-VSV-G, 140 ng of pCS7-GagPol, 85 ng of pCS7-Ai9-GagCas9, and 270 ng of the U6 vector encoding the Ai9

gRNA. Ai9-GagCas9 – Ai9 VLPs were made exactly as stated above with the substitution of the U6 vector for 270 ng of non-coding plasmid DNA. Cre VLPs were made using 25 ng of VSV-g and 475 ng of Gag-Cre or using 25 ng of VSV-g, 335 ng of Gag-Cre and 140 ng of GagPol. All other VLPs were made and used as described in Chapter 3.

Indirect immunofluorescent detection for cells in monolayer:

For detection of HA, cells were fixed with 4% PFA in PBS for ten minutes and then washed three times with PBS. Cells were incubated with a blocking solution of 0.2% Triton and 30% goat serum in PBS for ten minutes. The anti-HA antibody (Cell Signaling Technology C29F4) was diluted 1:1600 in 0.05% Triton and 10% goat serum in PBS. Incubation was for 1 hour at room temperature, followed by three washes with PBS. Alexa Fluor® 594 AffiniPure Donkey Anti-Rabbit IgG (H+L) (Jackson Immunoresearch 711-585-152) was diluted 1:1000 in 0.05% Triton and 10% goat serum.

For combined immunodetection of Wtap/HA, cells were fixed with 2% PFA for ten minutes. Cells were incubated with a blocking solution of 1% Triton and 90% goat serum in PBS for ten minutes. The anti-Wtap antibody (Santa Cruz D-7) was diluted 1:50 in 0.1% Triton and 10% goat serum. The anti-HA antibody (Cell Signaling Technology C29F4) was diluted 1:1200 in 0.1% Triton and 10% goat serum. Primary antibodies were incubated for 1 hour. After the anti-Wtap antibody was removed, cells were washed three times, then the anti-HA antibody was added for 1 hour. Alexa Fluor® 594 AffiniPure Donkey Anti-Rabbit IgG (H+L) (Jackson Immunoresearch 711-585-152) was diluted 1:1000 in 0.1% Triton and 10% goat serum. Alexa Fluor® 488 AffiniPure Goat Anti-Mouse IgG (H+L) was diluted 1:400 in 0.1% Triton and 10% goat serum. Secondary antibodies were incubated for 30 minutes.

Flow Cytometry:

All cells were trypsinized, quenched with DMEM with 10% FBS and transferred to 1.7 mL tubes. All centrifugation steps were 500 × g for five minutes. All steps were performed at room temperature unless otherwise indicated. All cells were passed through a 40 µm FlowMi filter (Bel-Art H13680-0040) prior to flow analysis.

Cells Assayed for tdTomato Fluorescence:

Cells were harvested 72 hours after infection. Cells were incubated with 4% PFA in PBS for ten minutes, then washed three times with PBS.

WTAP and HA co-staining:

Cells were harvested 72 hours after infection. The cell pellet was resuspended in 30% goat serum, 0.1% Pluronic F-68 Non-ionic Surfactant (F-68) (Gibco 24040-032), 0.3% Tween-20, 1ug/mL DNaseI for 10 minutes. Cells were 500 × g for five minutes and the supernatant was aspirated. 200 µL of PBS with 0.1% F-68 was added to the cells and they were resuspended by pipetting. Cells were 500 × g for five minutes. The anti-WTAP antibody was diluted 1:50 in 0.1% Tween-20, 0.1% F-68, 10% goat serum. 200ul of diluted antibody was added to the cells for 1hr. The anti-HA antibody was diluted 1:1200 in 0.1% Tween-20, 0.1% F-68, 10% goat serum. 200ul of diluted antibody was added to cells and they were incubated for 1 hour. The negative control received 200ul of 0.1% Tween-20, 0.1% F-68, 10% goat serum. VLPs that received both antibodies were incubated with anti-WTAP antibody, washed once with 0.1% F-68 in PBS, and then incubated for an hour with the anti-HA antibody. After incubations were complete, cells were washed twice with 0.1% F-68 in PBS. The secondary antibody anti-mouse Alexa Fluor 488 was diluted 1:400 and anti-rabbit Alexa Fluor 594 was diluted 1:1000 in PBS with 0.1% Tween-20, 0.1% F-68, 10% goat serum. 200 µL of the secondary antibody (or both

secondaries) was added to cells and incubated for 30 minutes. After incubations were complete, cells were washed twice with 0.1% F-68 in PBS and resuspended in 900ul of PBS.

Mouse Retina Infections

Cre VLPs were made by the transfection of HEK293T cells in 24 well plates with 25 ng of pCMV-VSV-g, 5 ng of pCS7-Gag-Prot, and 480 ng of pCS7-Gag-Cre using VirusGen (Mirus). At 24 hours post-transfection, the media was aspirated and replaced with 1 mL of prewarmed media composed of DMEM with 20% horse serum. At 48 hours post-transfection, 4 replicate wells were combined and filtered through a 0.45 micron surfactant-free cellulose acetate filter (Thermo Scientific 723-2545). Cre VLPs were concentrated 75-fold by centrifugation at 20,000 x g for 2.5 hours at 4°C. They were resuspended in 40 µL of media composed of DMEM with 20% horse serum spiked with polybrene at 8 µg/mL. A small aliquot of the concentrated Cre VLPs were applied to Ai9 MDF cells where they were found to have a titer of 2x10^7 IU/ml. Retinas were incubated with the remainder of the VLP concentrate in 96 well plates for ten hours at 37°C prior to transferring explants onto support membranes. After three days, retinas were fixed and sectioned.

Chapter 5 Conclusions and Future Directions

VLPs based on MLV have been previously described as vehicles for protein transduction into cells (Mangeot et al. 2019; Voelkel et al. 2010; D.-T. Wu and Roth 2014; Bobis-Wozowicz et al. 2015), as well as in a recent paper (Banskota et al. 2022). Our work with new VLPs based on the MMLV has yielded a new vehicle for protein transduction and insights into the features that are important for efficient protein delivery, including the delivery of genome editing proteins. We have shown that codon optimization and addition of a Kozak sequence for MMLV Gag improve Cas9 VLP efficacy by increasing GagCas9 protein production, and likely also by increasing protein localization to the cell membrane and multimerization into VLPs through the increasing production of the full-length GagCas9 protein. We have shown that viral protease processing likely increases the efficacy of Cas9 VLPs and that the viral protease can be provided from a GagPol construct that does not contain functional reverse transcriptase or integrase. We have shown that the Cas9 prime editor can be fused to Gag and retain its function by combining it with a pegRNA that introduces eGFP into the human *LMNA* locus, which to our knowledge is the largest prime editing insertion to date. We have shown the first example of prime editor delivery via VLP with a pegRNA that inserts the HA epitope tag into human *LMNA* or mouse *Actb* genes. Although the efficiencies of these first-generation PE VLPs are very modest, the fact that they function at all and can be readily assayed through immunohistochemistry and next-generation sequencing means that they can be modified by several different mechanisms, as discussed below, and assayed to determine which engineered modifications may improve their function.

5.1 Implications of RNA-targeting by NmeCas9

CRISPR-Cas (CRISPR associated) multicomponent systems are used by microbes as an adaptive immune system to protect them against mobile genetic elements such as phages. The Type II CRISPR-Cas system contains the endonuclease Cas9 which is guided by RNA. Cas9 is canonically a DNA-targeting enzyme. However, we in chapter 2 (Rousseau et al. 2018) and others (Dugar et al. 2018) have shown that NmeCas9 and CjeCas9 have natural RNA-targeting abilities in vitro. Cas9 could theoretically be used in mammalian cells as both a DNA and RNA-targeting enzyme for the rapid silencing of genes by disrupting the coding region of a gene and also silencing mRNA transcripts by directly cleaving them. Transcript knockdown by dCas9 has been shown in vivo (Batra et al. 2017), but a nuclease deficient Cas9 would not be able to also induce a mutation at the genomic locus from where the transcript is derived. This may have special application to post-mitotic cells which do not have the benefit of transcript dilution and rapid gene turnover more commonly seen in rapidly dividing cells. However, simultaneous DNA and RNA targeting by Cas9 in mammalian cells has not yet been shown in the literature. This is a complex challenge likely explained by a few different factors such as the ability to target RNA in an RNA-guided manner by Cas9 or other CRISPR-Cas enzymes, the challenge of limiting Cas9 off-target activity in vivo, and the difficulty of finding a model system to show the utility of this approach.

Knock-down of transcripts by RNA interference (RNAi) is a naturally occurring method of gene regulation in eukaryotes that has been coopted as a useful tool in the laboratory for the effective knock-down of a gene through the introduction a small interfering RNA (siRNA) (Paddison and Vogt 2008). In recent years, RNAi has shown therapeutic potential for cancers and other diseases (Kara, Calin, and Ozpolat 2022). If RNAi were used in conjunction with a

Cas9 enzyme targeted against the same gene, with Cas9 editing to interrupt the coding region and knockdown of existing transcripts by RNAi could be achieved, obviating the need for Cas9 enzymes loaded with a guide RNA capable of targeting the gene and transcripts to perform both functions simultaneously. However, the delivery efficiency for RNPs and siRNAs may differ, depending on the mechanism of delivery. In the CRISPR toolbox, CRISPR-Cas13 is an RNA-guide RNA-targeting enzyme that has been used in mammalian cells for the knock-down of RNA transcripts (Abudayyeh et al. 2017). Cas9 and Cas13 could both potentially be delivered as RNPs, or possibly as the cargo of VLPs, to achieve simultaneous genome editing and transcript degradation. Both RNA-targeting by RNAi and CRISPR-Cas are established methods used for the interrogation of gene product functions in research, and with Cas9 able to target a genomic locus, using Cas9 with either RNA silencing technique might be a better strategy than investing in the search for a naturally occurring Cas9 enzyme that can target both DNA and RNA with high fidelity and high efficiency.

If Cas9 were to be used as an all-in-one enzyme to mediate gene disruption and transcript silencing, it would need to do so with low off-target activity, and with the ability to localize to both the cytoplasm and the nucleus to reach its RNA and DNA targets. Cas9 is known to cause off-target cleavage events (see Chapter 1.3) which can be mitigated with the use of a higher fidelity Cas9, the use of engineered guide RNAs, or by delivering Cas9 as an RNP. Although RNA-targeting in cells by dCas9 has been described, the RNA targeted was composed of microsatellite repeats, which are plentiful and suggest that SpyCas9's affinity for RNA is not very high (Batra et al. 2017). RNA-targeting in mammalian cells has not yet been reported for either NmeCas9 or CjeCas9. Again, this suggests that their affinity for RNA may not be high enough to have a biological effect. If a naturally occurring Cas9 that strongly targets RNA

cannot be found, then it might be possible to engineer one. However, engineering of Cas9 enzymes in order to make Cas9 more permissive in targeting (i.e. permit Cas9 to recognize a PAM sequence such as 5'-NG-3' versus the more restrictive 5'-NGG-3') can result in Cas9 enzymes with greater off-target activity (Walton et al. 2020). This might be the case also if a Cas9 enzyme were engineered to target RNA. A careful balance between effective RNA-targeting and high-fidelity genome editing would need to be achieved. Furthermore, Cas9 is normally localized to the nucleus with NLS sequences. It may be possible to augment the localization of Cas9 to be predominantly in the cytoplasm (where there may be more transcripts depending on the targeted RNA) but still maintain some presence in the nucleus by fusing Cas9 to peptides from viruses that have the ability to change their localization (Tessier et al. 2019). While perhaps not insurmountable, these challenges to make Cas9 both a DNA and RNA-targeting enzyme coupled with alternate existing paths that may achieve simultaneous genome editing and RNA knock-down suggest that the goal of finding or creating a Cas9 with both excellent genome editing fidelity and high efficacy in RNA interference is perhaps not the best route to pursue.

5.2 Implications of Efficient Delivery of Cas9 by Virus-Like Particles

Many cell types are difficult to transfect, particularly primary cells and stem cells. These cell types are of great importance for basic research, for studies of diseases, and for creating or testing potential treatments for illnesses. In addition to facilitating Cas9 delivery to primary cells, protein transduction of Cas9 or the prime editor with VLPs has several advantages over delivery by other methods (see Chapter 1.5). In particular, the dose of Cas9 can be more easily controlled, Cas9 ribonucleoproteins can act without the delay necessary for synthesis after delivery by nucleic acids or viral vectors, and the Cas9 protein is expressed for a limited period of time.

Banskota et. al. demonstrated that VLPs loaded with the base editor were able to edit a sufficient fraction of mouse hepatocytes in vivo that they improved the mouse serum profile to a therapeutically relevant extent (Banskota et al. 2022). They also used base editor VLPs to improve visual function in a mouse model of genetic blindness. Although we did not achieve our goal of efficiently editing somatic cells in a murine retina, the pilot experiment in Chapter 4 shows that MMLV VLPs can deliver genomic editing proteins to embryonic retinal cells in explant culture, and support the possibility of using this method in the future for Cas9 delivery. There are strong indicators for engineering strategies that could successfully increase the efficacy of Cas9 and prime editing VLPs as discussed below and in previous chapters. Such improvements could allow genome editing in a sufficient fraction of somatic cells in vivo to obviate the need for a genetically modified mouse model to study some genes relevant to development. This may be especially valuable for the analysis of genes that are embryonically lethal and can only be disrupted in somatic cells at later ages with a conditional knock-out approach.

5.3 Considerations in the Design of Gag-Fusion Proteins

Unlike putting a new CD into a CD-player, proteins fused to Gag are not so easily interchangeable. We found that Cre VLPs had good titers when produced with our first GagCre construct, and that the presence of the viral protease was not required for Cre function. On the other hand, Gag fusions to Cas9, Cre-Cas9, or the prime editor protein yielded suboptimal VLP titers even after codon-optimization of the Gag protein, and inclusion of the viral protease was important for maximal Cas9 VLP effectiveness. The lower Cas9 VLP titers are likely because there are multiple factors affecting the efficiency of Cas9-containing VLPs, including protein production and localization to the cell membrane, the efficacy of guide RNA loading onto Cas9

prior to VLP budding, the efficiency with which Cas9 is cleaved from Gag, and the efficiency with which Cas9 enters the nucleus of the target cell. Our data, in conjunction with the data from Banskota et. al. (2022), lead us to suggest that four changes could be made to improve future Cas9 and prime editing VLPs.

The first change to consider is the addition of a nuclear export sequence (NES) at the end of Gag but before the viral protease cleavage site. This should direct the Gag-fusion to the cell membrane and improve the efficiency that Gag-fusion proteins are loaded into VLPs. Our design for GagCas9 has one SV40 large T antigen NLS sequence preceding the first methionine in Cas9 and one nucleoplasmin NLS sequence at the end of the protein. These NLS's are very effective for encouraging protein translocation to the nucleus (Lange et al. 2007). If GagCas9 is in the nucleus, it cannot be at the cell membrane and will not load into VLPs. The addition of a NES could localize more GagCas9 to the cytoplasm where it would be more likely to load into particles.

Apart from the NES, it may be possible to borrow sequences from proteins that have regulated nuclear export and import. The Rabies virus P-protein has an N-terminal NES that strongly induces nuclear export, but it also has a unique set of overlapping NES and NLS sequences (Oksayan et al. 2012). The NES masks the presence of the NLS until the NES is cleaved in half which reveals the presence of a very strong NLS. It may be possible to substitute these overlapping NES/NLS sequences so that they are separated by a viral protease cleavage site in a Gag-fusion protein. Before cleavage, the NLS would be masked by the presence of the NES. After cleavage, the NLS would be able to translocate the protein to the nucleus. This might make it possible to remove the bipartite NLS sequences from Cas9 entirely, thus further improving the ability of the protein to localize in a spatially and temporally appropriate fashion.

Aside from the Rabies virus, many other viruses have complex mechanisms for shuttling between the nucleus and the cytoplasm (Tessier et al. 2019). One or more of them may provide a solution to the problem of Gag-fusion localization.

A second possible change is to vary the viral protease cleavage site between Gag and the fusion protein. Two constructs with new sites could be made. The site used in eVLPs (TSTLL/MENSS) may prove useful (Banskota et al. 2022). Another site to be tested would be the MA/P12 sequence (ASLLTRSSLY/PALPT) that we found to be better than the NC/Prot sequence (see Chapter 4.2.5). VLPs using these alternate protease sites could be assayed using the Ai9 MDF cell line and Cas9 VLPs loaded with the Ai9 sgRNA. PE VLPs could be tested either with the HA-lamin-pegRNA or the HA-ACTB-pegRNA in either human or mouse cells respectively. A western blot could be used to assess potential differences in cleavage efficiency for the different cleavage sites. The best viral cleavage site may not be a naturally occurring one, but one that can be found through screening a small library. Testing Cas9 VLPs loaded with the Ai9 sgRNA and used to infect Ai9 MDF cells would be a quick and simple method of screening through a reasonable number of candidate cleavage sites to determine the best one for delivering Cas9 VLPs.

A third potential change would be to alter the stoichiometry of the plasmids co-transfected to make VLPs. While we made an attempt to do this (see Chapter 4.2.6), a more systematic approach using optimized constructs would likely result in better outcomes. Making Cas9 VLPs loaded with the Ai9 sgRNA and using them to infect MDF cells might be the best way to test different ratios of the plasmids needed to create Cas9 VLPs. The ratio of the vectors encoding GagCas9 and another vector that is the source of Gag and the viral protease (i.e. GagPol, GagPol-RT$^-$IN$^-$, or Gag-Stop-Prot) could be altered until an optimal titer is achieved.

The results of those experiments may translate into enhanced efficiency when applied to PE VLPs as well. If so, this would be fortunate because PE VLPs take longer to assay. Cells (mouse or human) infected with PE VLPs are normally assayed 96 hours after infection whereas Ai9 MDF cells infected with Cas9 VLPs targeting the Ai9 locus are normally assayed at 72 hours post-infection. However, it is possible that different proteins will always have some different preferences in their stoichiometry due to steric hindrance. As we found, the stoichiometry that worked best for Cre VLPs did not work as well for Cas9 VLPs.

A fourth potential change would be to alter the guide RNAs for Cas9. It remains possible that Cas9 in VLPs is not completely loaded with guide RNA (see Chapters 3 and 4). With our current transfection conditions, the amount of the U6 sgRNA vector used in the cotransfection is half of the total amount of vector DNA in the transfection. The vector that expresses the guide RNA is 89% backbone, with only 11% comprising the U6 cassette that expresses the guide RNA sequence. One method of removing the vector backbone and thereby increasing the amount of the necessary U6 cassette and guide RNA would be to amplify that segment by PCR to create a small dsDNA template for transfection. The inclusion of phosphorothioate bonds at the ends of the PCR primers could help make the linear dsDNA resistant to degradation when it is transfected into cells. This may allow the use of less DNA dedicated to expressing the sgRNA in the transfection, and allow for inclusion of additional vector DNA to express the other components necessary to Cas9 VLP formation. Ensuring that guide RNA loading onto the GagCas9 fusion is as complete as possible ensures that the Cas9 in VLPs is being delivered in a form that is functional. Furthermore, for longer guide RNAs, such as prime editing guide RNAs, it is possible that degradation during the long journey from one cell to another compromises the integrity of the guide RNA. The guide RNA could be stabilized by the introduction of a hairpin

at the 3’ end of the guide RNA that may prevent it from being degraded by a 3’ exonuclease. There are sequences known to do this. We added ET-208, predicted to form a hairpin-like structure, to the end of the eGFP-lamin-pegRNA (data not shown) and found that it worked almost as well as the original pegRNA. However, because we did not use that pegRNA in the making of VLPs, we do not know what effect it would have had on PE VLP efficacy. Chemically synthesized sgRNAs with stabilizing chemical modifications also could be used to make VLPs. Cas9 has been shown to efficiently edit cells when delivered as a RNP loaded with a guide RNA that has chemical modifications (Mir, Alterman, et al. 2018). VLPs produced with such stabilized sgRNAs may allow the sgRNA to persist longer, leading to more functional Cas9/sgRNA delivered to target cells, which should increase genome editing.

In vivo editing by Cas9 VLPs (Mangeot et al. 2019; Banskota et al. 2022) has shown the promise of VLPs as a method of delivering Cas9 and Cas9-derivatives as RNPs. This allows for tighter control over the dose of Cas9 delivered as well as limits the duration of time that Cas9 is present in the cell and avoids the needs to borrow host-cell transcription and translational machinery, unlike delivery by AAV or lentivirus. However, we found that there was still room for improvement for our VLPs in our pilot experiment with the infection of embryonic murine retinas in explant (see Chapter 4.2.8). The whole mounts from our pilot experiment (Fig. 28A) suggested that Cre VLPs preferentially infected cells along the edges of the retina, with fewer cells infected across the broad outer surface of the retina. This suggests that part of the reason we saw suboptimal infection was because the VLPs were not pseudotyped appropriately to take advantage of the receptors available on the outer surface of the retina (comprised mostly of progenitor cell projections in the embryonic retina). Eco Env may be an alternative to VSV-g in retinas, as retroviral vectors with Eco Env are known to infect murine retinas, or possibly a

combination of both VSV-g and Eco Env. However, we did not attempt these infections because of the apparently limited titers for VLPs prepared with Eco Env, based on results in the Ai9 MDF cell line (see Chapter 3.2.1). Alternately, there are other viral envelope proteins that could be used to pseudotype VLPs. The envelope protein from the Hantavirus might be the best candidate. Hantaviruses are negative sense ssRNA viruses in the Bunyavirales order. There are two types of Hantaviruses, New World (e.g. Andes Virus) and Old World (e.g. Haantanvirus, Dobrava-Belgrade virus) (Slough, Chandran, and Jangra 2019). Apart from their difference in origin, the two types of viruses differ in whether they bud at the golgi (Old World) or at the plasma membrane (New World). For MLV derived VLPs, it would be reasonable to use an envelope from a New World Hantavirus. The Hantavirus envelope protein (Gn/Gc) has been used to pseudotype replication competent or incompetent vesicular stomatitis virus (VSV) particles (Ray et al. 2010). In one example, Andes Virus enveloped VLPs infected a variety of cell types (Vero, BHK-1, HEK293T, Hela, RK-13, etc.) about as well as the same VLPs pseudotyped with VSV-G. Protocadherin1 (PCDH1) is the receptor that mediates entry of New World Hantavirus Gn/Gc glycoproteins (Jangra et al. 2018). PCDH1 is expressed in the murine retina from E15 (Redies et al. 2008). Taken together, it is likely that the Hantavirus Gn/Gc envelope protein would be a good candidate to use to pseudotype Cas9 VLPs. It could be initially tested by infection of the Ai9 MDF cell line. There are likely still unexplored ways to improve the efficacy of infection in the retina by VLPs, such as the VLP modifications suggested above (e.g. the addition of a NES to Gag, the variation of the protease cleavage site, altering the stoichiometry of the proteins composing VLPs, increasing the efficiency of sgRNA loading, or pseudotyping with different envelope proteins). These coupled with the fact that we have a quick and facile method for testing the effects of modifications on VLP efficiency (e.g. infecting Ai9

MDF cells with Cas9 VLPs loaded with the Ai9 sgRNA) makes this a very exciting and promising area of research that will see great progress in the near future.

Bibliography

Abudayyeh, Omar O., Jonathan S. Gootenberg, Patrick Essletzbichler, Shuo Han, Julia Joung, Joseph J. Belanto, Vanessa Verdine, et al. 2017. "RNA Targeting with CRISPR–Cas13." *Nature* 550 (7675): 280–84. https://doi.org/10.1038/nature24049.

Abudayyeh, Omar O., Jonathan S. Gootenberg, Silvana Konermann, Julia Joung, Ian M. Slaymaker, David B. T. Cox, Sergey Shmakov, et al. 2016. "C2c2 Is a Single-Component Programmable RNA-Guided RNA-Targeting CRISPR Effector." *Science* 353 (6299): aaf5573. https://doi.org/10.1126/science.aaf5573.

Andrawiss, Mariam, Yasuhiro Takeuchi, Lindsay Hewlett, and Mary Collins. 2003. "Murine Leukemia Virus Particle Assembly Quantitated by Fluorescence Microscopy: Role of Gag-Gag Interactions and Membrane Association." *Journal of Virology* 77 (21): 11651–60. https://doi.org/10.1128/JVI.77.21.11651-11660.2003.

Annunziato, Stefano, Catrin Lutz, Linda Henneman, Jinhyuk Bhin, Kim Wong, Bjørn Siteur, Bas Gerwen, et al. 2020. "*In Situ* CRISPR-Cas9 Base Editing for the Development of Genetically Engineered Mouse Models of Breast Cancer." *The EMBO Journal* 39 (5). https://doi.org/10.15252/embj.2019102169.

Anzalone, Andrew V., Peyton B. Randolph, Jessie R. Davis, Alexander A. Sousa, Luke W. Koblan, Jonathan M. Levy, Peter J. Chen, et al. 2019. "Search-and-Replace Genome Editing without Double-Strand Breaks or Donor DNA." *Nature* 576 (7785): 149–57. https://doi.org/10.1038/s41586-019-1711-4.

Auerbach, Marcy R., Kristy R. Brown, and Ila R. Singh. 2007. "Mutational Analysis of the N-Terminal Domain of Moloney Murine Leukemia Virus Capsid Protein." *Journal of Virology* 81 (22): 12337–47. https://doi.org/10.1128/JVI.01286-07.

Banskota, Samagya, Aditya Raguram, Susie Suh, Samuel W. Du, Jessie R. Davis, Elliot H. Choi, Xiao Wang, et al. 2022. "Engineered Virus-like Particles for Efficient in Vivo Delivery of Therapeutic Proteins." *Cell* 185 (2): 250-265.e16. https://doi.org/10.1016/j.cell.2021.12.021.

Barrangou, Rodolphe, Christophe Fremaux, Hélène Deveau, Melissa Richards, Patrick Boyaval, Sylvain Moineau, Dennis A. Romero, and Philippe Horvath. 2007. "CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes." *Science* 315 (5819): 1709–12. https://doi.org/10.1126/science.1138140.

Bartels, Hanni, and Jeremy Luban. 2014. "Gammaretroviral Pol Sequences Act in Cis to Direct Polysome Loading and NXF1/NXT-Dependent Protein Production by Gag-Encoded RNA," 22.

Batra, Ranjan, David A. Nelles, Elaine Pirie, Steven M. Blue, Ryan J. Marina, Harrison Wang, Isaac A. Chaim, et al. 2017. "Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9." *Cell* 170 (5): 899-912.e10. https://doi.org/10.1016/j.cell.2017.07.010.

Battini, J L, J M Heard, and O Danos. 1992. "Receptor Choice Determinants in the Envelope Glycoproteins of Amphotropic, Xenotropic, and Polytropic Murine Leukemia Viruses." *Journal of Virology* 66 (3): 1468–75. https://doi.org/10.1128/jvi.66.3.1468-1475.1992.

Bhaya, Devaki, Michelle Davison, and Rodolphe Barrangou. 2011. "CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation."

Annual Review of Genetics 45 (1): 273–97. https://doi.org/10.1146/annurev-genet-110410-132430.

Bobis-Wozowicz, Sylwia, Melanie Galla, Jamal Alzubi, Johannes Kuehle, Christopher Baum, Axel Schambach, and Toni Cathomen. 2015. "Non-Integrating Gamma-Retroviral Vectors as a Versatile Tool for Transient Zinc-Finger Nuclease Delivery." *Scientific Reports* 4 (1): 4656. https://doi.org/10.1038/srep04656.

Bogdanove, Adam J., and Daniel F. Voytas. 2011. "TAL Effectors: Customizable Proteins for DNA Targeting." *Science* 333 (6051): 1843–46. https://doi.org/10.1126/science.1204094.

Bouard, D, N Alazard-Dany, and F-L Cosset. 2009. "Viral Vectors: From Virology to Transgene Expression: Viral Vectors: From Virology to Transgene Expression." *British Journal of Pharmacology* 157 (2): 153–65. https://doi.org/10.1038/bjp.2008.349.

Briner, Alexandra E., Paul D. Donohoue, Ahmed A. Gomaa, Kurt Selle, Euan M. Slorach, Christopher H. Nye, Rachel E. Haurwitz, Chase L. Beisel, Andrew P. May, and Rodolphe Barrangou. 2014. "Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality." *Molecular Cell* 56 (2): 333–39. https://doi.org/10.1016/j.molcel.2014.09.019.

Brinkman, Eva K., Tao Chen, Mario Amendola, and Bas van Steensel. 2014. "Easy Quantitative Assessment of Genome Editing by Sequence Trace Decomposition." *Nucleic Acids Research* 42 (22): e168–e168. https://doi.org/10.1093/nar/gku936.

Brouns, Stan J. J., Matthijs M. Jore, Magnus Lundgren, Edze R. Westra, Rik J. H. Slijkhuis, Ambrosius P. L. Snijders, Mark J. Dickman, Kira S. Makarova, Eugene V. Koonin, and John van der Oost. 2008. "Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes." *Science* 321 (5891): 960–64. https://doi.org/10.1126/science.1159689.

Bunnell, Tina M., Brandon J. Burbach, Yoji Shimizu, and James M. Ervasti. 2011. "β-Actin Specifically Controls Cell Growth, Migration, and the G-Actin Pool." Edited by Paul Forscher. *Molecular Biology of the Cell* 22 (21): 4047–58. https://doi.org/10.1091/mbc.e11-06-0582.

Cai, Yujia, Rasmus O Bak, and Jacob Giehm Mikkelsen. 2014. "Targeted Genome Editing by Lentiviral Protein Transduction of Zinc-Finger and TAL-Effector Nucleases." *ELife* 3 (April): e01911. https://doi.org/10.7554/eLife.01911.

Cai, Yujia, Anders Laustsen, Yan Zhou, Chenglong Sun, Mads Valdemar Anderson, Shengting Li, Niels Uldbjerg, Yonglun Luo, Martin R Jakobsen, and Jacob Giehm Mikkelsen. 2016. "Targeted, Homology-Driven Gene Insertion in Stem Cells by ZFN-Loaded 'All-in-One' Lentiviral Vectors." *ELife* 5 (June): e12213. https://doi.org/10.7554/eLife.12213.

Cai, Yujia, and Jacob Giehm Mikkelsen. 2014. "Driving DNA Transposition by Lentiviral Protein Transduction." *Mobile Genetic Elements* 4 (3): e29591. https://doi.org/10.4161/mge.29591.

Canver, Matthew C., Daniel E. Bauer, Abhishek Dass, Yvette Y. Yien, Jacky Chung, Takeshi Masuda, Takahiro Maeda, Barry H. Paw, and Stuart H. Orkin. 2014. "Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells*." *Journal of Biological Chemistry* 289 (31): 21312–24. https://doi.org/10.1074/jbc.M114.564625.

Chang, Howard H. Y., Nicholas R. Pannunzio, Noritaka Adachi, and Michael R. Lieber. 2017. "Non-Homologous DNA End Joining and Alternative Pathways to Double-Strand Break Repair." *Nature Reviews Molecular Cell Biology* 18 (8): 495–506. https://doi.org/10.1038/nrm.2017.48.

Chen, Janice S, Yavuz S Dagdas, Benjamin P Kleinstiver, Moira M Welch, Alexander A Sousa, Lucas B Harrington, Samuel H Sternberg, J Keith Joung, Ahmet Yildiz, and Jennifer A Doudna. 2018. “Enhanced Proofreading Governs CRISPR-Cas9 Targeting Accuracy,” 25.

Chen, Qiubing, Ying Zhang, and Hao Yin. 2021. “Recent Advances in Chemical Modifications of Guide RNA, MRNA and Donor Template for CRISPR-Mediated Genome Editing.” *Advanced Drug Delivery Reviews* 168 (January): 246–58. https://doi.org/10.1016/j.addr.2020.10.014.

Cho, Seung Woo, Sojung Kim, Jong Min Kim, and Jin-Soo Kim. 2013. “Targeted Genome Engineering in Human Cells with the Cas9 RNA-Guided Endonuclease.” *Nature Biotechnology* 31 (3): 230–32. https://doi.org/10.1038/nbt.2507.

Chylinski, Krzysztof, Kira S. Makarova, Emmanuelle Charpentier, and Eugene V. Koonin. 2014. “Classification and Evolution of Type II CRISPR-Cas Systems.” *Nucleic Acids Research* 42 (10): 6091–6105. https://doi.org/10.1093/nar/gku241.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, et al. 2013. “Multiplex Genome Engineering Using CRISPR/Cas Systems.” *Science* 339 (6121): 819–23. https://doi.org/10.1126/science.1231143.

Cox, David B T, Jonathan S Gootenberg, Omar O Abudayyeh, Brian Franklin, Max J Kellner, Julia Joung, and Feng Zhang. 2017. “RNA Editing with CRISPR-Cas13,” 10.

Cronin, James, Xian-Yang Zhang, and Jakob Reiser. 2005. “Altering the Tropism of Lentiviral Vectors through Pseudotyping.” *Current Gene Therapy* 5 (4): 387–98. https://doi.org/10.2174/1566523054546224.

Czaplinski, Kevin. 2014. “Understanding MRNA Trafficking: Are We There Yet?” *Seminars in Cell & Developmental Biology* 32 (August): 63–70. https://doi.org/10.1016/j.semcdb.2014.04.025.

Danielsson, Frida, McKenzie Peterson, Helena Caldeira Araújo, Franziska Lautenschläger, and Annica Gad. 2018. “Vimentin Diversity in Health and Disease.” *Cells* 7 (10): 147. https://doi.org/10.3390/cells7100147.

Datta, S. A. K., X. Zuo, P. K. Clark, S. J. Campbell, Y.-X. Wang, and A. Rein. 2011. “Solution Properties of Murine Leukemia Virus Gag Protein: Differences from HIV-1 Gag.” *Journal of Virology* 85 (23): 12733–41. https://doi.org/10.1128/JVI.05889-11.

Deltcheva, Elitza, Krzysztof Chylinski, Cynthia M. Sharma, Karine Gonzales, Yanjie Chao, Zaid A. Pirzada, Maria R. Eckert, Jörg Vogel, and Emmanuelle Charpentier. 2011. “CRISPR RNA Maturation by Trans-Encoded Small RNA and Host Factor RNase III.” *Nature* 471 (7340): 602–7. https://doi.org/10.1038/nature09886.

Dhar, Arun K., Refugio Robles-Sikisaka, Vanvimon Saksmerprome, and Dilip K. Lakshman. 2014. “Biology, Genome Organization, and Evolution of Parvoviruses in Marine Shrimp.” In *Advances in Virus Research*, 89:85–139. Elsevier. https://doi.org/10.1016/B978-0-12-800172-1.00003-3.

Doench, John G, Nicolo Fusi, Meagan Sullender, Mudra Hegde, Emma W Vaimberg, Katherine F Donovan, Ian Smith, et al. 2016. “Optimized SgRNA Design to Maximize Activity and Minimize Off-Target Effects of CRISPR-Cas9.” *Nature Biotechnology* 34 (2): 184–91. https://doi.org/10.1038/nbt.3437.

Dugar, Gaurav, Ryan T. Leenay, Sara K. Eisenbart, Thorsten Bischler, Belinda U. Aul, Chase L. Beisel, and Cynthia M. Sharma. 2018. “CRISPR RNA-Dependent Binding and Cleavage

of Endogenous RNAs by the Campylobacter Jejuni Cas9." *Molecular Cell* 69 (5): 893-905.e7. https://doi.org/10.1016/j.molcel.2018.01.032.
Durkin, Marian, Xiaolan Qian, Nicholas Popescu, and Douglas Lowy. 2013. "Isolation of Mouse Embryo Fibroblasts." *BIO-PROTOCOL* 3 (18). https://doi.org/10.21769/BioProtoc.908.
East-Seletsky, Alexandra, Mitchell R. O'Connell, Spencer C. Knight, David Burstein, Jamie H. D. Cate, Robert Tjian, and Jennifer A. Doudna. 2016. "Two Distinct RNase Activities of CRISPR-C2c2 Enable Guide-RNA Processing and RNA Detection." *Nature* 538 (7624): 270–73. https://doi.org/10.1038/nature19802.
Esvelt, Kevin M, Prashant Mali, Jonathan L Braff, Mark Moosburner, Stephanie J Yaung, and George M Church. 2013. "Orthogonal Cas9 Proteins for RNA-Guided Gene Regulation and Editing." *Nature Methods* 10 (11): 1116–21. https://doi.org/10.1038/nmeth.2681.
Fehér, Anita, Péter Boross, Tamás Sperka, Gabriella Miklóssy, János Kádas, Péter Bagossi, Stephen Oroszlan, Irene T. Weber, and József Tözsér. 2006. "Characterization of the Murine Leukemia Virus Protease and Its Comparison with the Human Immunodeficiency Virus Type 1 Protease." *Journal of General Virology* 87 (5): 1321–30. https://doi.org/10.1099/vir.0.81382-0.
Feng, Y X, H Yuan, A Rein, and J G Levin. 1992. "Bipartite Signal for Read-through Suppression in Murine Leukemia Virus MRNA: An Eight-Nucleotide Purine-Rich Sequence Immediately Downstream of the Gag Termination Codon Followed by an RNA Pseudoknot." *Journal of Virology* 66 (8): 5127–32. https://doi.org/10.1128/jvi.66.8.5127-5132.1992.
Fukusumi, Yoshiyasu, Chie Naruse, and Masahide Asano. 2008. "Wtap Is Required for Differentiation of Endoderm and Mesoderm in the Mouse Embryo." *Developmental Dynamics* 237 (3): 618–29. https://doi.org/10.1002/dvdy.21444.
Garneau, Josiane E., Marie-Ève Dupuis, Manuela Villion, Dennis A. Romero, Rodolphe Barrangou, Patrick Boyaval, Christophe Fremaux, Philippe Horvath, Alfonso H. Magadán, and Sylvain Moineau. 2010. "The CRISPR/Cas Bacterial Immune System Cleaves Bacteriophage and Plasmid DNA." *Nature* 468 (7320): 67–71. https://doi.org/10.1038/nature09523.
Gasiunas, G., R. Barrangou, P. Horvath, and V. Siksnys. 2012. "Cas9-CrRNA Ribonucleoprotein Complex Mediates Specific DNA Cleavage for Adaptive Immunity in Bacteria." *Proceedings of the National Academy of Sciences* 109 (39): E2579–86. https://doi.org/10.1073/pnas.1208507109.
Geurts, Maarten H., Eyleen de Poel, Gimano D. Amatngalim, Rurika Oka, Fleur M. Meijers, Evelien Kruisselbrink, Peter van Mourik, et al. 2020. "CRISPR-Based Adenine Editors Correct Nonsense Mutations in a Cystic Fibrosis Organoid Biobank." *Cell Stem Cell* 26 (4): 503-510.e7. https://doi.org/10.1016/j.stem.2020.01.019.
Goff, Stephen P., and Leslie I. Lobel. 1987. "Mutants of Murine Leukemia Viruses and Retroviral Replication." *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 907 (2): 93–123. https://doi.org/10.1016/0304-419X(87)90001-1.
Gonsky, Jason, Eran Bacharach, and Stephen P. Goff. 2001. "Identification of Residues of the Moloney Murine Leukemia Virus Nucleocapsid Critical for Viral DNA Synthesis In Vivo." *Journal of Virology* 75 (6): 2616–26. https://doi.org/10.1128/JVI.75.6.2616-2626.2001.

Grimm, C., E. Lund, and J. E. Dahlberg. 1997. “Selection and Nuclear Immobilization of Exportable RNAs.” *Proceedings of the National Academy of Sciences* 94 (19): 10122–27. https://doi.org/10.1073/pnas.94.19.10122.

Gruber, A. R., R. Lorenz, S. H. Bernhart, R. Neubock, and I. L. Hofacker. 2008. “The Vienna RNA Websuite.” *Nucleic Acids Research* 36 (Web Server): W70–74. https://doi.org/10.1093/nar/gkn188.

Gul, Saima, Sher Bahadar Khan, Inayat Ur Rehman, Murad Ali Khan, and M. I. Khan. 2019. “A Comprehensive Review of Magnetic Nanomaterials Modern Day Theranostics.” *Frontiers in Materials* 6 (July): 179. https://doi.org/10.3389/fmats.2019.00179.

Gupta, Shweta, Tanya M-Redmond, Fan Meng, Andrew Tidball, Huda Akil, Stanley Watson, Jack M. Parent, and Michael Uhler. 2018. “Fibroblast Growth Factor 2 Regulates Activity and Gene Expression of Human Post-mitotic Excitatory Neurons.” *Journal of Neurochemistry* 145 (3): 188–203. https://doi.org/10.1111/jnc.14255.

Hamard-Peron, E., F. Juillard, J. S. Saad, C. Roy, P. Roingeard, M. F. Summers, J.-L. Darlix, C. Picart, and D. Muriaux. 2010. “Targeting of Murine Leukemia Virus Gag to the Plasma Membrane Is Mediated by PI(4,5)P $_2$ /PS and a Polybasic Region in the Matrix.” *Journal of Virology* 84 (1): 503–15. https://doi.org/10.1128/JVI.01134-09.

Hamilton, Holly L., and Joseph P. Dillard. 2006. “Natural Transformation of *Neisseria Gonorrhoeae* : From DNA Donation to Homologous Recombination: Natural Transformation of *Neisseria Gonorrhoeae*.” *Molecular Microbiology* 59 (2): 376–85. https://doi.org/10.1111/j.1365-2958.2005.04964.x.

Hanlon, Killian S., Benjamin P. Kleinstiver, Sara P. Garcia, Mikołaj P. Zaborowski, Adrienn Volak, Stefan E. Spirig, Alissa Muller, et al. 2019. “High Levels of AAV Vector Integration into CRISPR-Induced DNA Breaks.” *Nature Communications* 10 (1): 4439. https://doi.org/10.1038/s41467-019-12449-2.

Harrington, Lucas B., Kevin W. Doxzen, Enbo Ma, Jun-Jie Liu, Gavin J. Knott, Alireza Edraki, Bianca Garcia, et al. 2017. “A Broad-Spectrum Inhibitor of CRISPR-Cas9.” *Cell* 170 (6): 1224-1233.e15. https://doi.org/10.1016/j.cell.2017.07.037.

Herdeis, Lorenz, Daniel Gerlach, Darryl B. McConnell, and Dirk Kessler. 2021. “Stopping the Beating Heart of Cancer: KRAS Reviewed.” *Current Opinion in Structural Biology* 71 (December): 136–47. https://doi.org/10.1016/j.sbi.2021.06.013.

Hou, Z., Y. Zhang, N. E. Propson, S. E. Howden, L.-F. Chu, E. J. Sontheimer, and J. A. Thomson. 2013. “Efficient Genome Engineering in Human Pluripotent Stem Cells Using Cas9 from Neisseria Meningitidis.” *Proceedings of the National Academy of Sciences* 110 (39): 15644–49. https://doi.org/10.1073/pnas.1313587110.

Hsieh-Feng, Vicki, and Yinong Yang. 2020. “Efficient Expression of Multiple Guide RNAs for CRISPR/Cas Genome Editing.” *ABIOTECH* 1 (2): 123–34. https://doi.org/10.1007/s42994-019-00014-w.

Hwang, Woong Y, Yanfang Fu, Deepak Reyon, Morgan L Maeder, Shengdar Q Tsai, Jeffry D Sander, Randall T Peterson, J-R Joanna Yeh, and J Keith Joung. 2013. “Efficient Genome Editing in Zebrafish Using a CRISPR-Cas System.” *Nature Biotechnology* 31 (3): 227–29. https://doi.org/10.1038/nbt.2501.

Hynes, Alexander P., Geneviève M. Rousseau, Marie-Laurence Lemay, Philippe Horvath, Dennis A. Romero, Christophe Fremaux, and Sylvain Moineau. 2017. “An Anti-CRISPR from a Virulent Streptococcal Phage Inhibits Streptococcus Pyogenes Cas9.” *Nature Microbiology* 2 (10): 1374–80. https://doi.org/10.1038/s41564-017-0004-7.

Ishino, Y, H Shinagawa, K Makino, M Amemura, and A Nakata. 1987. “Nucleotide Sequence of the Iap Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in Escherichia Coli, and Identification of the Gene Product.” *Journal of Bacteriology* 169 (12): 5429–33. https://doi.org/10.1128/jb.169.12.5429-5433.1987.

Jangra, Rohit K., Andrew S. Herbert, Rong Li, Lucas T. Jae, Lara M. Kleinfelter, Megan M. Slough, Sarah L. Barker, et al. 2018. “Protocadherin-1 Is Essential for Cell Entry by New World Hantaviruses.” *Nature* 563 (7732): 559–63. https://doi.org/10.1038/s41586-018-0702-1.

Jarrosson-Wuilleme, Loraine, Caroline Goujon, Jeanine Bernaud, Dominique Rigal, Jean-Luc Darlix, and Andrea Cimarelli. 2006. “Transduction of Nondividing Human Macrophages with Gammaretrovirus-Derived Vectors.” *Journal of Virology* 80 (3): 1152–59. https://doi.org/10.1128/JVI.80.3.1152-1159.2006.

Jensen, Kristopher Torp, Lasse Fløe, Trine Skov Petersen, Jinrong Huang, Fengping Xu, Lars Bolund, Yonglun Luo, and Lin Lin. 2017. “Chromatin Accessibility and Guide Sequence Secondary Structure Affect CRISPR-Cas9 Gene Editing Efficiency.” *FEBS Letters* 591 (13): 1892–1901. https://doi.org/10.1002/1873-3468.12707.

Jiang, Wenyan, David Bikard, David Cox, Feng Zhang, and Luciano A Marraffini. 2013. “RNA-Guided Editing of Bacterial Genomes Using CRISPR-Cas Systems.” *Nature Biotechnology* 31 (3): 233–39. https://doi.org/10.1038/nbt.2508.

Jiang, Wenyan, Poulami Samai, and Luciano A. Marraffini. 2016. “Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity.” *Cell* 164 (4): 710–21. https://doi.org/10.1016/j.cell.2015.12.053.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. “A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity.” *Science* 337 (6096): 816–21. https://doi.org/10.1126/science.1225829.

Jinek, Martin, Alexandra East, Aaron Cheng, Steven Lin, Enbo Ma, and Jennifer Doudna. 2013. “RNA-Programmed Genome Editing in Human Cells.” *ELife* 2 (January): e00471. https://doi.org/10.7554/eLife.00471.

Johnson, S. F., J. T. Collins, V. M. D’Souza, and A. Telesnitsky. 2014. “Determinants of Moloney Murine Leukemia Virus Gag-Pol and Genomic RNA Proportions.” *Journal of Virology* 88 (13): 7267–75. https://doi.org/10.1128/JVI.03513-13.

Kaczmarczyk, S. J., K. Sitaraman, H. A. Young, S. H. Hughes, and D. K. Chatterjee. 2011. “Protein Delivery Using Engineered Virus-like Particles.” *Proceedings of the National Academy of Sciences* 108 (41): 16998–3. https://doi.org/10.1073/pnas.1101874108.

Kantor, Ariel, Michelle McClements, and Robert MacLaren. 2020. “CRISPR-Cas9 DNA Base-Editing and Prime-Editing.” *International Journal of Molecular Sciences* 21 (17): 6240. https://doi.org/10.3390/ijms21176240.

Kara, Goknur, George A. Calin, and Bulent Ozpolat. 2022. “RNAi-Based Therapeutics and Tumor Targeted Delivery in Cancer.” *Advanced Drug Delivery Reviews* 182 (March): 114113. https://doi.org/10.1016/j.addr.2022.114113.

Kazlauskiene, Migle, Georgij Kostiuk, Česlovas Venclovas, Gintautas Tamulaitis, and Virginijus Siksnys. 2017. “A Cyclic Oligonucleotide Signaling Pathway in Type III CRISPR-Cas Systems.” *Science* 357 (6351): 605–9. https://doi.org/10.1126/science.aao0100.

Kheirollah, Alireza, Dept. of Biochemistry, Medical School, Cellular and Molecular Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran, Mostafa

Chashmpoosh, Department of Biochemistry, Faculty of Medical Sciences, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Ira, Asma Mohammadi, Department of Biochemistry, Faculty of Medical Sciences, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Ira, Neda Abdovis, Department of Biochemistry, Faculty of Medical Sciences, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Ira, Hamide Ghasempoor, and Department of Biochemistry, Faculty of Medical Sciences, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Ira. 2016. "Isolation and Culture of Mouse Newborn Skin Fibroblast Cells in a Simple and Inexpensive Method." *Pars of Jahrom University of Medical Sciences* 14 (3): 35–42. https://doi.org/10.29252/jmj.14.3.35.

Kim, Eunji, Taeyoung Koo, Sung Wook Park, Daesik Kim, Kyoungmi Kim, Hee-Yeon Cho, Dong Woo Song, et al. 2017. "In Vivo Genome Editing with a Small Cas9 Orthologue Derived from Campylobacter Jejuni." *Nature Communications* 8 (1): 14500. https://doi.org/10.1038/ncomms14500.

Kim, Goo-Young, Jeong-Mi Moon, Ji-Hye Han, Kyung-Hee Kim, and Hyangshuk Rhim. 2011. "The SCMV IE Enhancer/Promoter System for High-Level Expression and Efficient Functional Studies of Target Genes in Mammalian Cells and Zebrafish." *Biotechnology Letters* 33 (7): 1319–26. https://doi.org/10.1007/s10529-011-0589-5.

Kleinstiver, Benjamin P., Vikram Pattanayak, Michelle S. Prew, Shengdar Q. Tsai, Nhu T. Nguyen, Zongli Zheng, and J. Keith Joung. 2016. "High-Fidelity CRISPR–Cas9 Nucleases with No Detectable Genome-Wide off-Target Effects." *Nature* 529 (7587): 490–95. https://doi.org/10.1038/nature16526.

Knopp, Yvonne, Franziska K. Geis, Dirk Heckl, Stefan Horn, Thomas Neumann, Johannes Kuehle, Janine Meyer, et al. 2018. "Transient Retrovirus-Based CRISPR/Cas9 All-in-One Particles for Efficient, Targeted Gene Knockout." *Molecular Therapy - Nucleic Acids* 13 (December): 256–74. https://doi.org/10.1016/j.omtn.2018.09.006.

Kol, Nitzan, Micha Gladnikoff, David Barlam, Roni Z. Shneck, Alan Rein, and Itay Rousso. 2006. "Mechanical Properties of Murine Leukemia Virus Particles: Effect of Maturation." *Biophysical Journal* 91 (2): 767–74. https://doi.org/10.1529/biophysj.105.079657.

Koonin, Eugene V, Kira S Makarova, and Feng Zhang. 2017. "Diversity, Classification and Evolution of CRISPR-Cas Systems." *Current Opinion in Microbiology* 37 (June): 67–78. https://doi.org/10.1016/j.mib.2017.05.008.

Lange, Allison, Ryan E. Mills, Christopher J. Lange, Murray Stewart, Scott E. Devine, and Anita H. Corbett. 2007. "Classical Nuclear Localization Signals: Definition, Function, and Interaction with Importin α." *Journal of Biological Chemistry* 282 (8): 5101–5. https://doi.org/10.1074/jbc.R600026200.

Lee, Ciaran M, Thomas J Cradick, and Gang Bao. 2016. "The Neisseria Meningitidis CRISPR-Cas9 System Enables Specific Genome Editing in Mammalian Cells." *Molecular Therapy* 24 (3): 645–54. https://doi.org/10.1038/mt.2016.8.

Li, Fei, Jing Jin, Christin Herrmann, and Walther Mothes. 2013. "Basic Residues in the Matrix Domain and Multimerization Target Murine Leukemia Virus Gag to the Virological Synapse." *Journal of Virology* 87 (12): 7113–26. https://doi.org/10.1128/JVI.03263-12.

Li, Xuan, and Wolf-Dietrich Heyer. 2008. "Homologous Recombination in DNA Repair and DNA Damage Tolerance." *Cell Research* 18 (1): 99–113. https://doi.org/10.1038/cr.2008.1.

Lindemann, D, M Bock, M Schweizer, and A Rethwilm. 1997. “Efficient Pseudotyping of Murine Leukemia Virus Particles with Chimeric Human Foamy Virus Envelope Proteins.” *Journal of Virology* 71 (6): 4815–20. https://doi.org/10.1128/jvi.71.6.4815-4820.1997.

Lino, Christopher A., Jason C. Harper, James P. Carney, and Jerilyn A. Timlin. 2018. “Delivering CRISPR: A Review of the Challenges and Approaches.” *Drug Delivery* 25 (1): 1234–57. https://doi.org/10.1080/10717544.2018.1474964.

Liu, Yao, Xiangyang Li, Siting He, Shuhong Huang, Chao Li, Yulin Chen, Zhen Liu, Xingxu Huang, and Xiaolong Wang. 2020. “Efficient Generation of Mouse Models with the Prime Editing System.” *Cell Discovery* 6 (1): 27. https://doi.org/10.1038/s41421-020-0165-z.

Lu, Chi-Wei, and Monica J. Roth. 2001. “Functional Characterization of the N Termini of Murine Leukemia Virus Envelope Proteins.” *Journal of Virology* 75 (9): 4357–66. https://doi.org/10.1128/JVI.75.9.4357-4366.2001.

Lund, E., and J.E. Dahlberg. 2006. “Substrate Selectivity of Exportin 5 and Dicer in the Biogenesis of MicroRNAs.” *Cold Spring Harbor Symposia on Quantitative Biology* 71 (0): 59–66. https://doi.org/10.1101/sqb.2006.71.050.

Ma, Enbo, Lucas B. Harrington, Mitchell R. O’Connell, Kaihong Zhou, and Jennifer A. Doudna. 2015. “Single-Stranded DNA Cleavage by Divergent CRISPR-Cas9 Enzymes.” *Molecular Cell* 60 (3): 398–407. https://doi.org/10.1016/j.molcel.2015.10.030.

Madisen, Linda, Theresa A Zwingman, Susan M Sunkin, Seung Wook Oh, Hatim A Zariwala, Hong Gu, Lydia L Ng, et al. 2010. “A Robust and High-Throughput Cre Reporting and Characterization System for the Whole Mouse Brain.” *Nature Neuroscience* 13 (1): 133–40. https://doi.org/10.1038/nn.2467.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church. 2013. “RNA-Guided Human Genome Engineering via Cas9.” *Science* 339 (6121): 823–26. https://doi.org/10.1126/science.1232033.

Mali, Prashant, John Aach, P Benjamin Stranges, Kevin M Esvelt, Mark Moosburner, Sriram Kosuri, Luhan Yang, and George M Church. 2013. “CAS9 Transcriptional Activators for Target Specificity Screening and Paired Nickases for Cooperative Genome Engineering.” *Nature Biotechnology* 31 (9): 833–38. https://doi.org/10.1038/nbt.2675.

Mangeot, Philippe E., Valérie Risson, Floriane Fusil, Aline Marnef, Emilie Laurent, Juliana Blin, Virginie Mournetas, et al. 2019. “Genome Editing in Primary Cells and in Vivo Using Viral-Derived Nanoblades Loaded with Cas9-SgRNA Ribonucleoproteins.” *Nature Communications* 10 (1): 45. https://doi.org/10.1038/s41467-018-07845-z.

Marraffini, Luciano A. 2015. “CRISPR-Cas Immunity in Prokaryotes.” *Nature* 526 (7571): 55–61. https://doi.org/10.1038/nature15386.

Marraffini, Luciano A., and Erik J. Sontheimer. 2008. “CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA.” *Science* 322 (5909): 1843–45. https://doi.org/10.1126/science.1165771.

Mayo, Keith, Jason McDermott, and Eric Barklis. 2002. “Hexagonal Organization of Moloney Murine Leukemia Virus Capsid Proteins.” *Virology* 298 (1): 30–38. https://doi.org/10.1006/viro.2002.1452.

Mir, Aamir, Julia F. Alterman, Matthew R. Hassler, Alexandre J. Debacker, Edward Hudgens, Dimas Echeverria, Michael H. Brodsky, Anastasia Khvorova, Jonathan K. Watts, and Erik J. Sontheimer. 2018. “Heavily and Fully Modified RNAs Guide Efficient SpyCas9-

Mediated Genome Editing." *Nature Communications* 9 (1): 2641. https://doi.org/10.1038/s41467-018-05073-z.
Mir, Aamir, Alireza Edraki, Jooyoung Lee, and Erik J. Sontheimer. 2018. "Type II-C CRISPR-Cas9 Biology, Mechanism, and Application." *ACS Chemical Biology* 13 (2): 357–65. https://doi.org/10.1021/acschembio.7b00855.
Montagna, Claudia, Gianluca Petris, Antonio Casini, Giulia Maule, Gian Marco Franceschini, Ilaria Zanella, Luciano Conti, et al. 2018. "VSV-G-Enveloped Vesicles for Traceless Delivery of CRISPR-Cas9." *Molecular Therapy - Nucleic Acids* 12 (September): 453–62. https://doi.org/10.1016/j.omtn.2018.05.010.
Muriaux, Delphine, Sylvain Costes, Kunio Nagashima, Jane Mirro, Ed Cho, Stephen Lockett, and Alan Rein. 2004. "Role of Murine Leukemia Virus Nucleocapsid Protein in Virus Assembly." *Journal of Virology* 78 (22): 12378–85. https://doi.org/10.1128/JVI.78.22.12378-12385.2004.
Muriaux, Delphine, Jane Mirro, Kunio Nagashima, Demetria Harvin, and Alan Rein. 2002. "Murine Leukemia Virus Nucleocapsid Mutant Particles Lacking Viral RNA Encapsidate Ribosomes." *Journal of Virology* 76 (22): 11405–13. https://doi.org/10.1128/JVI.76.22.11405-11413.2002.
Nelles, David A., Mark Y. Fang, Stefan Aigner, and Gene W. Yeo. 2015. "Applications of Cas9 as an RNA-Programmed RNA-Binding Protein." *BioEssays* 37 (7): 732–39. https://doi.org/10.1002/bies.201500001.
Nelles, David A., Mark Y. Fang, Mitchell R. O'Connell, Jia L. Xu, Sebastian J. Markmiller, Jennifer A. Doudna, and Gene W. Yeo. 2016. "Programmable RNA Tracking in Live Cells with CRISPR/Cas9." *Cell* 165 (2): 488–96. https://doi.org/10.1016/j.cell.2016.02.054.
Nie, Linghu, Meghna Das Thakur, Yumei Wang, Qin Su, Yongliang Zhao, and Yunfeng Feng. 2010. "Regulation of U6 Promoter Activity by Transcriptional Interference in Viral Vector-Based RNAi." *Genomics, Proteomics & Bioinformatics* 8 (3): 170–79. https://doi.org/10.1016/S1672-0229(10)60019-8.
Niewoehner, Ole, Carmela Garcia-Doval, Jakob T. Rostøl, Christian Berk, Frank Schwede, Laurent Bigler, Jonathan Hall, Luciano A. Marraffini, and Martin Jinek. 2017. "Type III CRISPR–Cas Systems Produce Cyclic Oligoadenylate Second Messengers." *Nature* 548 (7669): 543–48. https://doi.org/10.1038/nature23467.
Nishimasu, Hiroshi, Xi Shi, Soh Ishiguro, Linyi Gao, Seiichi Hirano, Sae Okazaki, Taichi Noda, et al. 2018. "Engineered CRISPR-Cas9 Nuclease with Expanded Targeting Space." *Science* 361 (6408): 1259–62. https://doi.org/10.1126/science.aas9129.
Nooraei, Saghi, Howra Bahrulolum, Zakieh Sadat Hoseini, Camellia Katalani, Abbas Hajizade, Andrew J. Easton, and Gholamreza Ahmadian. 2021. "Virus-like Particles: Preparation, Immunogenicity and Their Roles as Nanovaccines and Drug Nanocarriers." *Journal of Nanobiotechnology* 19 (1): 59. https://doi.org/10.1186/s12951-021-00806-7.
O'Connell, Mitchell R., Benjamin L. Oakes, Samuel H. Sternberg, Alexandra East-Seletsky, Matias Kaplan, and Jennifer A. Doudna. 2014. "Programmable RNA Recognition and Cleavage by CRISPR/Cas9." *Nature* 516 (7530): 263–66. https://doi.org/10.1038/nature13769.
Oksayan, Sibil, Linda Wiltzer, Caitlin L. Rowe, Danielle Blondel, David A. Jans, and Gregory W. Moseley. 2012. "A Novel Nuclear Trafficking Module Regulates the Nucleocytoplasmic Localization of the Rabies Virus Interferon Antagonist, P Protein."

Journal of Biological Chemistry 287 (33): 28112–21. https://doi.org/10.1074/jbc.M112.374694.

O'Leary, Monique N., Katherine H. Schreiber, Yong Zhang, Anne-Cécile E. Duc, Shuyun Rao, J. Scott Hale, Emmeline C. Academia, et al. 2013. "The Ribosomal Protein Rpl22 Controls Ribosome Composition by Directly Repressing Expression of Its Own Paralog, Rpl22l1." Edited by Gregory P. Copenhaver. *PLoS Genetics* 9 (8): e1003708. https://doi.org/10.1371/journal.pgen.1003708.

O'Reilly, Lucille, and Monica J. Roth. 2003. "G541R within the 4070A TM Protein Regulates Fusion in Murine Leukemia Viruses." *Journal of Virology* 77 (22): 12011–21. https://doi.org/10.1128/JVI.77.22.12011-12021.2003.

Oshima, Masamichi, Delphine Muriaux, Jane Mirro, Kunio Nagashima, Kelly Dryden, Mark Yeager, and Alan Rein. 2004. "Effects of Blocking Individual Maturation Cleavages in Murine Leukemia Virus Gag." *Journal of Virology* 78 (3): 1411–20. https://doi.org/10.1128/JVI.78.3.1411-1420.2004.

Paddison, Patrick J., and Peter K. Vogt, eds. 2008. *RNA Interference*. Current Topics in Microbiology and Immunology 320. Berlin Heidelberg: Springer.

Pawluk, April, Nadia Amrani, Yan Zhang, Bianca Garcia, Yurima Hidalgo-Reyes, Jooyoung Lee, Alireza Edraki, et al. 2016. "Naturally Occurring Off-Switches for CRISPR-Cas9." *Cell* 167 (7): 1829-1838.e9. https://doi.org/10.1016/j.cell.2016.11.017.

Petri, Karl, Weiting Zhang, Junyan Ma, Andrea Schmidts, Hyunho Lee, Joy E. Horng, Daniel Y. Kim, et al. 2021. "CRISPR Prime Editing with Ribonucleoprotein Complexes in Zebrafish and Primary Human Cells." *Nature Biotechnology*, April. https://doi.org/10.1038/s41587-021-00901-y.

Pommer, Ansgar J, Santiago Cal, Anthony H Keeble, Daniel Walker, Steven J Evans, Ulrike C Kühlmann, Alan Cooper, et al. 2001. "Mechanism and Cleavage Specificity of the H-N-H Endonuclease Colicin E9 1 1Edited by J. Karn." *Journal of Molecular Biology* 314 (4): 735–49. https://doi.org/10.1006/jmbi.2001.5189.

Proudfoot, Nick J. 2011. "Ending the Message: Poly(A) Signals Then and Now." *Genes & Development* 25 (17): 1770–82. https://doi.org/10.1101/gad.17268411.

Radici, Lucia, Marzia Bianchi, Rita Crinelli, and Mauro Magnani. 2013. "Ubiquitin C Gene: Structure, Function, and Transcriptional Regulation." *Advances in Bioscience and Biotechnology* 04 (12): 1057–62. https://doi.org/10.4236/abb.2013.412141.

Ranade, Devika, Roopali Pradhan, Muhunden Jayakrishnan, Sushmitha Hegde, and Kundan Sengupta. 2019. "Lamin A/C and Emerin Depletion Impacts Chromatin Organization and Dynamics in the Interphase Nucleus." *BMC Molecular and Cell Biology* 20 (1): 11. https://doi.org/10.1186/s12860-019-0192-5.

Rauch, Benjamin J., Melanie R. Silvis, Judd F. Hultquist, Christopher S. Waters, Michael J. McGregor, Nevan J. Krogan, and Joseph Bondy-Denomy. 2017. "Inhibition of CRISPR-Cas9 with Bacteriophage Proteins." *Cell* 168 (1–2): 150-158.e10. https://doi.org/10.1016/j.cell.2016.12.009.

Ray, Neelanjana, Jillian Whidby, Shaun Stewart, Jay W. Hooper, and Andrea Bertolotti-Ciarlet. 2010. "Study of Andes Virus Entry and Neutralization Using a Pseudovirion System." *Journal of Virological Methods* 163 (2): 416–23. https://doi.org/10.1016/j.jviromet.2009.11.004.

Redies, Christoph, Jessica Heyder, Tomáš Kohoutek, Katrien Staes, and Frans Van Roy. 2008. "Expression of Protocadherin-1 (Pcdh1) during Mouse Development." *Developmental Dynamics* 237 (9): 2496–2505. https://doi.org/10.1002/dvdy.21650.

Rees, Holly A., Alex C. Minella, Cameron A. Burnett, Alexis C. Komor, and Nicole M. Gaudelli. 2021. "CRISPR-Derived Genome Editing Therapies: Progress from Bench to Bedside." *Molecular Therapy* 29 (11): 3125–39. https://doi.org/10.1016/j.ymthe.2021.09.027.

Rein, Alan. 2011. "Murine Leukemia Viruses: Objects and Organisms." *Advances in Virology* 2011: 1–14. https://doi.org/10.1155/2011/403419.

Richardson, Christopher D, Graham J Ray, Mark A DeWitt, Gemma L Curie, and Jacob E Corn. 2016. "Enhancing Homology-Directed Genome Editing by Catalytically Active and Inactive CRISPR-Cas9 Using Asymmetric Donor DNA." *Nature Biotechnology* 34 (3): 339–44. https://doi.org/10.1038/nbt.3481.

Risco, Cristina, Luis Menéndez-Arias, Terry D Copeland, and Stephen Oroszlan. 1995. "Intracellular Transport of the Murine Leukemia Virus during Acute Infection of NIH 3T3 Cells: Nuclear Import of Nucleocapsid Protein and Integrase." *Journal of Cell Science* 108: 3039–50.

Robert, Marc-André, Viktoria Lytvyn, Francis Deforet, Rénald Gilbert, and Bruno Gaillet. 2017. "Virus-Like Particles Derived from HIV-1 for Delivery of Nuclear Proteins: Improvement of Production and Activity by Protein Engineering." *Molecular Biotechnology* 59 (1): 9–23. https://doi.org/10.1007/s12033-016-9987-1.

Rousseau, Beth A., Zhonggang Hou, Max J. Gramelspacher, and Yan Zhang. 2018. "Programmable RNA Cleavage and Recognition by a Natural CRISPR-Cas9 System from Neisseria Meningitidis." *Molecular Cell* 69 (5): 906-914.e4. https://doi.org/10.1016/j.molcel.2018.01.025.

Roy, Christine, Naceur Tounekti, MarylÈne Mougel, Jean-Luc Darlix, Claude Paoletti, Chantal Ehresmann, Bernard Ehresmann, and Jacques Paoletti. 1990. "An Analytical Study of the Dimerization of *in Vitro* Generated RNA of Moloney Murine Leukemia Virus MoMuLV." *Nucleic Acids Research* 18 (24): 7287–92. https://doi.org/10.1093/nar/18.24.7287.

Rupp, R A, L Snider, and H Weintraub. 1994. "Xenopus Embryos Regulate the Nuclear Localization of XMyoD." *Genes & Development* 8 (11): 1311–23. https://doi.org/10.1101/gad.8.11.1311.

Samaan, George, Danielle Yugo, Sangeetha Rajagopalan, Jonathan Wall, Robert Donnell, Dan Goldowitz, Rajaram Gopalakrishnan, and Sundaresan Venkatachalam. 2010. "Foxn3 Is Essential for Craniofacial Development in Mice and a Putative Candidate Involved in Human Congenital Craniofacial Defects." *Biochemical and Biophysical Research Communications* 400 (1): 60–65. https://doi.org/10.1016/j.bbrc.2010.07.142.

Samai, Poulami, Nora Pyenson, Wenyan Jiang, Gregory W. Goldberg, Asma Hatoum-Aslan, and Luciano A. Marraffini. 2015. "Co-Transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity." *Cell* 161 (5): 1164–74. https://doi.org/10.1016/j.cell.2015.04.027.

Sampson, Timothy R., Sunil D. Saroj, Anna C. Llewellyn, Yih-Ling Tzeng, and David S. Weiss. 2013. "A CRISPR/Cas System Mediates Bacterial Innate Immune Evasion and Virulence." *Nature* 497 (7448): 254–57. https://doi.org/10.1038/nature12048.

Samulski, R.J., X. Zhu, X. Xiao, J.D. Brook, D.E. Housman, N. Epstein, and L.A. Hunter. 1991. "Targeted Integration of Adeno-Associated Virus (AAV) into Human Chromosome 19." *The EMBO Journal* 10 (12): 3941–50. https://doi.org/10.1002/j.1460-2075.1991.tb04964.x.

Sanson, Kendall R., Ruth E. Hanna, Mudra Hegde, Katherine F. Donovan, Christine Strand, Meagan E. Sullender, Emma W. Vaimberg, et al. 2018. "Optimized Libraries for CRISPR-Cas9 Genetic Screens with Multiple Modalities." *Nature Communications* 9 (1): 5416. https://doi.org/10.1038/s41467-018-07901-8.

Sanz, E., L. Yang, T. Su, D. R. Morris, G. S. McKnight, and P. S. Amieux. 2009. "Cell-Type-Specific Isolation of Ribosome-Associated MRNA from Complex Tissues." *Proceedings of the National Academy of Sciences* 106 (33): 13939–44. https://doi.org/10.1073/pnas.0907143106.

Schene, Imre F., Indi P. Joore, Rurika Oka, Michal Mokry, Anke H. M. van Vugt, Ruben van Boxtel, Hubert P. J. van der Doef, et al. 2020. "Prime Editing for Functional Repair in Patient-Derived Disease Models." *Nature Communications* 11 (1): 5352. https://doi.org/10.1038/s41467-020-19136-7.

Schmidt, Moritz J., Ashish Gupta, Christien Bednarski, Stefanie Gehrig-Giannini, Florian Richter, Christian Pitzler, Michael Gamalinda, et al. 2021. "Improved CRISPR Genome Editing Using Small Highly Active and Specific Engineered RNA-Guided Nucleases." *Nature Communications* 12 (1): 4219. https://doi.org/10.1038/s41467-021-24454-5.

Schnierle, B. S., J. Stitz, V. Bosch, F. Nocken, H. Merget-Millitzer, M. Engelstadter, R. Kurth, B. Groner, and K. Cichutek. 1997. "Pseudotyping of Murine Leukemia Virus with the Envelope Glycoproteins of HIV Generates a Retroviral Vector with Specificity of Infection for CD4-Expressing Cells." *Proceedings of the National Academy of Sciences* 94 (16): 8640–45. https://doi.org/10.1073/pnas.94.16.8640.

Scott, Kenneth L., and Sharon E. Plon. 2005. "CHES1/FOXN3 Interacts with Ski-Interacting Protein and Acts as a Transcriptional Repressor." *Gene* 359 (October): 119–26. https://doi.org/10.1016/j.gene.2005.06.014.

Senturk, Serif, Nitin H. Shirole, Dawid G. Nowak, Vincenzo Corbo, Debjani Pal, Alexander Vaughan, David A. Tuveson, Lloyd C. Trotman, Justin B. Kinney, and Raffaella Sordella. 2017. "Rapid and Tunable Method to Temporally Control Gene Editing Based on Conditional Cas9 Stabilization." *Nature Communications* 8 (1): 14370. https://doi.org/10.1038/ncomms14370.

Skene, Peter J, Jorja G Henikoff, and Steven Henikoff. 2018. "Targeted in Situ Genome-Wide Profiling with High Efficiency for Low Cell Numbers." *Nature Protocols* 13 (5): 1006–19. https://doi.org/10.1038/nprot.2018.015.

Skipper, Kristian Alsbjerg, Mathias Gaarde Nielsen, Sofie Andersen, Laura Barrett Ryø, Rasmus O. Bak, and Jacob Giehm Mikkelsen. 2018. "Time-Restricted PiggyBac DNA Transposition by Transposase Protein Delivery Using Lentivirus-Derived Nanoparticles." *Molecular Therapy - Nucleic Acids* 11 (June): 253–62. https://doi.org/10.1016/j.omtn.2018.02.006.

Slaymaker, Ian M., Linyi Gao, Bernd Zetsche, David A. Scott, Winston X. Yan, and Feng Zhang. 2016. "Rationally Engineered Cas9 Nucleases with Improved Specificity." *Science* 351 (6268): 84–88. https://doi.org/10.1126/science.aad5227.

Slough, Megan M., Kartik Chandran, and Rohit K. Jangra. 2019. "Two Point Mutations in Old World Hantavirus Glycoproteins Afford the Generation of Highly Infectious

Recombinant Vesicular Stomatitis Virus Vectors." Edited by Anne Moscona. *MBio* 10 (1). https://doi.org/10.1128/mBio.02372-18.
Soneoka, Y, S M Kingsman, and A J Kingsman. 1997. "Mutagenesis Analysis of the Murine Leukemia Virus Matrix Protein: Identification of Regions Important for Membrane Localization and Intracellular Transport." *Journal of Virology* 71 (7): 5549–59. https://doi.org/10.1128/jvi.71.7.5549-5559.1997.
Staahl, Brett T, Madhurima Benekareddy, Claire Coulon-Bainier, Ashwin A Banfal, Stephen N Floor, Jennifer K Sabo, Cole Urnes, Gabriela Acevedo Munares, Anirvan Ghosh, and Jennifer A Doudna. 2017. "Efficient Genome Editing in the Mouse Brain by Local Delivery of Engineered Cas9 Ribonucleoprotein Complexes." *Nature Biotechnology* 35 (5): 431–34. https://doi.org/10.1038/nbt.3806.
Sternberg, Samuel H., Benjamin LaFrance, Matias Kaplan, and Jennifer A. Doudna. 2015. "Conformational Control of DNA Target Cleavage by CRISPR–Cas9." *Nature* 527 (7576): 110–13. https://doi.org/10.1038/nature15544.
Sternberg, Samuel H., Sy Redding, Martin Jinek, Eric C. Greene, and Jennifer A. Doudna. 2014. "DNA Interrogation by the CRISPR RNA-Guided Endonuclease Cas9." *Nature* 507 (7490): 62–67. https://doi.org/10.1038/nature13011.
Stewart, Martin P., Robert Langer, and Klavs F. Jensen. 2018. "Intracellular Delivery by Membrane Disruption: Mechanisms, Strategies, and Concepts." *Chemical Reviews* 118 (16): 7409–7531. https://doi.org/10.1021/acs.chemrev.7b00678.
Strutt, Steven C, Rachel M Torrez, Emine Kaya, Oscar A Negrete, and Jennifer A Doudna. 2018. "RNA-Dependent RNA Targeting by CRISPR-Cas9." *ELife* 7 (January): e32724. https://doi.org/10.7554/eLife.32724.
Telesnitsky, A., and S.P. Goff. 1993. "Two Defective Forms of Reverse Transcriptase Can Complement to Restore Retroviral Infectivity." *The EMBO Journal* 12 (11): 4433–38. https://doi.org/10.1002/j.1460-2075.1993.tb06128.x.
Tessier, Tanner M., Mackenzie J. Dodge, Martin A. Prusinkiewicz, and Joe S. Mymryk. 2019. "Viral Appropriation: Laying Claim to Host Nuclear Transport Machinery." *Cells* 8 (6): 559. https://doi.org/10.3390/cells8060559.
Tian, Shouwei, Linjian Jiang, Xiaxia Cui, Jie Zhang, Shaogui Guo, Maoying Li, Haiying Zhang, et al. 2018. "Engineering Herbicide-Resistant Watermelon Variety through CRISPR/Cas9-Mediated Base-Editing." *Plant Cell Reports* 37 (9): 1353–56. https://doi.org/10.1007/s00299-018-2299-0.
Todorova, Roumiana. 2011. "Comparative Analysis of the Methods of Drug and Protein Delivery for the Treatment of Cancer, Genetic Diseases and Diagnostics." *Drug Delivery* 18 (8): 586–98. https://doi.org/10.3109/10717544.2011.600783.
Tsai, Shengdar Q, Zongli Zheng, Nhu T Nguyen, Matthew Liebers, Ved V Topkar, Vishal Thapar, Nicolas Wyvekens, et al. 2015. "GUIDE-Seq Enables Genome-Wide Profiling of off-Target Cleavage by CRISPR-Cas Nucleases." *Nature Biotechnology* 33 (2): 187–97. https://doi.org/10.1038/nbt.3117.
Turner, D L, and H Weintraub. 1994. "Expression of Achaete-Scute Homolog 3 in Xenopus Embryos Converts Ectodermal Cells to a Neural Fate." *Genes & Development* 8 (12): 1434–47. https://doi.org/10.1101/gad.8.12.1434.
Vakulskas, Christopher A., and Mark A. Behlke. 2019. "Evaluation and Reduction of CRISPR Off-Target Cleavage Events." *Nucleic Acid Therapeutics* 29 (4): 167–74. https://doi.org/10.1089/nat.2019.0790.

Vakulskas, Christopher A., Daniel P. Dever, Garrett R. Rettig, Rolf Turk, Ashley M. Jacobi, Michael A. Collingwood, Nicole M. Bode, et al. 2018. "A High-Fidelity Cas9 Mutant Delivered as a Ribonucleoprotein Complex Enables Efficient Gene Editing in Human Hematopoietic Stem and Progenitor Cells." *Nature Medicine* 24 (8): 1216–24. https://doi.org/10.1038/s41591-018-0137-0.

Voelkel, Christine, Melanie Galla, Tobias Maetzig, Eva Warlich, Johannes Kuehle, Daniela Zychlinski, Juergen Bode, Tobias Cantz, Axel Schambach, and Christopher Baum. 2010. "Protein Transduction from Retroviral Gag Precursors." *Proceedings of the National Academy of Sciences* 107 (17): 7805–10. https://doi.org/10.1073/pnas.0914517107.

Walton, Russell T., Kathleen A. Christie, Madelynn N. Whittaker, and Benjamin P. Kleinstiver. 2020. "Unconstrained Genome Targeting with Near-PAMless Engineered CRISPR-Cas9 Variants." *Science* 368 (6488): 290–96. https://doi.org/10.1126/science.aba8853.

Wang, Zeng-Feng, Allen M Sirotkin, Gregory M Buchold, Arthur I Skoultchi, and William F Marzluff. 1997. "The Mouse Histone H1 Genes: Gene Organization and Differential Regulation." *Journal of Molecular Biology* 271 (1): 124–38. https://doi.org/10.1006/jmbi.1997.1166.

Williams, Michael R., Catherine J. Fricano-Kugler, Stephanie A. Getz, Patrick D. Skelton, Jeonghoon Lee, Christian P. Rizzuto, Joseph S. Geller, Meijie Li, and Bryan W. Luikart. 2016. "A Retroviral CRISPR-Cas9 System for Cellular Autism-Associated Phenotype Discovery in Developing Neurons." *Scientific Reports* 6 (1): 25611. https://doi.org/10.1038/srep25611.

Wright, Addison V., James K. Nuñez, and Jennifer A. Doudna. 2016. "Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering." *Cell* 164 (1–2): 29–44. https://doi.org/10.1016/j.cell.2015.12.035.

Wu, Dai-Tze, and Monica J. Roth. 2014. "MLV Based Viral-like-Particles for Delivery of Toxic Proteins and Nuclear Transcription Factors." *Biomaterials* 35 (29): 8416–26. https://doi.org/10.1016/j.biomaterials.2014.06.006.

Wu, Ke, Juan He, Wenchen Pu, and Yong Peng. 2018. "The Role of Exportin-5 in MicroRNA Biogenesis and Cancer." *Genomics, Proteomics & Bioinformatics* 16 (2): 120–26. https://doi.org/10.1016/j.gpb.2017.09.004.

Xia, X. G. 2003. "An Enhanced U6 Promoter for Synthesis of Short Hairpin RNA." *Nucleic Acids Research* 31 (17): 100e–1100. https://doi.org/10.1093/nar/gng098.

Xu, Christine L., Merry Z. C. Ruan, Vinit B. Mahajan, and Stephen H. Tsang. 2019. "Viral Delivery Systems for CRISPR." *Viruses* 11 (1): 28. https://doi.org/10.3390/v11010028.

Yang, Han, Shuling Ren, Siyuan Yu, Haifeng Pan, Tingdong Li, Shengxiang Ge, Jun Zhang, and Ningshao Xia. 2020. "Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks." *International Journal of Molecular Sciences* 21 (18): 6461. https://doi.org/10.3390/ijms21186461.

Yang, Yoosoo, Yeonsun Hong, Eunji Cho, Gi Beom Kim, and In-San Kim. 2018. "Extracellular Vesicles as a Platform for Membrane-Associated Therapeutic Protein Delivery." *Journal of Extracellular Vesicles* 7 (1): 1440131. https://doi.org/10.1080/20013078.2018.1440131.

Yuan, B. 1999. "Mutations Altering the Moloney Murine Leukemia Virus P12 Gag Protein Affect Virion Production and Early Events of the Virus Life Cycle." *The EMBO Journal* 18 (17): 4700–4710. https://doi.org/10.1093/emboj/18.17.4700.

Yuan, Bing, Stephen Campbell, Eran Bacharach, Alan Rein, and Stephen P. Goff. 2000. "Infectivity of Moloney Murine Leukemia Virus Defective in Late Assembly Events Is Restored by Late Assembly Domains of Other Retroviruses." *Journal of Virology* 74 (16): 7250–60. https://doi.org/10.1128/JVI.74.16.7250-7260.2000.

Zhang, Huanqing, Pei Zhuang, Ryan M. Welchko, Manhong Dai, Fan Meng, and David L. Turner. 2022. "Regulation of Retinal Amacrine Cell Generation by MiR-216b and Foxn3." *Development* 149 (2): dev199484. https://doi.org/10.1242/dev.199484.

Zhang, Jian-Ping, Xiao-Lan Li, Amanda Neises, Wanqiu Chen, Lin-Ping Hu, Guang-Zhen Ji, Jun-Yao Yu, et al. 2016. "Different Effects of SgRNA Length on CRISPR-Mediated Gene Knockout Efficiency." *Scientific Reports* 6 (1): 28566. https://doi.org/10.1038/srep28566.

Zhang, Xiya, Puping Liang, Chenhui Ding, Zhen Zhang, Jianwen Zhou, Xiaowei Xie, Rui Huang, et al. 2016. "Efficient Production of Gene-Modified Mice Using Staphylococcus Aureus Cas9." *Scientific Reports* 6 (1): 32565. https://doi.org/10.1038/srep32565.

Zhang, Yan. 2017. "The CRISPR-Cas9 System in Neisseria Spp." *Pathogens and Disease* 75 (4). https://doi.org/10.1093/femspd/ftx036.

Zhang, Yan, Nadja Heidrich, Biju Joseph Ampattu, Carl W. Gunderson, H. Steven Seifert, Christoph Schoen, Jörg Vogel, and Erik J. Sontheimer. 2013. "Processing-Independent CRISPR RNAs Limit Natural Transformation in Neisseria Meningitidis." *Molecular Cell* 50 (4): 488–503. https://doi.org/10.1016/j.molcel.2013.05.001.

Zhang, Yan, Rakhi Rajan, H. Steven Seifert, Alfonso Mondragón, and Erik J. Sontheimer. 2015. "DNase H Activity of Neisseria Meningitidis Cas9." *Molecular Cell* 60 (2): 242–55. https://doi.org/10.1016/j.molcel.2015.09.020.

Zhang, Yan, and Long-Chuan Yu. 2008. "Single-Cell Microinjection Technology in Cell Biology." *BioEssays* 30 (6): 606–10. https://doi.org/10.1002/bies.20759.

Zhi, Shengyao, Yuxi Chen, Guanglan Wu, Jinkun Wen, Jinni Wu, Qianyi Liu, Yang Li, et al. 2021. "Dual-AAV Delivering Split Prime Editor System for in Vivo Genome Editing." *Molecular Therapy*, July, S1525001621003658. https://doi.org/10.1016/j.ymthe.2021.07.011.

Zhuang, Pei, Huanqing Zhang, Ryan M. Welchko, Robert C. Thompson, Shunbin Xu, and David L. Turner. 2020. "Combined MicroRNA and MRNA Detection in Mammalian Retinas by in Situ Hybridization Chain Reaction." *Scientific Reports* 10 (1): 351. https://doi.org/10.1038/s41598-019-57194-0.