Title
Engineering Adeno-Associated Viral Vectors for Gene Therapy, Gene Targeting, and Gene Manipulation

Author
Sun, Sabrina

Publication Date
2019

Peer reviewed|Thesis/dissertation

Engineering Adeno-Associated Viral Vectors for Gene Therapy, Gene Targeting, and Gene Manipulation

By

Sabrina I Sun

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Chemical Engineering
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor David V. Schaffer, Chair
Professor Markita P. Landry
Professor Britt A. Glaunsinger

Summer 2019

Abstract

Engineering Adeno-Associated Viral Vectors
for Gene Therapy, Gene Targeting, and Gene Manipulation

by

Sabrina I Sun

Doctor of Philosophy in Chemical Engineering

University of California, Berkeley

Professor David V. Schaffer, Chair

Gene therapy, the delivery of therapeutic genetic material to diseased cells, has demonstrated clinical efficacy for numerous inherited disorders in the last decade. Recombinant adeno-associated viral (AAV) vectors, in particular, have enabled the development of breakthrough drugs for several rare genetic diseases, including biallelic *RPE65*-associated retinal dystrophy, spinal muscular atrophy Type I, and hemophilia B, due to their ability to safely mediate long-term transgene expression in human patients. While this clinical success has inspired widespread enthusiasm about the promise of AAV-mediated gene delivery to treat a variety of diseases, broad applicability of AAV vectors for gene therapy demands enhanced transduction efficiency and targeting of therapeutically relevant tissues. Innovative strategies that grant greater functional benefit, in conjunction with improved disease models, are also pivotal for realizing the full potential of AAV-based therapies in treating human disorders. Such challenges have driven my dissertation research in two major directions: engineering new AAV capsid variants to address unmet clinical needs, and combining AAV vectors and CRISPR/Cas9 technologies to efficiently generate preclinical models of human disorders and to correct disease-causing mutations through genome editing.

The AAV delivery problem has been tackled by engineering the viral capsid to evade immune detection, circumvent biological barriers, and exhibit high infectivity of clinically important tissues. In this dissertation, I describe the latest trends in AAV vector design that harness cell-type-specific selection schemes and our rapidly accumulating knowledge of AAV capsid structure and function to overcome longstanding obstacles in the gene therapy field. Of note, these techniques have yielded novel AAV variants with desirable biodistribution in the central nervous system, stimulating strong interest in adapting them for sparse neuronal labeling and physiological studies of the brain and spinal cord. I also discuss next-generation sequencing platforms that have powered unprecedented, high-throughput phenotypic profiling of AAV vectors and lent invaluable insight to AAV capsid engineering.

In addition to the development of more potent vectors, the efficient generation of preclinical animal models is key to continued progress in gene therapy. Specifically, genetically engineered mouse models of disease enable early assessment of safety and efficacy of new therapeutics. These mouse lines harbor large sequence insertions or modifications and thus require

complex genomic manipulation; however, existing methods to produce such animals remain laborious and costly. To address this, we developed an approach called CRISPR-READI (CRISPR RNP Electroporation and AAV Donor Infection) that couples AAV-mediated donor delivery with Cas9/sgRNA ribonucleoprotein (RNP) electroporation to induce large homology-directed, site-specific modifications in the mouse genome with high efficiency and throughput. We successfully targeted a 774 bp fluorescent reporter, a 2.1 kb CreERT2 driver, and a 3.3 kb expression cassette into endogenous loci in both embryos and live mice, demonstrating that CRISPR-READI is applicable to most widely used knock-in schemes with potential applications in other mammalian species.

AAV vectors and CRISPR-based tools can also be leveraged for permanent correction of debilitating genetic diseases through genome editing in somatic tissues. Duchenne muscular dystrophy (DMD) is a particularly attractive disease target, as the massive *DMD* gene limits the feasibility of simple gene addition as a therapeutic approach. To date, most gene editing strategies for DMD employ reframing or exon skipping mediated by non-homologous end joining (NHEJ) to bypass mutations that disrupt the *DMD* open reading frame and produce a truncated but partially functional dystrophin protein. However, restoration of full-length dystrophin expression has not yet been achieved. To overcome this challenge, we devised a novel "excise-and-replace" approach that harnesses AAV-mediated co-delivery of CRISPR/Cas9 and a homology-independent targeted integration (HITI) donor to correct the *DMD* locus. Through dual gRNA directed deletion and repair of a target exon, our strategy yields wildtype dystrophin transcripts in an *in vitro* human cardiomyocyte model of DMD. To improve gene targeting efficiency, we also examine the ability of a third gRNA to preferentially drive desired transgene insertion over other genomic arrangements. Our gene correction method can potentially be applied to various exon deletions and small mutations in *DMD* to rescue full-length dystrophin expression and confer maximal clinical benefit in afflicted muscle cells.

In summary, my dissertation research focuses on the ongoing evolution of new AAV capsids and the creation of advanced genome editing strategies for generating preclinical animal models and treating human disease. Together, these technologies make important contributions to the development of improved AAV-mediated gene therapies.

Dedication

In loving memory of my father

Kochiu (Joe) Sun

6/3/1957-6/29/2018

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my adviser, David Schaffer, for inviting me to be a part of the reinvigoration of AAV-mediated gene therapy and the revolution of gene editing. The expertise that I have gained as a result of being immersed in these two exciting fields for the last six years has positioned me exceptionally well for the next step of my career. This was all made possible when you took a chance on me after just a handful of meetings in the fall of 2013. Since then, your positive reinforcement and support for my interest in opportunities outside of the lab have enabled me to grow as a scientist and to identify future directions for my career development. Finally, thank you for offering me multiple opportunities to work with other academic labs; these were the most eye-opening and rewarding experiences of my PhD and helped me recognize my potential in collaborative environments.

I would also like to thank my other two thesis committee members, Britt Glaunsinger and Markita del Carpio Landry, for their feedback and time. You are both forces to be reckoned with and are influential role models for young women in STEM by being so sharp and yet so kind. Thank you for the inspiration to pursue leadership roles in science and mentor the next generation of female scientists.

It is simply impossible to fully capture my gratitude for our lab manager, Noem Ramey, in a few words. In the six years that I've been in the lab, you have never neglected a scientific request of mine, no matter how overwhelmed you were with imminent deadlines and long to-do lists. More importantly, your unfaltering dedication to build a culture of warmth and kindness in the lab has been pivotal in countering the rocky nature of research and has fostered a strong sense of camaraderie within the group. On a personal level, you have re-directed me when I have wandered off the path and supported me wholeheartedly when I suffered the most painful loss of my life. Thank you for your invaluable perspective, patience, empathy, and love throughout my PhD.

I am extremely thankful to Jorge Santiago-Ortiz and David Ojala, who of their own volition took me under their wing when I was a naïve first-year grad student floundering in the lab. From both of you, I learned the basics of molecular biology, AAV vector production, and everything in between. Our many scientific discussions collectively catalyzed the turning point of my graduate career, at which I could think critically about problems on my own. I want to also thank you for lifting my spirits in moments of weakness and literally rushing to my side in times of frustration. I still cannot believe that you both came down to the NAF to help me with tail vein injections when I was convinced my experiment would fail – your generosity is unparalleled. I feel so fortunate that we had and continue to have each other to lean on.

Throughout my PhD, I had rich learning experiences working with collaborators from within and outside the lab on a range of interesting multidisciplinary projects. I want to thank David Booth, Heather Middleton, Lily Helfrich, and Nicole King for welcoming me into their lab as though I were one of their own, providing an open environment that encourages discussion, and always challenging me to think deeper. I also thoroughly enjoyed working side-by-side with Sean Chen, whose ambition and drive are awe-inspiring, and Lin He, whose advice was always insightful. We made such an efficient and productive team, and I'm so proud of the technology that we developed together. Finally, thanks to Xiaoping Bao, whose creativity and vision have

been crucial for bringing a challenging project to fruition. Our complementary skillsets created a great partnership, and I continue to learn from your wealth of knowledge.

I have so many labmates to thank for their scientific guidance and friendship. Prajit Limsirichai was always ready to offer a helping hand, along with humorous "pep talks". You are by far the most resourceful person that I know, and we were all so lucky to have benefited from your selfless generosity. Anusuya Ramasubramanian was a wonderful role model who made time for me even on her way out of the lab and continues to be supportive while living on the opposite coast. I have engaged in many stimulating discussions with Tim Day, Cecile Fortuny, and Leah Byrne, who were our gene therapy counterparts in John Flannery's group, and appreciated the lovely dynamic that we shared. Thom Gaj provided invaluable mentorship to me and so many others by bringing us up to speed on genome editing and pushing the lab to make novel contributions at the forefront of an exciting and rapidly evolving field. Rocio Sampayo and Kira Mosher have repeatedly advised me with disarming patience and positivity, and it is no surprise to me that they have risen as valued mentors in the lab. I want to also thank my former labmates – Dawn Spelke, Maroof Adil, Barbara Ekerdt, Alyssa Rosenbloom, Sisi Chen, and Lukasz Bugaj – for their encouragement and perspective, especially during the early years of my PhD. Finally, thanks to colleagues from inside and outside of the lab – Goncalo Rodrigues, Sean McFarland, Yusuke Okubo, Mike Kang, Marc Casas, George Lin, and Joe Chen – for their help and insight.

I want to give a special shout-out to my fellow crusty ol' 6th year grad students in the lab: Phil Kang, Nicole Repina, Riya Muckom, Andrew Bremer, and Olivia Scheideler (and honorary member, Christina Fuentes). My heart is full, as I reflect on moments when we felt lost and helped each other find the path forward again. I'm so grateful that our bond remained steadfast through the ups and downs of grad school, despite how large our class was. Andrew, you inspire me constantly with your passion for science policy and I can't wait for the day that I address you as Congressman Bremer. Thank you for being willing to help at any given moment and for blessing me with your contagious laughter. I will forever think back to all the raw conversations that we shared with bittersweet nostalgia. Olivia and Christina, suffice it to say that separate cards are needed to fully capture my gratitude for your unwavering support and (best) friendship. But in brief, I am incredibly thankful that I can always count on you two to be great listeners and personal cheerleaders. While I will sorely miss seeing you both every day, I find comfort in knowing that we have developed a special and lasting sisterhood that will keep us connected as we tumble through post-grad life and on.

The work described in this dissertation would not be possible without the help of my undergraduate assistants. I am so glad that I recognized the potential of Clancy Lee while she was a work study student in our lab and helped her transition into research. You were absolutely exceptional, and your attention to detail and ability to pick up new skills are impressive. What a luxury it was that I could show you a technique once and trust that you could apply it independently and accurately thereafter. My most productive time in grad school was when we worked together, so thank you for your diligence, resilience, and positivity. I also had the pleasure of mentoring Aaron Espinosa, who helped me throughout the past year with great enthusiasm and cheer. I truly admired and appreciated your work ethic and receptiveness to constructive criticism.

Thank you to the faculty of the Department of Chemical and Biomolecular Engineering who challenged me to step beyond my comfort zone and the staff who make all student affairs possible. Carlet Altamirano has patiently resolved every crisis that I have brought to her attention and somehow manages much more than what a single person should be asked to handle with

admirable composure. To my friends in the class of 2013 – Michelle Liu, Nahyun Cho, Meron Tesfaye, Danielle Pesko, Ksenia Timachova, Audrey Harker, and Alice Yeh – I could not have endured the stress of prelims and the turbulence of transitioning into research without your support.

Last but certainly not least, I want to thank David Ojala and my family for their unconditional love and support through the last six years. David, thank you for helping me troubleshoot my research over dinner, for providing perspective when I needed it the most, and for your compassion and understanding during my darkest moments. I am also so grateful to my parents, who raised me to be self-driven yet selfless and to pursue lofty goals. My father in particular, to whom this dissertation is dedicated, consistently met my self-doubt with reassurance and exuded pride in my every achievement. I would have given up early on in my PhD without his encouragement. I deeply miss your comforting and loving presence, but I know you would have been exceptionally proud, Daddy.

Chapter 1: Introduction

1.1 AAV Vectors for Gene Therapy

Gene therapy is the delivery of a therapeutic gene cassette to treat a disease or deficiency[1]. This approach is especially well-suited for treating monogenic disorders, or diseases that are caused by a mutation in a single gene, due to the known functional benefit of providing a correct copy of the gene[2, 3]. Importantly, efficient transgene delivery to target cells and tissues is necessary to induce functional improvement and accurately assess the safety and efficacy of the treatment[4]. While non-viral vehicles can offer some attractive safety features and potentially carry larger cargo[5], millions of years of evolution have refined the ability of natural viruses to circumvent biological transport barriers and transfer their genetic payload to cells[5].

Recombinant adeno-associated viral (AAV) vectors, in particular, have irrefutably taken the lead in gene therapy clinical trials, gaining overwhelming popularity in the past decade. The FDA's approval of Luxturna in 2017 for a rare genetic form of retinal dystrophy and the wave of Phase 1/2 studies coming down the clinical pipeline are compelling testaments to the therapeutic promise of this breakthrough technology. Its minimal risk of insertional mutagenesis, low immunogenicity, and ability to safely mediate long-term transgene expression in both dividing and non-dividing cells all contribute to its strong potential for single-administration treatments[6]. Nonetheless, improvements in delivery efficiency, target specificity, and resistance to neutralizing antibodies of AAV vectors[7] in humans are crucial for expanding patient eligibility and adapting the technology to a broad range of diseases. These shortcomings have driven and continue to fuel the engineering of next-generation AAV vectors that surmount delivery challenges barring clinical translation, bringing AAV-based gene therapies closer to fruition.

AAV Biology

Adeno-associated virus is a non-enveloped single-stranded DNA virus that requires the presence of a helper virus to replicate[8]. Flanked by inverted terminal repeats (ITRs) that serve as the viral origin of replication, two viral genes *rep* and *cap* occupy the 4.7-kilobase (kb) packaging capacity of the AAV genome and encode multiple proteins through alternative open reading frames (ORFs) and splicing[8]. The *rep* open reading frame (ORF) produces four non-structural proteins that are critical for viral replication, targeted integration of natural AAV genomes into the host genome, transcriptional regulation of promoter activity, and viral particle assembly[8, 9]. The *cap* ORF encodes three structural proteins (VP1-3), present in a 1:1:10 ratio, that assemble into a 60-mer icosahedral capsid with a diameter of approximately 20 nm[8]. Finally, an alternative ORF within *cap* generates the assembly-activating protein (AAP), which recruits newly synthesized VPs to the nucleolus of the host cell and facilitates capsid assembly[10]. Wide dissemination of native AAV genomes in human and non-human primate tissues has led to the discovery of over 100 unique isolates, or serotypes, and implicated the utility of the non-pathogenic virus in human gene therapy[11]. These natural variants differ in tropism, or the ability to transduce specific types of cells and tissues, and other infectious properties because of the distinct amino acid composition of their capsids[12].

Due to their favorable safety and transduction profiles, AAVs have been repurposed as recombinant vectors for delivery of therapeutic transgenes to diseased cell populations. As the ITRs are the only *cis* elements necessary for viral packaging, the native AAV genome can be gutted and replaced with a customized expression cassette, while *rep* and *cap* are supplied *in trans* during recombinant AAV (rAAV) vector production. These vectors primarily persist as stable episomal DNA structures upon infection of post-mitotic cells and mediate transgene expression that can remarkably last for years[13]. The identification of AAV serotypes with high infectivity of specific target tissues has enabled therapeutic success in AAV-based preclinical and clinical studies for disorders afflicting the retina, blood, liver, central nervous system, skeletal muscle, and more[9].

Directed Evolution of the AAV Capsid

In spite of several clinical milestones achieved in AAV-mediated gene therapy, viral uptake and transgene expression in many target cells and tissues must be improved to unlock the full therapeutic potential of recombinant AAV technology. The obstacles of inadequate localization to certain cell types, as well as widespread dispersion to off-target tissues, stem from the simple consideration that properties conferring fitness for natural viral infections are distinct from those needed for treating human disease. Furthermore, pre-existing anti-AAV immunity necessitates vectors with antibody evading capabilities to treat patients who would otherwise be excluded. AAV's minimalistic genome, in which all capsid proteins are encoded by a single gene, has enabled accelerated development of novel capsid design strategies, as the virus's tropism and other infectious properties can be manipulated through modification of *cap* alone[9]. AAV capsid engineering has two main flavors: rational design and directed evolution. While our knowledge of AAV capsid structure continues to grow, dissecting the intricate structure-function relationships of multimeric viral particles remains challenging[9]. In contrast, directed evolution does not require pre-existing structural or mechanistic understanding of virus-cell interactions and has been extensively and successfully employed for the generation of enhanced AAV vectors.

Directed evolution is a powerful, high-throughput tool that was initially devised to engineer antibodies with enhanced binding affinity and enzymes with improved catalytic activity[14]. This technique has since been applied to engineer novel AAV variants with altered receptor binding, antibody-evading capabilities, improved *in vitro* and *in vivo* infectivity of previously non-permissive cells, and enhanced retrograde transport efficiency in neurons *in vivo*[15-20]. Like its natural counterpart, directed evolution entails genetic diversification and iterative rounds of phenotypic selection to drive convergence toward the fittest variants (Figure 1). The paradigm begins with generating large *cap* libraries (>10^6 in genetic diversity) using a range of molecular biology methods including error-prone polymerase chain reaction (PCR)[21], saturation mutagenesis of key regions[18], peptide insertion[22], DNA shuffling[17], structure-guided recombination[23, 24], transposon-based insertional mutagenesis[25], and ancestral reconstruction[26]. These mutant libraries are assembled into viral particles, in which each viral capsid encapsulates the genome encoding it[19,27]. The virions are then subjected to a functional selective pressure, and *cap* variants conferring therapeutically relevant properties are recovered from each round. After a suitable number of iterations, individual AAV variants are isolated, validated, and harnessed for therapeutic gene delivery in disease models. Directed evolution efforts have culminated in recent FDA clearance of

several synthetic AAVs for use in clinical studies, reinforcing the utility of this method for improving vector performance.

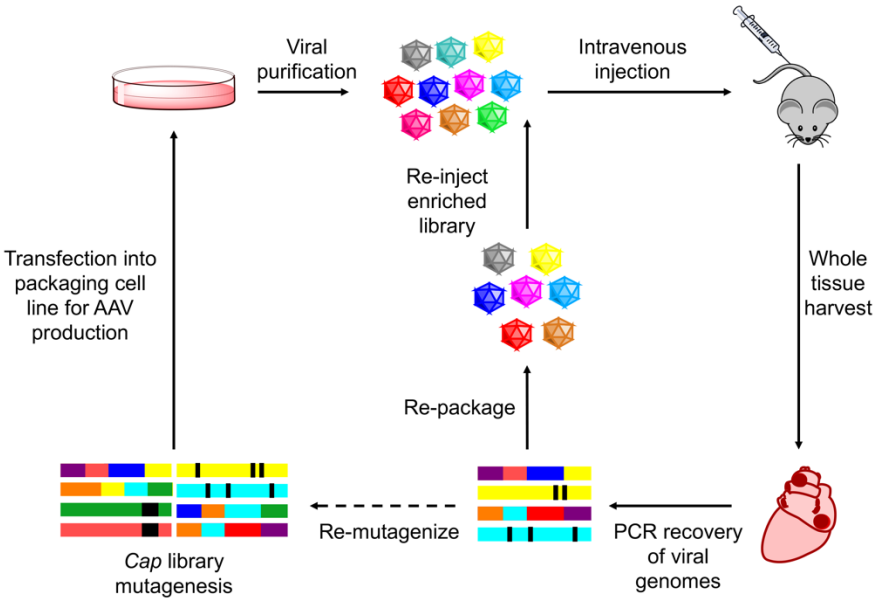


Figure 1.1. AAV capsid engineering using directed evolution.

Application of directed evolution to the AAV capsid begins with genetic diversification of the AAV *cap* gene using a variety of molecular biology techniques. The mutant *cap* library is co-transfected with helper plasmids into a packaging cell line for viral particle assembly. Virions are then purified and subjected to a user-defined selective pressure. In this example, the AAV library is injected by systemic administration into mice and viral genomes are harvested by PCR amplification from the target tissue. This pool of *cap* genes is subsequently re-packaged into viral particles, and iterative rounds of selection are conducted to enrich for the top performing variants. Alternatively, recovered clones can undergo further mutagenesis prior to the next selection round to introduce greater genetic diversity.

1.2 AAV-Mediated Delivery of Gene Editing Tools

While long-term expression of therapeutic transgenes from AAV genomes has indeed demonstrated efficacy for numerous rare monogenic disorders, the limited capacity and episomal nature of AAV vectors have stimulated interest in leveraging gene editing technology to modify endogenous loci. Therapeutic gene editing can potentially treat a wider range of diseases, including those with target genes that exceed AAV's native packaging capacity (4.9 kb) and those affecting mitotic cell types in which AAV genomes undergo degradation and dilution[28]. The approach is further bolstered by the prospect of permanent gene correction, as the persistence of AAV episomes in transduced cells continues to be studied[13, 29]. Gene editing has also been employed for the generation of knock-in mice for modeling human disease and assessing the safety and efficacy of gene therapies[30]. In particular, knock-in mice harboring large sequence insertions or substitutions are essential for a variety of applications including endogenous gene tagging, conditional gene knockout, site-specific transgene insertion, and gene replacement.

Gene editing has classically relied on homologous recombination (HR), a DNA repair mechanism that can be exploited to introduce mutations into a genomic locus by supplying a donor substrate flanked by sequences with homology to the target region. In spite of its accuracy, HR is limited in efficiency, typically achieving gene targeting frequencies of 10^{-6} or lower[31]. To overcome this obstacle, endogenous DNA repair mechanisms have been leveraged through artificial, nuclease-mediated induction of double strand breaks (DSBs)[32, 33]. As DSBs are hazardous to the cell, this molecular event triggers repair through two major pathways: 1) non-homologous end joining (NHEJ) in which the two break ends are directly ligated and error-prone repair can result in heterogenous small insertions or deletions, also known as "indels", that often disrupt the open reading frame (ORF) of the target gene, and 2) homology-directed repair (HDR) in which an exogenous sequence flanked by "arms" homologous to the target region is used as a donor template to precisely repair the genomic lesion[34] (Figure 2). The emergence and refinement of nucleases with programmable sequence specificity, including zinc finger nucleases (ZFNs)[35] and transcription-activator-like effector nucleases (TALENs)[36], has enabled the introduction of site-specific DSBs to activate NHEJ- or HDR-based repair and improve the efficiency of gene knockout and knock-in by two to three orders of magnitude in mammalian cells[37]. The target specificity of these technologies is indeed remarkable[38, 39]; however, nucleotide triplet recognition by zinc finger proteins (ZFPs) requires extensive engineering of ZFNs to target a desired genomic sequence, while the highly repetitive nature and large size of TALEs present molecular cloning and delivery challenges[40].

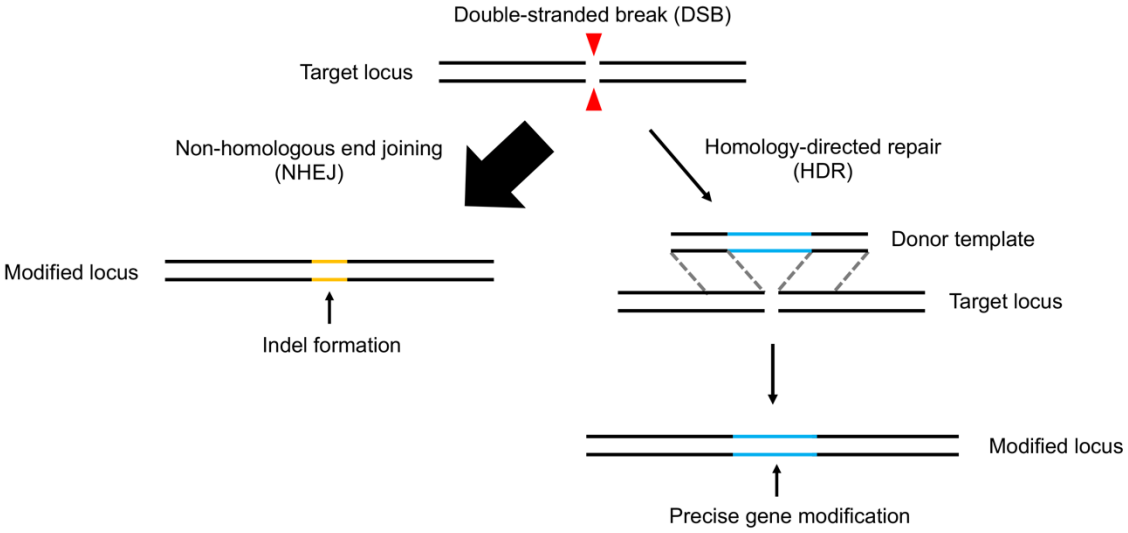


Figure 1.2. Endogenous cell repair pathways in mammalian cells.

Double-stranded breaks (DSBs) in genomic DNA are repaired through two major endogenous pathways in mammalian cells: non-homologous end joining (NHEJ) and homology-directed repair (HDR). During NHEJ (left), the two ends of the lesion are directly ligated. However, error-prone repair can lead to indel formation at the DSB, often resulting in a frameshift that disrupts the open reading frame of the target gene. During HDR (right), an exogenous donor template with homology to the target locus serves as a repair substrate to facilitate precise gene modification. NHEJ occurs more efficiently than HDR and is active in both proliferating and post-mitotic cells. In contrast, HDR relies on cell repair machinery present in S and G2 phases of cell cycle and thus has limited potential for gene correction and replacement in dividing cells.

CRISPR/Cas9-Based Genome Engineering

Since the discovery of their role in bacterial adaptive immunity, Class 2 Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) systems have revolutionized genome engineering in mammalian cells as a result of their unparalleled simplicity and programmability. The technical ease of adapting CRISPR/Cas9 technology to target any genomic locus is largely due to its unique modularity. Unlike previously discussed enzymes that utilize protein motifs for sequence-specific DNA recognition and binding, CRISPR-based editing tools are comprised of two major components: 1) a Cas9 endonuclease with the ability to cleave double stranded DNA, and 2) a chimeric single guide RNA (sgRNA) containing a 20 nucleotide "protospacer" sequence that is complementary to the genomic target site and a scaffold sequence necessary for Cas9 binding[41]. Cas9 activity also requires the presence of a protospacer adjacent motif (PAM) immediately adjacent to the genomic target site[41]. Notably, modification of the gRNA protospacer alone re-targets Cas9 to a different genomic sequence[42-44]. As target specificity is programmed by the gRNA, simultaneous engineering of multiple genomic loci can potentially be achieved through multiplex expression of gRNAs[45]. Furthermore, the DNA binding and nuclease domains of Cas9 function as independent modules[41], facilitating the development of nucleolytically inactive, "dead" Cas9 (dCas9) mutants that can be fused to transcriptional regulators and targeted to promoter regions for robust activation[46-51] or repression[52] of endogenous genes. Chromatin modifiers[53-57] or cytidine and adenine deaminases[58-60] can also be genetically conjugated to dCas9 to modulate gene expression or mediate precise single nucleotide conversion, respectively, without a DSB at the genomic target site. Finally, the engineering of enhanced Cas9-based enzymes, including paired nickases[61] and high fidelity nucleases that confer higher target specificity[62-66] and engineered nuclease variants with non-native PAM recognition that further broaden the technology's applicability[67-70], continues at an accelerated pace.

The attractive complementarity of AAV-mediated gene delivery and CRISPR/Cas9 editing tools has driven the combination of the two technologies for gene correction and gene targeting. While *S. pyogenes* Cas9 (SpCas9) remains the most popular CRISPR endonuclease for genome engineering due to its high on-target activity and the abundance of its simple PAM sequence[71], its large size (~4 kb) necessitates the use of minimal but suboptimal promoters and polyadenylation signals to efficiently produce AAV vectors expressing SpCas9[72, 73]. To bypass AAV's packaging limit, split-intein moieties have been incorporated into each of the SpCas9 halves and shown to induce *trans*-splicing-mediated reconstitution of the full-length SpCas9 protein[74-76]. Furthermore, the discovery of a more compact Cas9 ortholog derived from *S. aureus* (SaCas9) has propelled the exploration of *in vivo* genome editing and regulation using AAV vectors and CRISPR-based tools[77-81]. Recently, pre-existing adaptive immunity against SpCas9 and SaCas9 was found to be highly prevalent in human sera[82], raising important questions about the safety of sustained Cas9 expression mediated by AAV and urging the identification of small orthologs from bacterial species that do not infect humans.

AAV Donor Vectors

Not only can AAV vectors be employed to deliver gene editing tools into target tissues, they can also serve as donor templates for DNA repair. In fact, AAV vectors alone can stimulate gene targeting with efficiencies 1000-fold greater than plasmid and other viral vector systems[31], achieving up to 1% knock-in frequencies in treated cells[83]. Moreover, the presence of a nuclease-

induced DSB can boost AAV-mediated gene targeting by up to 100-fold in human cells[84, 85]. Due to their recombinogenic nature and broad tropism, AAV vectors have been paired with engineered site-specific nucleases to produce homology-directed genome modifications in mammalian cells for a variety of applications, including stable cell line generation *in vitro*[86] and therapeutic gene correction *ex vivo*[87] and *in vivo*[88]. In this approach, the exogenous repair template is provided by the AAV donor vector, and the targeted nuclease technology is also vectorized[89], or alternatively delivered as mRNA[90-92] or ribonucleoproteins (RNPs)[93] that exhibit rapid cleavage activity and turnover. AAV-based delivery of CRISPR/Cas9 components and HDR donors has also been found to facilitate efficient targeted gene insertion in mouse embryos *ex vivo*, demonstrating the potential of AAV vectors to generate mouse models with far superior throughput than conventional microinjection[94]. Finally, a recent study showed that that the constraint on donor length imposed by AAV's small packaging capacity can be bypassed by splitting the transgene into two AAV donor vectors that induce sequential recombination events to reconstitute the full-length cassette, thus further expanding the editing capabilities of AAV-mediated HDR[95].

1.3 Scope of Dissertation

This dissertation centers on the development of improved gene therapies, with a particular focus on the intersection of two breakthrough technologies: AAV vectors and CRISPR/Cas9 editing tools. While AAV-mediated delivery of therapeutic transgenes has already reached important clinical milestones for several genetic disorders, the applicability of AAV-based gene therapies to a broad range of diseases demands advancements on the following major fronts: 1) efficient delivery to target cells and tissues, 2) generation of disease models for preclinical characterization of safety and efficacy, and 3) the design of novel strategies that achieve enhanced clinical benefit. The work presented here advances the field on all three fronts and in turn illustrates the potency of unifying AAV and CRISPR/Cas9 technologies. **Chapter 1** reviews the latest innovations in AAV vector engineering that aim to overcome the specific delivery challenges thwarting clinical translation of AAV-based gene therapies. Emerging vector characterization methods and barriers to product development are also discussed to provide a high-level perspective on the promise of commercializing novel AAV variants. **Chapter 2** dives into the application of synthetic AAVs to the central nervous system and the new opportunities that they offer for characterizing the identity and connectivity of single neurons within complex neural networks and modulating neuronal activity via delivery of functional genetic tools. **Chapter 3** reports the development of a novel approach called CRISPR-READI that combines AAV-mediated HDR donor delivery with Cas9/gRNA RNP electroporation to engineer large site-specific modifications in the mouse genome with high efficiency and throughput. Finally, **Chapter 4** discusses the design and application of new AAV-based "excise-and-replace" repair strategies for correction of the mutant dystrophin locus to restore full-length dystrophin protein expression in a human cardiomyocyte model of Duchenne muscular dystrophy.

1.4 References

1. Schaffer, D.V., J.T. Koerber, and K.I. Lim, *Molecular engineering of viral gene delivery vehicles.* Annu Rev Biomed Eng, 2008. **10**: p. 169-94.
2. Russell, S., et al., *Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial.* Lancet, 2017. **390**(10097): p. 849-860.
3. Al-Zaidy, S., et al., *Health outcomes in spinal muscular atrophy type 1 following AVXS-101 gene replacement therapy.* Pediatr Pulmonol, 2019. **54**(2): p. 179-185.
4. Hulot, J.S., K. Ishikawa, and R.J. Hajjar, *Gene therapy for the treatment of heart failure: promise postponed.* Eur Heart J, 2016. **37**(21): p. 1651-8.
5. Yin, H., et al., *Non-viral vectors for gene-based therapy.* Nat Rev Genet, 2014. **15**(8): p. 541-55.
6. McCarty, D.M., S.M. Young, Jr., and R.J. Samulski, *Integration of adeno-associated virus (AAV) and recombinant AAV vectors.* Annu Rev Genet, 2004. **38**: p. 819-45.
7. Colella, P., G. Ronzitti, and F. Mingozzi, *Emerging Issues in AAV-Mediated In Vivo Gene Therapy.* Mol Ther Methods Clin Dev, 2018. **8**: p. 87-104.
8. Weitzman, M.D. and R.M. Linden, *Adeno-associated virus biology.* Methods Mol Biol, 2011. **807**: p. 1-23.
9. Kotterman, M.A. and D.V. Schaffer, *Engineering adeno-associated viruses for clinical gene therapy.* Nat Rev Genet, 2014. **15**(7): p. 445-51.
10. Naumer, M., et al., *Properties of the adeno-associated virus assembly-activating protein.* J Virol, 2012. **86**(23): p. 13038-48.
11. Gao, G., et al., *Clades of Adeno-associated viruses are widely disseminated in human tissues.* J Virol, 2004. **78**(12): p. 6381-8.
12. Zincarelli, C., et al., *Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection.* Mol Ther, 2008. **16**(6): p. 1073-80.
13. Buchlis G, e.a., *Factor IX expression in skeletal muscle of a severe hemophilia B patient 10 years after AAV-mediated gene transfer.* Blood, 2012. **119**(13): p. 3038-3041.
14. Romero, P.A. and F.H. Arnold, *Exploring protein fitness landscapes by directed evolution.* Nat Rev Mol Cell Biol, 2009. **10**(12): p. 866-76.
15. Asuri, P., et al., *Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells.* Mol Ther, 2012. **20**(2): p. 329-38.
16. Jang, J.H., et al., *An evolved adeno-associated viral variant enhances gene delivery and gene targeting in neural stem cells.* Mol Ther, 2011. **19**(4): p. 667-75.
17. Koerber, J.T., J.H. Jang, and D.V. Schaffer, *DNA shuffling of adeno-associated virus yields functionally diverse viral progeny.* Mol Ther, 2008. **16**(10): p. 1703-9.

18. Koerber, J.T., et al., *Molecular evolution of adeno-associated virus for enhanced glial gene delivery.* Mol Ther, 2009. **17**(12): p. 2088-95.

19. Maheshri, N., et al., *Directed evolution of adeno-associated virus yields enhanced gene delivery vectors.* Nat Biotechnol, 2006. **24**(2): p. 198-204.

20. Tervo, D.G., et al., *A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons.* Neuron, 2016. **92**(2): p. 372-382.

21. Koerber, J.T., et al., *Construction of diverse adeno-associated viral libraries for directed evolution of enhanced gene delivery vehicles.* Nat Protoc, 2006. **1**(2): p. 701-6.

22. Muller, O.J., et al., *Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors.* Nat Biotechnol, 2003. **21**(9): p. 1040-6.

23. Silberg, J.J., J.B. Endelman, and F.H. Arnold, *SCHEMA-guided protein recombination.* Methods Enzymol, 2004. **388**: p. 35-42.

24. Ojala, D.S., et al., *In Vivo Selection of a Computationally Designed SCHEMA AAV Library Yields a Novel Variant for Infection of Adult Neural Stem Cells in the SVZ.* Mol Ther, 2017.

25. Koerber, J.T. and D.V. Schaffer, *Transposon-based mutagenesis generates diverse adeno-associated viral libraries with novel gene delivery properties.* Methods Mol Biol, 2008. **434**: p. 161-70.

26. Santiago-Ortiz, J., et al., *AAV ancestral reconstruction library enables selection of broadly infectious viral variants.* Gene Ther, 2015. **22**(12): p. 934-46.

27. Nonnenmacher, M., et al., *High capsid-genome correlation facilitates creation of AAV libraries for directed evolution.* Mol Ther, 2015. **23**(4): p. 675-82.

28. Nakai, H., et al., *Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction in vivo.* J Virol, 2001. **75**(15): p. 6969-76.

29. Nathwani, A.C., et al., *Long-term safety and efficacy following systemic administration of a self-complementary AAV vector encoding human FIX pseudotyped with serotype 5 and 8 capsid proteins.* Mol Ther, 2011. **19**(5): p. 876-85.

30. Gurumurthy, C.B. and K.C.K. Lloyd, *Generating mouse models for biomedical research: technological advances.* Dis Model Mech, 2019. **12**(1).

31. Russell, D.W. and R.K. Hirata, *Human gene targeting by viral vectors.* Nat Genet, 1998. **18**(4): p. 325-30.

32. Rouet, P., F. Smih, and M. Jasin, *Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells.* Proc Natl Acad Sci U S A, 1994. **91**(13): p. 6064-8.

33. Porteus, M.H. and D. Baltimore, *Chimeric nucleases stimulate gene targeting in human cells.* Science, 2003. **300**(5620): p. 763.

34. Wyman, C. and R. Kanaar, *DNA double-strand break repair: all's well that ends well.* Annu Rev Genet, 2006. **40**: p. 363-83.

35. Urnov, F.D., et al., *Genome editing with engineered zinc finger nucleases.* Nat Rev Genet, 2010. **11**(9): p. 636-46.

36. Joung, J.K. and J.D. Sander, *TALENs: a widely applicable technology for targeted genome editing.* Nat Rev Mol Cell Biol, 2013. **14**(1): p. 49-55.

37. Rouet, P., F. Smih, and M. Jasin, *Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease.* Mol Cell Biol, 1994. **14**(12): p. 8096-106.

38. Miller, J.C., et al., *An improved zinc-finger nuclease architecture for highly specific genome editing.* Nat Biotechnol, 2007. **25**(7): p. 778-85.

39. Kim, Y., et al., *A library of TAL effector nucleases spanning the human genome.* Nat Biotechnol, 2013. **31**(3): p. 251-8.

40. Gupta, R.M. and K. Musunuru, *Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9.* J Clin Invest, 2014. **124**(10): p. 4154-61.

41. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.

42. Jinek, M., et al., *RNA-programmed genome editing in human cells.* Elife, 2013. **2**: p. e00471.

43. Mali, P., et al., *RNA-guided human genome engineering via Cas9.* Science, 2013. **339**(6121): p. 823-6.

44. Cho, S.W., et al., *Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease.* Nat Biotechnol, 2013. **31**(3): p. 230-2.

45. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.

46. Perez-Pinera, P., et al., *RNA-guided gene activation by CRISPR-Cas9-based transcription factors.* Nat Methods, 2013. **10**(10): p. 973-6.

47. Maeder, M.L., et al., *CRISPR RNA-guided activation of endogenous human genes.* Nat Methods, 2013. **10**(10): p. 977-9.

48. Mali, P., et al., *CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering.* Nat Biotechnol, 2013. **31**(9): p. 833-8.

49. Chavez, A., et al., *Highly efficient Cas9-mediated transcriptional programming.* Nat Methods, 2015. **12**(4): p. 326-8.

50. Gilbert, L.A., et al., *Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation.* Cell, 2014. **159**(3): p. 647-61.

51. Konermann, S., et al., *Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex.* Nature, 2015. **517**(7536): p. 583-8.

52. Gilbert, L.A., et al., *CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes.* Cell, 2013. **154**(2): p. 442-51.

53. Hilton, I.B., et al., *Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers.* Nat Biotechnol, 2015. **33**(5): p. 510-7.

54. Thakore, P.I., et al., *Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements.* Nat Methods, 2015. **12**(12): p. 1143-9.

55. Kearns, N.A., et al., *Functional annotation of native enhancers with a Cas9-histone demethylase fusion.* Nat Methods, 2015. **12**(5): p. 401-403.

56. Cano-Rodriguez, D., et al., *Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner.* Nat Commun, 2016. **7**: p. 12284.

57. Kwon, D.Y., et al., *Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC.* Nat Commun, 2017. **8**: p. 15315.

58. Komor, A.C., et al., *Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage.* Nature, 2016. **533**(7603): p. 420-4.

59. Rees, H.A., et al., *Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery.* Nat Commun, 2017. **8**: p. 15790.

60. Gaudelli, N.M., et al., *Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage.* Nature, 2017. **551**(7681): p. 464-471.

61. Ran, F.A., et al., *Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity.* Cell, 2013. **154**(6): p. 1380-9.

62. Slaymaker, I.M., et al., *Rationally engineered Cas9 nucleases with improved specificity.* Science, 2016. **351**(6268): p. 84-8.

63. Kleinstiver, B.P., et al., *High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects.* Nature, 2016. **529**(7587): p. 490-5.

64. Chen, J.S., et al., *Enhanced proofreading governs CRISPR-Cas9 targeting accuracy.* Nature, 2017. **550**(7676): p. 407-410.

65. Casini, A., et al., *A highly specific SpCas9 variant is identified by in vivo screening in yeast.* Nat Biotechnol, 2018. **36**(3): p. 265-271.

66. Vakulskas, C.A., et al., *A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells.* Nat Med, 2018. **24**(8): p. 1216-1224.

67. Kleinstiver, B.P., et al., *Engineered CRISPR-Cas9 nucleases with altered PAM specificities.* Nature, 2015. **523**(7561): p. 481-5.

68. Kleinstiver, B.P., et al., *Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition.* Nat Biotechnol, 2015. **33**(12): p. 1293-1298.

69. Hu, J.H., et al., *Evolved Cas9 variants with broad PAM compatibility and high DNA specificity.* Nature, 2018. **556**(7699): p. 57-63.

70. Nishimasu, H., et al., *Engineered CRISPR-Cas9 nuclease with expanded targeting space.* Science, 2018. **361**(6408): p. 1259-1262.

71. Sander, J.D. and J.K. Joung, *CRISPR-Cas systems for editing, regulating and targeting genomes.* Nat Biotechnol, 2014. **32**(4): p. 347-55.

72. Swiech, L., et al., *In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9.* Nat Biotechnol, 2015. **33**(1): p. 102-6.

73. Senis, E., et al., *CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox.* Biotechnol J, 2014. **9**(11): p. 1402-12.

74. Fine, E.J., et al., *Trans-spliced Cas9 allows cleavage of HBB and CCR5 genes in human cells using compact expression cassettes.* Sci Rep, 2015. **5**: p. 10777.

75. Truong, D.J., et al., *Development of an intein-mediated split-Cas9 system for gene therapy.* Nucleic Acids Res, 2015. **43**(13): p. 6450-8.

76. Chew, W.L., et al., *A multifunctional AAV-CRISPR-Cas9 and its host response.* Nat Methods, 2016. **13**(10): p. 868-74.

77. Ran, F.A., et al., *In vivo genome editing using Staphylococcus aureus Cas9.* Nature, 2015. **520**(7546): p. 186-91.

78. Tabebordbar, M., et al., *In vivo gene editing in dystrophic mouse muscle and muscle stem cells.* Science, 2016. **351**(6271): p. 407-411.

79. Nelson, C.E., et al., *In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy.* Science, 2016. **351**(6271): p. 403-7.

80. Thakore, P.I., et al., *RNA-guided transcriptional silencing in vivo with S. aureus CRISPR-Cas9 repressors.* Nat Commun, 2018. **9**(1): p. 1674.

81. Gaj, T., et al., *In vivo genome editing improves motor function and extends survival in a mouse model of ALS.* Sci Adv, 2017. **3**(12): p. eaar3952.

82. Charlesworth, C.T., et al., *Identification of preexisting adaptive immunity to Cas9 proteins in humans.* Nat Med, 2019. **25**(2): p. 249-254.

83. Hirata, R., et al., *Targeted transgene insertion into human chromosomes by adeno-associated virus vectors.* Nat Biotechnol, 2002. **20**(7): p. 735-8.

84. Miller, D.G., L.M. Petek, and D.W. Russell, *Human gene targeting by adeno-associated virus vectors is enhanced by DNA double-strand breaks.* Mol Cell Biol, 2003. **23**(10): p. 3550-7.

85. Porteus, M.H., et al., *Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks.* Mol Cell Biol, 2003. **23**(10): p. 3558-65.

86. Gaj, T., et al., *Targeted gene knock-in by homology-directed genome editing using Cas9 ribonucleoprotein and AAV donor delivery.* Nucleic Acids Res, 2017. **45**(11): p. e98.

87. Dever, D.P., et al., *CRISPR/Cas9 beta-globin gene targeting in human haematopoietic stem cells.* Nature, 2016. **539**(7629): p. 384-389.

88. Sharma, R., et al., *In vivo genome editing of the albumin locus as a platform for protein replacement therapy.* Blood, 2015. **126**(15): p. 1777-84.

89. Yang, Y., et al., *A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice.* Nat Biotechnol, 2016. **34**(3): p. 334-8.

90. Sather, B.D., et al., *Efficient modification of CCR5 in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template.* Sci Transl Med, 2015. **7**(307): p. 307ra156.

91. Wang, J., et al., *Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors.* Nat Biotechnol, 2015. **33**(12): p. 1256-1263.

92. Yin, H., et al., *Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo.* Nat Biotechnol, 2016. **34**(3): p. 328-33.

93. Martin, R.M., et al., *Highly Efficient and Marker-free Genome Editing of Human Pluripotent Stem Cells by CRISPR-Cas9 RNP and AAV6 Donor-Mediated Homologous Recombination.* Cell Stem Cell, 2019. **24**(5): p. 821-828 e5.

94. Yoon, Y., et al., *Streamlined ex vivo and in vivo genome editing in mouse embryos using recombinant adeno-associated viruses.* Nat Commun, 2018. **9**(1): p. 412.

95. Bak, R.O. and M.H. Porteus, *CRISPR-Mediated Integration of Large Gene Cassettes Using AAV Donor Vectors.* Cell Rep, 2017. **20**(3): p. 750-756.

Chapter 2: The Ongoing Evolution and Translational Potential of Novel AAV Vectors for Human Gene Therapy

This chapter is adapted from a manuscript published as

S. Sun, D.V. Schaffer. The Ongoing Evolution and Translational Potential of Novel AAV Vectors for Human Gene Therapy. *Cell and Gene Therapy Insights*, 2019. **5**(4): p.429-442.

2.1 Introduction

The potential of AAV-mediated gene therapy for long-term efficacy has irrevocably disrupted the rare disease market, with full realization brought forth by the FDA's recent landmark approval of Luxturna for biallelic RPE65-associated retinal dystrophies. Due to their relative safety and ability to achieve persistent transgene expression, natural AAV serotypes, re-purposed as recombinant gene delivery vectors to treat debilitating genetic disorders, have yielded promising results in additional clinical trials for hemophilia A[1] and B[2], spinal muscular atrophy type 1[3], and choroideremia[4]. While these important studies demonstrate the strong therapeutic potential of AAV vectors, numerous delivery barriers must be overcome to apply this breakthrough technology to larger patient cohorts and a wider range of genetic and idiopathic diseases. The need for enhanced transduction efficiency and target cell specificity has stimulated the development of innovative vector engineering approaches and led to insightful discoveries that will guide the ongoing evolution of AAV capsid design.

Initial efforts to create novel and more potent AAV variants focused on site-directed mutagenesis of capsid surface residues. These early rational design studies included generation of single point mutants with altered binding and antibody neutralization activity[5], tyrosine mutants with increased transgene expression[6], and chimeras with targeting peptide insertions[7] or alternate receptor footprints[8] for re-directing viral tropism. Limited knowledge of AAV capsid structure, however, has driven widespread adoption of directed evolution[9], a powerful high-throughput engineering approach that does not require mechanistic understanding of virus-cell or virus-antibody interactions. Several pioneering studies provided proof of principle that careful design of selective pressures can produce new AAV vectors that are better tailored to clinical needs[10-13], suggesting that directed evolution can be extended to a variety of gene delivery challenges. Indeed, this strategy has created synthetic AAVs with *in vitro* and *in vivo* infectivity of previously non-permissive cells[14-18], antibody-evading capabilities[19-21], superior transduction efficiency[22-25], improved target specificity[25, 26], and enhanced axonal transport in neurons[27].

Here, we highlight the latest and most impactful advances in AAV vector engineering that build on our growing understanding of AAV capsid structure and function. We also discuss exciting developments in vector characterization that have been enabled by next-generation sequencing (NGS) technology and implications for product development, as novel AAV variants progress from bench to bedside.

2.2 *In vivo* Selection Schemes for Delivery to Target Cell Populations

Inadequate biodistribution to therapeutically relevant cell populations has motivated vector engineering approaches to address the high cellular heterogeneity of target tissues and offer cell type selectivity. For example, transduction of photoreceptors within the outer retina is a longstanding gene delivery challenge that has typically been bypassed with direct subretinal injections to treat inherited blinding disorders[28]. Of late, less invasive intravitreal administration of AAV vectors has been widely pursued as an alternative that can potentially achieve pan-retinal transduction while avoiding procedure-induced retinal damage. To identify novel AAV variants that can penetrate multiple cell layers from the vitreous and reach the light-sensing parts of the eye, an *in vivo* directed evolution strategy, in which mouse rod photoreceptors were isolated to recover AAV mutants that localized to these target cells, was applied[29] (Figure 1A). With this selection scheme, an AAV2 variant called 7m8 was found to show broad and efficient infectivity of photoreceptors and retinal pigmental epithelium across the murine retina. Of note, 7m8 recently entered a clinical trial for age-related macular degeneration (clinicaltrials.gov identifier: NCT03748784).

More recently, a Cre-dependent approach was devised to recover AAV capsids that transduce pre-defined target cell populations in mouse without post-mortem sorting[30, 31] (Figure 1A). In this clever selection scheme, loxP sites were incorporated into AAV libraries to drive unidirectional inversion of the floxed sequence in Cre-expressing cells, and PCR primers were designed to selectively amplify recombined AAV genomes. The method was applied in Cre driver mouse lines to enrich for clones capable of crossing the blood brain barrier (BBB) and localizing to functional cell types within the central nervous system (CNS)[31]. Upon intravenous administration, an AAV9 variant called AAV-PHP.B was found to transduce the majority of astrocytes and neurons throughout the mouse brain, outperforming its parental serotype by at least 40-fold in gene transfer efficiency. Notably, further evolution on AAV-PHP.B, conducted through additional diversification of its heptamer peptide, yielded an even more potent variant called AAV-PHP.eB that retains its parental tropism but demonstrates enhanced infectivity of neurons and glia[32]. This powerful Cre-dependent strategy was also employed to produce a mosaic AAV variant called SCH9 that infects 60% of adult neural stem cells in both cerebral hemispheres following unilateral intracerebroventricular injection[30]. Collectively, these studies provide compelling proof-of-concept demonstration that Cre-mediated selectivity can be leveraged to engineer AAV capsids with improved delivery to specific murine cell populations.

2.3 Emergence of Semi-Rational AAV Vector Engineering Strategies

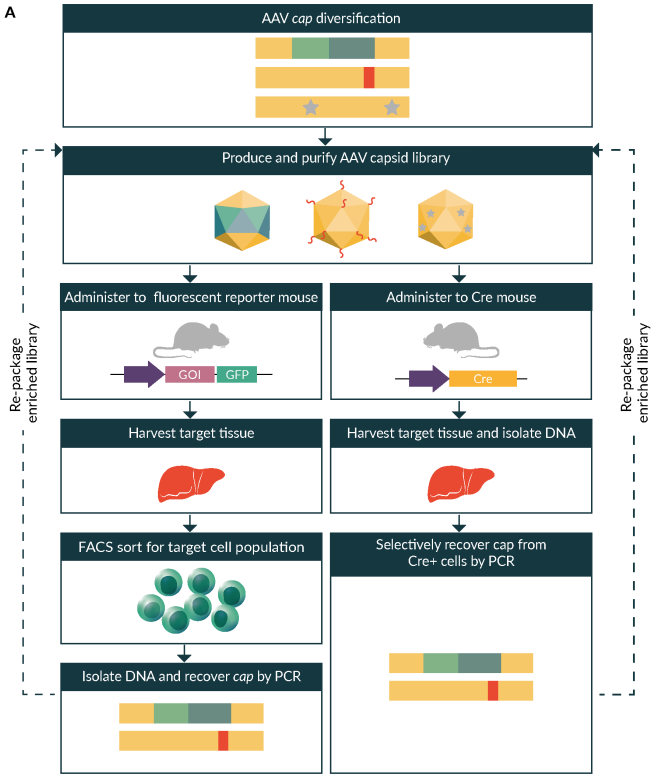
Cryo-electron microscopy (cryo-EM) and X-ray crystallography have been instrumental in providing 3D image reconstructions of AAV capsids[33-37] and the complexes they form with monoclonal antibodies (mAbs)[38-40]. This repertoire of structural models, in combination with insights gained from peptide scanning techniques, has enabled high-resolution epitope mapping and targeting of antigenic footprints to create novel AAV serotypes with disrupted antibody recognition sites[41]. In a structure-guided directed evolution approach, mutagenesis of antigenic hotspots on AAV1 was coupled with iterative selection on a permissive cell line to produce variants that resist mAb neutralization without compromising packaging and transduction

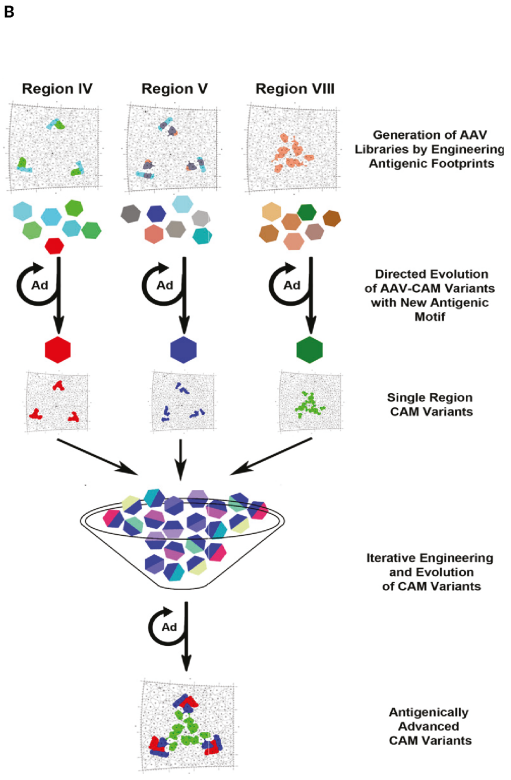
efficiency[42] (Figure 1B). Moreover, modularity of the targeted antigenic clusters facilitated the combinatorial engineering of multiple distinct footprints on a single capsid template, yielding a mutant that evades polyclonal anti-AAV1 sera from immunized mice and non-human primates (NHPs) up to 8-fold more effectively than AAV1.

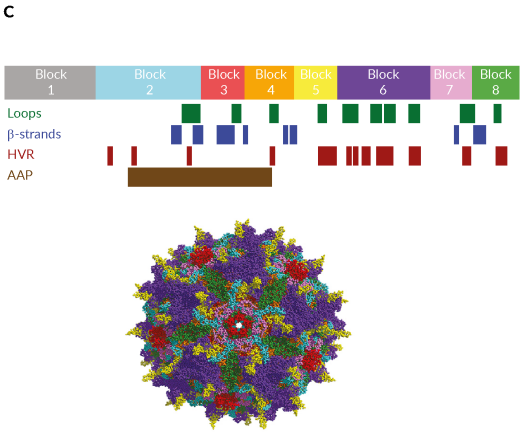
Structural data has also been utilized to inform the generation of functional AAV chimeras with high sequence diversity. SCHEMA, a computational algorithm that predicts optimal crossover points for DNA shuffling of homologous proteins by minimizing structural disruption[43], was applied to the AAV capsid to guide the recombination of distantly related serotypes while retaining core viral functions such as particle assembly and transduction[30, 44]. In contrast to traditional DNA shuffling techniques that preferentially place crossovers at regions of high homology, an AAV SCHEMA library constructed from six natural serotypes captured relatively equal representation from each parent and contained infectious clones with unique tropism[30] (Figure 1C). Interestingly, a SCHEMA variant called SCH9 displayed lower susceptibility to human polyclonal immunoglobulin compared to its parental strains, putatively due to disruption of key neutralizing antibody (NAb) epitopes[30]. The work described thus far showcases the power of integrating rational design with directed evolution to effectively narrow our search of the vast AAV capsid protein space to functionally rich regions and identify next-generation vectors that are more likely to suit clinical needs.

In silico reconstruction of putative "parental" AAV capsids has also proven to be a powerful capsid engineering approach (Figure 1D). In two independent studies, ancestral AAV libraries comprised of predicted evolutionary intermediates were synthesized *de novo* and displayed broad tropism and enhanced thermostability[45, 46]. These variants show distinct transduction profiles, including potent infectivity of cochlear cells *in vivo*[47, 48], and offer great potential as starting points for directed evolution due to their inherent promiscuity and plasticity[45]. In spite of their distinctive properties, ancestral capsids cross-react with antibodies raised against contemporary AAVs, suggesting that further evolution may be necessary to evade immune detection[45, 46].

Lastly, strategies to overcome humoral immunity without modifying the AAV capsid are actively being explored. For example, AAV vector preparations can be supplemented with empty capsid decoys to adsorb NAbs in a dose-dependent manner and block their inhibition of therapeutic efficacy, though this approach requires patient-specific optimization of full-to-empty capsid ratios and may adversely contribute to vector immunogenicity through elevated capsid antigen presentation[49]. Exosome-associated AAVs (exo-AAVs) have recently emerged as novel vectors with improved gene transfer to various target tissues, relative to their non-enveloped counterparts[50-52], and the ability to shield AAV capsids from circulating human NAbs[53, 54]. Despite the therapeutic promise of exo-AAVs, this technology is young, as challenges with scalable manufacturing and characterization of the composition and reproducibility of these microvesicles remain.


A
AAV cap diversification
Produce and purify AAV capsid library
Administer to fluorescent reporter mouse
GOI
GFP
Administer to Cre mouse
Cre
Re-package enriched library
Re-package enriched library
Harvest target tissue
Harvest target tissue and isolate DNA
FACS sort for target cell population
Selectively recover cap from Cre+ cells by PCR
Isolate DNA and recover cap by PCR


B
Region IV
Region V
Region VIII
Generation of AAV Libraries by Engineering Antigenic Footprints
Ad
Ad
Ad
Directed Evolution of AAV-CAM Variants with New Antigenic Motif
Single Region CAM Variants
Iterative Engineering and Evolution of CAM Variants
Ad
Antigenically Advanced CAM Variants


C
Block 1
Block 2
Block 3
Block 4
Block 5
Block 6
Block 7
Block 8
Loops
β-strands
HVR
AAP

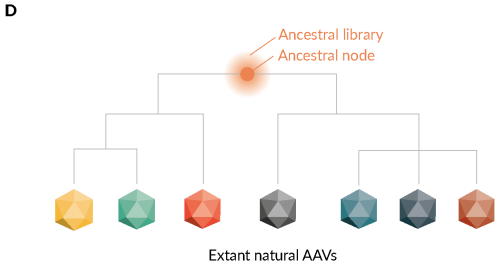

D
Ancestral library
Ancestral node
Extant natural AAVs

Figure 2.1. Novel AAV Selection Strategies.

A) Directed evolution of AAV for cell-type-specific transduction. The AAV *cap* gene is first diversified to generate a library of capsid mutants that is subsequently assembled into viral particles. For post-mortem sorting of a target cell population (left), AAV capsid libraries are administered into transgenic mice expressing a fluorescent protein in the cell type of interest. Following harvest and dissociation of the target tissue, the desired cell population is isolated by FACS and viral genomes are recovered by PCR amplification. For Cre-dependent recovery from a target cell population (right), AAV capsid libraries flanked by loxP sites are administered into transgenic mice expressing Cre recombinase in the cell type of interest. Cre mediates inversion of viral genomes that have localized to the target cell population, and inversion-specific primers are used to selectively amplify recombined genomes from the target tissue. Recovered variants are re-packaged into virions and enriched through iterative rounds of selection.

B) Structure-guided evolution of serologically distinct AAV variants[42]. Individual antigenic footprints are mapped onto a capsid template and mutagenized to generate AAV capsid libraries. Libraries are iteratively selected for infectivity of a permissive cell line to yield "single region" AAV mutants that escape monoclonal antibody neutralization. Antigenic clusters from these variants are then combined and further engineered and evolved to produce "advanced" AAVs with the potential to evade polyclonal sera.

C) SCHEMA-guided design of AAV chimeric library[30]. Crossover positions in the AAV *cap* gene are selected by SCHEMA to minimize structural disruption while maximizing functional diversity. Shuffled AAV *cap* sequence "blocks" are represented by colored boxes and aligned with capsid loops, β-strands, hypervariable regions (HVR), and the assembly-activating protein (AAP) to indicate structure-function correlations. Blocks are mapped onto the AAV2 crystal structure (PDB: 1LP3) in PyMOL to generate a three-dimensional model of the SCHEMA capsid design.

D) *In silico* reconstruction of putative AAV ancestors. A representative phylogenetic tree depicting extant natural AAVs and their relation to an inferred ancestral node is shown. A library of predicted AAV ancestors can be *de novo* synthesized to compensate for ambiguity in evolutionary relationships and capture sequence divergence at highly variable residues near a node.

2.4 NGS for High-Throughput Phenotypic Analysis

In the past five years, NGS technology has revolutionized AAV vector characterization by enabling massively parallel phenotypic analysis of barcoded AAV libraries. AAV Barcode-Seq, a method by which AAV mutants are individually tagged with unique barcodes and characterized with a range of *in vitro* and *in vivo* assays via multiplexed Illumina barcode sequencing, marked the first application of NGS to AAV vectors and allowed systematic dissection of capsid structure and function at the single amino acid level[55] (Figure 2). Since then, deep sequencing has been adopted to quantify vector genome biodistribution and transgene mRNA expression of AAV variants in a high-throughput manner[56]. This approach simplifies the costly and laborious process of *in vivo* vector characterization and may significantly reduce the ethical and financial burden of screening multiple candidate vectors in large animal models.

Not only has NGS been a powerful tool for evaluating biodistribution profiles and mapping functional domains within the AAV capsid, it has also been employed to track clonal enrichment and guide determination of evolutionary convergence during AAV vector selection. Changes in relative abundance of individual peptide insertions[57] and antigenic footprints[42] between iterative selection rounds were traced in two directed evolution studies, achieving unprecedented depth of characterization. Furthermore, analysis of clonal frequencies in off-target organs aided

identification of candidates that potentially demonstrate tissue-selective homing[57]. Importantly, NGS-based screening has lent invaluable insight into selection kinetics, in particular revealing misleading clonal dominance in early rounds[57]. Recombination-free methods to link capsid mutations to *in cis* molecular barcodes for subsequent deep sequencing are also in development[58, 59]. As Illumina's short read length has restricted its compatibility to AAV libraries containing highly localized genetic diversity, improvements to this technology will broaden applicability to chimeras and variants with point mutations distributed across the AAV capsid gene.

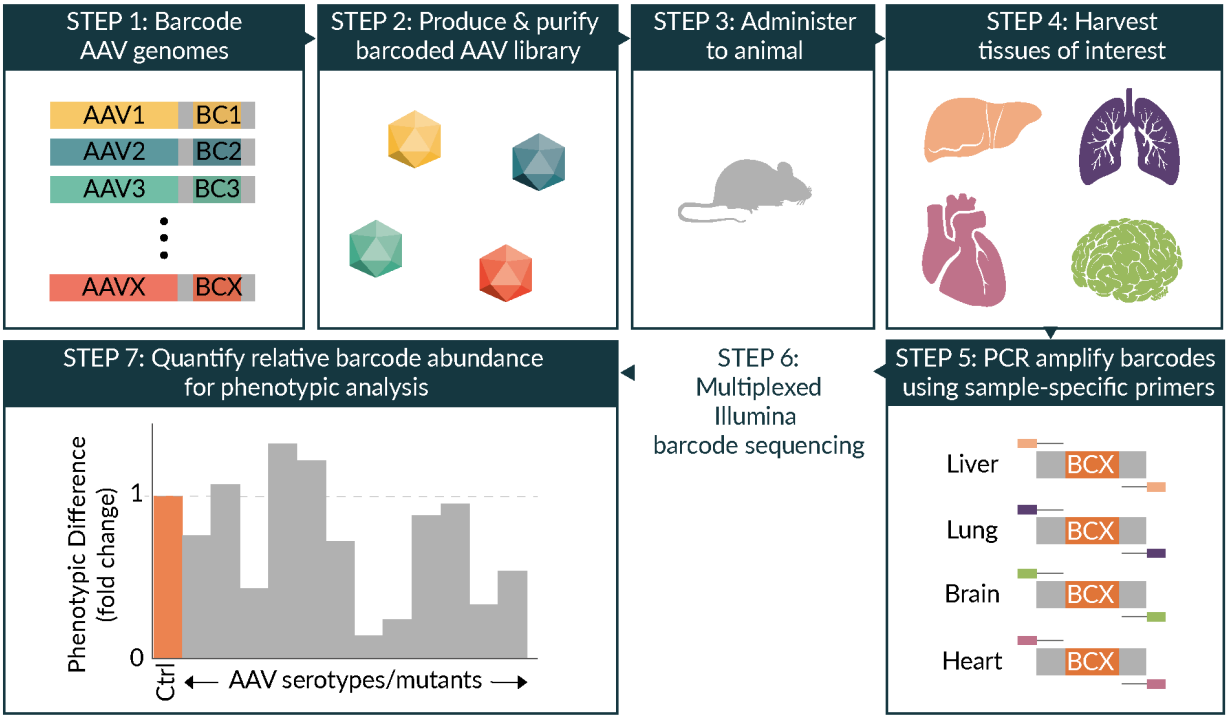


Figure 2.2. AAV Barcode-Seq for High-Throughput Phenotypic Analysis *in vivo*.

A panel of AAV serotypes and mutants is selected for characterization, and each viral genome is paired with a unique *in cis* DNA barcode (VBC) that serves as a proxy for its associated AAV clone. The resulting barcoded AAV library is assembled into viral particles, which are subsequently purified and administered to an animal. Various tissues of interest are then collected for DNA extraction, and VBCs are recovered from each tissue by PCR using primers with sample-specific identifiers (SBC). These SBCs enable massively parallel sequencing of PCR-amplified VBCs, pooled from different samples, using Illumina technology. Sequencing reads are then deconvoluted through the identity of the corresponding SBC, and VBC read counts for each sample are analyzed to systemically determine phenotypic differences between clones and draw high-resolution functional maps of the AAV capsid.

2.5 Manufacturability and Scalable Purification of Engineered AAV Vectors

An important consideration for commercialization of novel AAV variants is their production efficiency and associated manufacturing burden. As natural AAVs differ in vector yield[60, 61], single amino acid substitutions have been identified to increase viral titer of specific serotypes by putatively enhancing folding and stability of viral protein subunits[62] or restoring monomer-monomer interactions within the multimeric capsid[63]. These mutations, if functionally neutral in the context of an engineered AAV vector, may be introduced to rescue efficient production, especially when directed evolution has already enriched for variants that package with near-wildtype efficiency through iterative rounds of viral replication and functional selection. Of note, mosaic capsids recombined from multiple parental serotypes are more likely than point mutants to capture high sequence diversity while retaining core viral functions, including genome packaging[21, 23, 64]. In fact, several SCHEMA AAV chimeras, as previously discussed, preserved the ability to fold and assemble properly, despite their high structural disruption scores[44]. As recombination potentially disrupts the activity of AAV's assembly-activating protein (AAP)[30, 44, 65], a chaperone that regulates virion assembly and is encoded in an alternate open reading frame of the AAV capsid gene[66, 67], there has been immense interest in investigating its role in genome packaging[68-72] and engineering a universal AAP to be supplied *in trans* for less biased AAV library selections and improved viral titers[73]. Lastly, sites within the AAV capsid that tolerate peptide insertion and minimally interfere with viral particle assembly have been extensively studied to display targeting ligands on the capsid surface[74, 75]. It has, however, become evident that the identity of the peptide alone can dramatically alter packaging efficiency[31]. Thus, screening multiple candidates will likely be necessary to strike a balance between desired phenotypes and feasible vector production.

As the AAV vector toolkit expands, the demand for high-throughput and economically viable purification platforms to support commercialization of a diverse range of AAV serotypes grows more urgent. Due to its large binding capacity and ease of scalability, column chromatography has been widely adopted for cGMP vector manufacturing[76, 77]. To produce clinical grade preparations of serotypes that are not amenable to early affinity-based methods[78, 79], one- and two-step ion-exchange chromatography (IEX) has been successfully adapted for AAV purification[80, 81]. Importantly, it can be fine-tuned to discriminate between empty and full capsids[82], but its high sensitivity to capsid surface charge requires extensive serotype-specific optimization and hinders efficient standardization across alternate AAVs. The engineering of single-domain anti-AAV antibodies has led to commercial availability of single-capture affinity resins: AVB Sepharose, which selectively binds to a limited variety of AAV subclasses[83, 84], and POROS™ CaptureSelect™ AAVX, which recognizes an epitope conserved across a wider range of serotypes. These technologies, paired with IEX as a polishing step to separate full from empty viral particles, may enable more cost-effective and streamlined purification of multiple AAV serotypes.

2.6 Translation Insight

Driven by accelerated research and development of new AAV serotypes, numerous engineered variants including Spark100 (clinicaltrials.gov identifier: NCT03307980), LK03 (NCT03003533), 7m8 (NCT03748784), and AAV2tYF (NCT03316560), recently made their clinical debut, bolstering the promise of novel vector technologies. As the safety of AAV-mediated gene therapy in humans becomes increasingly compelling, discordance in delivery barriers[85, 86] across species is now apparent and limits translatability despite efficacy observed in preclinical studies. In the CUPID-2b gene therapy trial (clinicaltrials.gov identifier: NCT01643330) for advanced heart failure, AAV1 uptake in human myocardium was <0.2% of the vector copy number observed in animal models, preventing the possibility of any functional improvement[87]. In another sobering study, the striking transduction profile of AAV-PHP.B in mouse CNS[31] fell far short of expectations in NHP[88] and even mouse strains other than C57BL6/J[89]. These discrepancies motivate the development of AAV vector selection strategies in large animal models to more accurately recapitulate tissue architecture present in humans. Xenotransplantation of target human cells in mice may also better model gene delivery barriers[20, 24], though this approach is limited to cell types that can partially repopulate host tissue upon engraftment.

In vivo preclinical studies have also proven to be poor predictors of crucial virus-host interactions, in particular cytotoxic T-cell responses to the AAV capsid upon systemic vector administration[90]. Adoptive transfer models have been generated to more accurately study capsid antigen presentation and subsequent CD8+ T-cell mediated destruction of transduced cells[91, 92]. Immunosuppressive regimens[93] and synthetic vaccine particles encapsulating rapamycin (SVP[Rapa])[94] have also been applied to modulate vector immunogenicity. These potentially detrimental immune responses appear to occur in a vector dose-dependent manner[95, 96], suggesting that novel AAVs with improved delivery efficiency and target specificity may mitigate T-cell activation due to their reduced capsid antigen load.

Finally, as reviewed in Colella *et al*[90], pre-existing humoral immunity remains a profound hurdle to the efficacy of AAV-mediated gene transfer as a therapeutic modality. For programs requiring intravenous vector administration, patient eligibility is significantly restricted by stringent criteria for low or undetectable NAb titers[1, 93]. While the creation of AAV mutants that evade immune detection has strong potential to expand patient enrollment, cross-reactivity of anti-AAV serum against various serotypes[97] necessitates thorough screening of all vectors on a per-patient basis[98, 99] and standardization of sensitive NAb detection assays[100].

The prior application of a rationally engineered AAV variant in a Phase 1 trial for Duchenne muscular dystrophy[101] and recent FDA clearance of several novel AAVs for use in human patients have together sparked well-warranted optimism about emerging vector technologies. As we await highly anticipated clinical data that will hopefully uphold their superior performance, our expanding knowledge of basic AAV biology and structure has guided the development of creative vector evolution techniques, yielding properties that have long been sought. Furthermore, structural and phenotypic insight obtained from rising NGS-based analytical methods may enable more targeted and elegant engineering approaches in the near future. While AAV capsid design has advanced by leaps and bounds since its conception, species-specific delivery barriers and pre-existing humoral immunity remain two major hurdles that will require engineering ingenuity and greater understanding of virus-host interactions to overcome.

2.7 References

1. Rangarajan, S., et al., *AAV5-Factor VIII Gene Transfer in Severe Hemophilia A.* N Engl J Med, 2017. **377**(26): p. 2519-2530.
2. George, L.A., et al., *Hemophilia B Gene Therapy with a High-Specific-Activity Factor IX Variant.* N Engl J Med, 2017. **377**(23): p. 2215-2227.
3. Mendell, J.R., et al., *Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy.* N Engl J Med, 2017. **377**(18): p. 1713-1722.
4. Lam, B.L., et al., *Choroideremia Gene Therapy Phase 2 Clinical Trial: 24-Month Results.* Am J Ophthalmol, 2019. **197**: p. 65-73.
5. Lochrie, M.A., et al., *Mutations on the external surfaces of adeno-associated virus type 2 capsids that affect transduction and neutralization.* J Virol, 2006. **80**(2): p. 821-34.
6. Zhong, L., et al., *Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses.* Proc Natl Acad Sci U S A, 2008. **105**(22): p. 7827-32.
7. Shi, W., G.S. Arnold, and J.S. Bartlett, *Insertional mutagenesis of the adeno-associated virus type 2 (AAV2) capsid gene and generation of AAV2 vectors targeted to alternative cell-surface receptors.* Hum Gene Ther, 2001. **12**(14): p. 1697-711.
8. Asokan, A., et al., *Reengineering a receptor footprint of adeno-associated virus enables selective and systemic gene transfer to muscle.* Nat Biotechnol, 2010. **28**(1): p. 79-82.
9. Romero, P.A. and F.H. Arnold, *Exploring protein fitness landscapes by directed evolution.* Nat Rev Mol Cell Biol, 2009. **10**(12): p. 866-76.
10. Perabo, L., et al., *In vitro selection of viral vectors with modified tropism: the adeno-associated virus display.* Mol Ther, 2003. **8**(1): p. 151-7.
11. Muller, O.J., et al., *Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors.* Nat Biotechnol, 2003. **21**(9): p. 1040-6.
12. Maheshri, N., et al., *Directed evolution of adeno-associated virus yields enhanced gene delivery vectors.* Nat Biotechnol, 2006. **24**(2): p. 198-204.
13. Perabo, L., et al., *Combinatorial engineering of a gene therapy vector: directed evolution of adeno-associated virus.* J Gene Med, 2006. **8**(2): p. 155-62.
14. Excoffon, K.J., et al., *Directed evolution of adeno-associated virus to an infectious respiratory virus.* Proc Natl Acad Sci U S A, 2009. **106**(10): p. 3865-70.
15. Li, W., et al., *Generation of novel AAV variants by directed evolution for improved CFTR delivery to human ciliated airway epithelium.* Mol Ther, 2009. **17**(12): p. 2067-77.
16. Asuri, P., et al., *Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells.* Mol Ther, 2012. **20**(2): p. 329-38.
17. Jang, J.H., et al., *An evolved adeno-associated viral variant enhances gene delivery and gene targeting in neural stem cells.* Mol Ther, 2011. **19**(4): p. 667-75.

18. Varadi, K., et al., *Novel random peptide libraries displayed on AAV serotype 9 for selection of endothelial cell-directed gene transfer vectors.* Gene Ther, 2012. **19**(8): p. 800-9.

19. Li, C., et al., *Development of Patient-specific AAV Vectors After Neutralizing Antibody Selection for Enhanced Muscle Gene Transfer.* Mol Ther, 2016. **24**(1): p. 53-65.

20. Paulk, N.K., et al., *Bioengineered AAV Capsids with Combined High Human Liver Transduction In Vivo and Unique Humoral Seroreactivity.* Mol Ther, 2018. **26**(1): p. 289-303.

21. Li, W., et al., *Engineering and Selection of Shuffled AAV Genomes: A New Strategy for Producing Targeted Biological Nanoparticles.* Mol Ther, 2008. **16**(7): p. 1252-1260.

22. Choudhury, S.R., et al., *In Vivo Selection Yields AAV-B1 Capsid for Central Nervous System and Muscle Gene Therapy.* Mol Ther, 2016. **24**(7): p. 1247-57.

23. Grimm, D., et al., *In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses.* J Virol, 2008. **82**(12): p. 5887-911.

24. Lisowski, L., et al., *Selection and evaluation of clinically relevant AAV variants in a xenograft liver model.* Nature, 2014. **506**(7488): p. 382-6.

25. Klimczak, R.R., et al., *A novel adeno-associated viral variant for efficient and selective intravitreal transduction of rat Muller cells.* PLoS One, 2009. **4**(10): p. e7467.

26. Yang, L., et al., *A myocardium tropic adeno-associated virus (AAV) evolved by DNA shuffling and in vivo selection.* Proc Natl Acad Sci U S A, 2009. **106**(10): p. 3946-51.

27. Tervo, D.G., et al., *A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons.* Neuron, 2016. **92**(2): p. 372-382.

28. Schon, C., M. Biel, and S. Michalakis, *Retinal gene delivery by adeno-associated virus (AAV) vectors: Strategies and applications.* Eur J Pharm Biopharm, 2015. **95**(Pt B): p. 343-52.

29. Dalkara, D., et al., *In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous.* Sci Transl Med, 2013. **5**(189): p. 189ra76.

30. Ojala, D.S., et al., *In Vivo Selection of a Computationally Designed SCHEMA AAV Library Yields a Novel Variant for Infection of Adult Neural Stem Cells in the SVZ.* Mol Ther, 2017.

31. Deverman, B.E., et al., *Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain.* Nat Biotechnol, 2016. **34**(2): p. 204-9.

32. Chan, K.Y., et al., *Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems.* Nat Neurosci, 2017.

33. Mikals, K., et al., *The structure of AAVrh32.33, a novel gene delivery vector.* J Struct Biol, 2014. **186**(2): p. 308-17.

34. Halder, S., et al., *Structure of neurotropic adeno-associated virus AAVrh.8.* J Struct Biol, 2015. **192**(1): p. 21-36.

35. Drouin, L.M., et al., *Cryo-electron Microscopy Reconstruction and Stability Studies of the Wild Type and the R432A Variant of Adeno-associated Virus Type 2 Reveal that Capsid Structural Stability Is a Major Factor in Genome Packaging.* J Virol, 2016. **90**(19): p. 8542-51.

36. Burg, M., et al., *Atomic structure of a rationally engineered gene delivery vector, AAV2.5.* J Struct Biol, 2018. **203**(3): p. 236-241.

37. Tan, Y.Z., et al., *Sub-2 A Ewald curvature corrected structure of an AAV2 capsid variant.* Nat Commun, 2018. **9**(1): p. 3628.

38. Tseng, Y.S., et al., *Adeno-associated virus serotype 1 (AAV1)- and AAV5-antibody complex structures reveal evolutionary commonalities in parvovirus antigenic reactivity.* J Virol, 2015. **89**(3): p. 1794-808.

39. Bennett, A.D., et al., *AAV6 K531 serves a dual function in selective receptor and antibody ADK6 recognition.* Virology, 2018. **518**: p. 369-376.

40. Jose, A., et al., *High-Resolution Structural Characterization of a New Adeno-associated Virus Serotype 5 Antibody Epitope toward Engineering Antibody-Resistant Recombinant Gene Delivery Vectors.* J Virol, 2019. **93**(1).

41. Tseng, Y.S. and M. Agbandje-McKenna, *Mapping the AAV Capsid Host Antibody Response toward the Development of Second Generation Gene Delivery Vectors.* Front Immunol, 2014. **5**: p. 9.

42. Tse, L.V., et al., *Structure-guided evolution of antigenically distinct adeno-associated virus variants for immune evasion.* Proc Natl Acad Sci U S A, 2017. **114**(24): p. E4812-E4821.

43. Voigt, C.A., et al., *Protein building blocks preserved by recombination.* Nat Struct Biol, 2002. **9**(7): p. 553-8.

44. Ho, M.L., et al., *SCHEMA computational design of virus capsid chimeras: calibrating how genome packaging, protection, and transduction correlate with calculated structural disruption.* ACS Synth Biol, 2013. **2**(12): p. 724-33.

45. Santiago-Ortiz, J., et al., *AAV ancestral reconstruction library enables selection of broadly infectious viral variants.* Gene Ther, 2015. **22**(12): p. 934-46.

46. Zinn, E., et al., *In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector.* Cell Rep, 2015. **12**(6): p. 1056-68.

47. Pan, B., et al., *Gene therapy restores auditory and vestibular function in a mouse model of Usher syndrome type 1c.* Nat Biotechnol, 2017. **35**(3): p. 264-272.

48. Landegger, L.D., et al., *A synthetic AAV vector enables safe and efficient gene transfer to the mammalian inner ear.* Nat Biotechnol, 2017. **35**(3): p. 280-284.

49. Mingozzi, F., et al., *Overcoming preexisting humoral immunity to AAV using capsid decoys.* Sci Transl Med, 2013. **5**(194): p. 194ra92.

50. Wassmer, S.J., et al., *Exosome-associated AAV2 vector mediates robust gene delivery into the murine retina upon intravitreal injection.* Sci Rep, 2017. **7**: p. 45329.

51. Hudry, E., et al., *Exosome-associated AAV vector as a robust and convenient neuroscience tool.* Gene Ther, 2016. **23**(4): p. 380-92.

52. Gyorgy, B., et al., *Rescue of Hearing by Gene Delivery to Inner-Ear Hair Cells Using Exosome-Associated AAV.* Mol Ther, 2017. **25**(2): p. 379-391.

53. Meliani, A., et al., *Enhanced liver gene transfer and evasion of preexisting humoral immunity with exosome-enveloped AAV vectors.* Blood Adv, 2017. **1**(23): p. 2019-2031.

54. Gyorgy, B., et al., *Naturally enveloped AAV vectors for shielding neutralizing antibodies and robust gene delivery in vivo.* Biomaterials, 2014. **35**(26): p. 7598-609.

55. Adachi, K., et al., *Drawing a high-resolution functional map of adeno-associated virus capsid by massively parallel sequencing.* Nat Commun, 2014. **5**: p. 3075.

56. Marsic, D., H.R. Mendez-Gomez, and S. Zolotukhin, *High-accuracy biodistribution analysis of adeno-associated virus variants by double barcode sequencing.* Mol Ther Methods Clin Dev, 2015. **2**: p. 15041.

57. Korbelin, J., et al., *Pulmonary Targeting of Adeno-associated Viral Vectors by Next-generation Sequencing-guided Screening of Random Capsid Displayed Peptide Libraries.* Mol Ther, 2016. **24**(6): p. 1050-1061.

58. Davidsson, M., et al., *A novel process of viral vector barcoding and library preparation enables high-diversity library generation and recombination-free paired-end sequencing.* Sci Rep, 2016. **6**: p. 37563.

59. Davidsson, M.W., G.; Aldrin-Kirk, P.; Cardoso, T.; Nolbrant, S.; Hartnor, M.; Parmar, M.; Bjorklund, T., *Barcoded Rational AAV Vector Evolution enables systematic in vivo mapping of peptide binding motifs* bioRxiv, 2018.

60. Vandenberghe, L.H., et al., *Efficient serotype-dependent release of functional vector into the culture medium during adeno-associated virus manufacturing.* Hum Gene Ther, 2010. **21**(10): p. 1251-7.

61. Holehonnur, R., et al., *Adeno-associated viral serotypes produce differing titers and differentially transduce neurons within the rat basal and lateral amygdala.* BMC Neurosci, 2014. **15**: p. 28.

62. Wu, Z., et al., *Single amino acid changes can influence titer, heparin binding, and tissue tropism in different adeno-associated virus serotypes.* J Virol, 2006. **80**(22): p. 11393-7.

63. Vandenberghe, L.H., et al., *Naturally occurring singleton residues in AAV capsid impact vector performance and illustrate structural constraints.* Gene Ther, 2009. **16**(12): p. 1416-28.

64. Koerber, J.T., J.H. Jang, and D.V. Schaffer, *DNA shuffling of adeno-associated virus yields functionally diverse viral progeny.* Mol Ther, 2008. **16**(10): p. 1703-9.

65. Herrmann, A.K., et al., *Impact of the Assembly-Activating Protein on Molecular Evolution of Synthetic Adeno-Associated Virus Capsids.* Hum Gene Ther, 2019. **30**(1): p. 21-35.

66. Sonntag, F., K. Schmidt, and J.A. Kleinschmidt, *A viral assembly factor promotes AAV2 capsid formation in the nucleolus.* Proc Natl Acad Sci U S A, 2010. **107**(22): p. 10220-5.

67. Sonntag, F., et al., *The assembly-activating protein promotes capsid assembly of different adeno-associated virus serotypes.* J Virol, 2011. **85**(23): p. 12686-97.

68. Earley, L.F., et al., *Identification and characterization of nuclear and nucleolar localization signals in the adeno-associated virus serotype 2 assembly-activating protein.* J Virol, 2015. **89**(6): p. 3038-48.

69. Earley, L.F., et al., *Adeno-associated Virus (AAV) Assembly-Activating Protein Is Not an Essential Requirement for Capsid Assembly of AAV Serotypes 4, 5, and 11.* J Virol, 2017. **91**(3).

70. Grosse, S., et al., *Relevance of Assembly-Activating Protein for Adeno-associated Virus Vector Production and Capsid Protein Stability in Mammalian and Insect Cells.* J Virol, 2017. **91**(20).

71. Tse, L.V., et al., *Mapping and Engineering Functional Domains of the Assembly-Activating Protein of Adeno-associated Viruses.* J Virol, 2018. **92**(14).

72. Maurer, A.C., et al., *The Assembly-Activating Protein Promotes Stability and Interactions between AAV's Viral Proteins to Nucleate Capsid Assembly.* Cell Rep, 2018. **23**(6): p. 1817-1830.

73. Maurer, A.C., A.K. Cepeda Diaz, and L.H. Vandenberghe, *Residues on Adeno-associated Virus Capsid Lumen Dictate Interactions and Compatibility with the Assembly-Activating Protein.* J Virol, 2019. **93**(7).

74. Girod, A., et al., *Genetic capsid modifications allow efficient re-targeting of adeno-associated virus type 2.* Nat Med, 1999. **5**(12): p. 1438.

75. Rabinowitz, J.E., W. Xiao, and R.J. Samulski, *Insertional mutagenesis of AAV2 capsid and the production of recombinant virus.* Virology, 1999. **265**(2): p. 274-85.

76. Ayuso, E., F. Mingozzi, and F. Bosch, *Production, purification and characterization of adeno-associated vectors.* Curr Gene Ther, 2010. **10**(6): p. 423-36.

77. Qu, W., et al., *Scalable downstream strategies for purification of recombinant adeno-associated virus vectors in light of the properties.* Curr Pharm Biotechnol, 2015. **16**(8): p. 684-95.

78. Davidoff, A.M., et al., *Purification of recombinant adeno-associated virus type 8 vectors by ion exchange chromatography generates clinical grade vector stock.* J Virol Methods, 2004. **121**(2): p. 209-15.

79. Zolotukhin, S., et al., *Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors.* Methods, 2002. **28**(2): p. 158-67.

80. Okada, T., et al., *Scalable purification of adeno-associated virus serotype 1 (AAV1) and AAV8 vectors, using dual ion-exchange adsorptive membranes.* Hum Gene Ther, 2009. **20**(9): p. 1013-21.

81. Brument, N., et al., *A versatile and scalable two-step ion-exchange chromatography process for the purification of recombinant adeno-associated virus serotypes-2 and -5.* Mol Ther, 2002. **6**(5): p. 678-86.

82. Qu, G., et al., *Separation of adeno-associated virus type 2 empty particles from genome containing vectors by anion-exchange column chromatography.* J Virol Methods, 2007. **140**(1-2): p. 183-92.

83. Mietzsch, M., et al., *OneBac: platform for scalable and high-titer production of adeno-associated virus serotype 1-12 vectors for gene therapy.* Hum Gene Ther, 2014. **25**(3): p. 212-22.

84. Wang, Q., et al., *Identification of an adeno-associated virus binding epitope for AVB sepharose affinity resin.* Mol Ther Methods Clin Dev, 2015. **2**: p. 15040.

85. Liu, X., et al., *Species-specific differences in mouse and human airway epithelial biology of recombinant adeno-associated virus transduction.* Am J Respir Cell Mol Biol, 2006. **34**(1): p. 56-64.

86. Mussolino, C., et al., *AAV-mediated photoreceptor transduction of the pig cone-enriched retina.* Gene Ther, 2011. **18**(7): p. 637-45.

87. Hulot, J.S., K. Ishikawa, and R.J. Hajjar, *Gene therapy for the treatment of heart failure: promise postponed.* Eur Heart J, 2016. **37**(21): p. 1651-8.

88. Matsuzaki, Y., et al., *Intravenous administration of the adeno-associated virus-PHP.B capsid fails to upregulate transduction efficiency in the marmoset brain.* Neurosci Lett, 2018. **665**: p. 182-188.

89. Hordeaux, J., et al., *The Neurotropic Properties of AAV-PHP.B Are Limited to C57BL/6J Mice.* Mol Ther, 2018. **26**(3): p. 664-668.

90. Colella, P., G. Ronzitti, and F. Mingozzi, *Emerging Issues in AAV-Mediated In Vivo Gene Therapy.* Mol Ther Methods Clin Dev, 2018. **8**: p. 87-104.

91. Li, H., et al., *Adeno-associated virus vectors serotype 2 induce prolonged proliferation of capsid-specific CD8+ T cells in mice.* Mol Ther, 2011. **19**(3): p. 536-46.

92. Martino, A.T., et al., *Engineered AAV vector minimizes in vivo targeting of transduced hepatocytes by capsid-specific CD8+ T cells.* Blood, 2013. **121**(12): p. 2224-33.

93. Nathwani, A.C., et al., *Adenovirus-associated virus vector-mediated gene transfer in hemophilia B.* N Engl J Med, 2011. **365**(25): p. 2357-65.

94. Meliani, A., et al., *Antigen-selective modulation of AAV immunogenicity with tolerogenic rapamycin nanoparticles enables successful vector re-administration.* Nat Commun, 2018. **9**(1): p. 4098.

95. Pien, G.C., et al., *Capsid antigen presentation flags human hepatocytes for destruction after transduction by adeno-associated viral vectors.* J Clin Invest, 2009. **119**(6): p. 1688-95.

96. Kumar, S.R.P., et al., *The Balance between CD8(+) T Cell-Mediated Clearance of AAV-Encoded Antigen in the Liver and Tolerance Is Dependent on the Vector Dose.* Mol Ther, 2017. **25**(4): p. 880-891.

97. Boutin, S., et al., *Prevalence of serum IgG and neutralizing factors against adeno-associated virus (AAV) types 1, 2, 5, 6, 8, and 9 in the healthy population: implications for gene therapy using AAV vectors.* Hum Gene Ther, 2010. **21**(6): p. 704-12.

98. Falese, L., et al., *Strategy to detect pre-existing immunity to AAV gene therapy.* Gene Ther, 2017. **24**(12): p. 768-778.

99. Meliani, A., et al., *Determination of anti-adeno-associated virus vector neutralizing antibody titer with an in vitro reporter system.* Hum Gene Ther Methods, 2015. **26**(2): p. 45-53.

100. Kruzik, A., et al., *The detection of biologically relevant low-titer neutralizing antibodies against AAV require sensitive in vitro assays.* Hum Gene Ther Methods, 2019.

101. Mendell, J.R., et al., *Dystrophin immunity in Duchenne's muscular dystrophy.* N Engl J Med, 2010. **363**(15): p. 1429-37.

Chapter 3: Engineered Viral Vectors for Functional Interrogation, Deconvolution, and Manipulation of Neural Circuits

This chapter is adapted from a manuscript published as

S. Sun, D.V. Schaffer. Engineered Viral Vectors for Functional Interrogation, Deconvolution, and Manipulation of Neural Circuits. *Current Opinion in Neurobiology*, 2018. **50**: p.163-170.

3.1 Introduction

Since herpes simplex virus was first used as a polysynaptic viral tracer for deciphering the input and output connectivity of the brain[1], a range of neurotropic viral vectors has been developed to achieve controlled unidirectional tagging of first-order presynaptic or postsynaptic neurons (Figure 1). These transsynaptic viral tracers have enabled pathway-specific gene delivery and forward screening of structural and functional connections of target neuronal populations with unprecedented precision, though their virulence often limits studies to shorter terms[2]. In parallel, genetically encoded calcium and voltage indicators[3] and opsins[4], coupled with transgenic technologies to achieve cell-type-specific expression, have been exploited to monitor and modulate neural activity with high temporal and spatial control. However, the long-term study and manipulation of networks defined by their connectivity, rather than their genetic profile, have been thwarted by the challenge of developing viral vectors that mediate persistent and robust expression with minimal neurotoxicity.

Vectors based on adeno-associated virus (AAV) – which offer a favorable safety profile, low immunogenicity, and ability to transduce and persist as stable episomes in post-mitotic cells of the CNS – have emerged as safe and increasingly promising technologies for treating neurological disorders, yielding encouraging results in Phase I/II clinical trials for Parkinson's, Batten's, and Canavan's disease after a single vector administration[5, 6]. Furthermore, AAV capsid proteins are encoded by a single *cap* gene (2.2 kb), which the field has progressively shown can be modified to manipulate the virus's tropism and other infectious properties[7] (Figure 2a), leading to a strong interest in developing targeted AAV vectors for both enhanced therapeutics and long-term *in vivo* evaluation of neural circuits. AAV structural analyses[8-11], coupled with an understanding of mechanisms that determine biodistribution and delivery efficiency[12-14], can aid in the rational design of enhanced AAV vectors. In contrast, directed evolution is an approach that can achieve improvements in AAV vector performance without necessitating pre-existing structural and mechanistic knowledge of virus-cell interactions[15-20]. This technique involves genetic diversification of the *cap* gene[17, 21-26] (Figure 2b,c) and iterative rounds of phenotypic selection to drive convergence toward the fittest clones (Figure 2d).

This review focuses on strategies for engineering AAV variants with enhanced neurotropism and transport capabilities in the CNS. The resulting vectors, summarized in Table 1, can be coupled with a range of transgenes and gene regulatory elements, to visualize individual cells, track their activity, and direct their behavior.

3.2 Improving monosynaptic input mapping of neural circuits

Prior to the adoption of AAV for circuit mapping, replication-competent viral vectors capable of crossing synapses were widely employed for neuroanatomical studies. Glycoprotein (G)-deleted rabies virus (RV*dG*), pseudotyped with EnvA, was engineered to infect target cell types and restrict viral spread to direct presynaptic inputs[27]. Since the creation of RV*dG*, complementation with an optimized glycoprotein (oG) was found to improve efficiency of monosynaptic retrograde labeling by 20-fold[28]. Furthermore, another strain of rabies (RABV) called CVS-N2c was discovered to exhibit enhanced neurotropism and retrograde transfer with reduced immunogenicity, compared to the traditional SAD-B19 strain[29]. Vesicular stomatitis virus (VSV) can also be endowed with retrograde or anterograde polysynaptic tracing ability by pseudotyping with RABV-G or VSV-G, respectively, and made monosynaptic by supplying the G proteins *in trans*[30]. An analogous replication-conditional system has been devised for anterograde trafficking of herpes simplex virus (HSV), where thymidine kinase complementation rescues viral replication in the infected cell and mediates expression in monosynaptically connected output populations[31].

These engineered transsynaptic viral tracers have offered more complete illumination of neurons and their dendritic arbors than their non-viral counterparts and thus enabled many advances; however, the rapid onset of neurotoxicity associated with their infection and replication hamper their application to physiological and behavior studies. That said, the recent development of a self-inactivating RVdG (SiR) that leaves a permanent genetic signature shortly before being transcriptionally silenced has extended the temporal window for optical imaging and functional interventions from 5-17 days to several months post-infection[32]. In contrast, lentiviruses[33] and AAVs[34] enable persistent, non-toxic gene expression and can provide local delivery of functional tools to analyze primarily efferent pathways originating from the injection site and to manipulate the biology of targeted neurons[2]. Furthermore, lentiviruses pseudotyped with fusion glycoproteins derived from VSV-G and RABV-G have enabled retrograde labeling of infected neuronal populations[35]. The remainder of this review delves into the diverse and growing toolkit of neurotropic AAV variants that merge viral vectors with functional and single-cell approaches.

3.3 Engineered AAV vectors with enhanced target infectivity and specificity

Most natural AAV serotypes preferentially transduce neurons following intraparenchymal administration[36-39], a property that has been harnessed for anatomical studies of bulk population connectivity[34]. AAVs have also been valuable tools for investigating retinal circuits in physiological and pathological conditions through *in vivo* imaging and electrical stimulation[40]. Furthermore, AAV-based delivery of donor templates has recently shown CRISPR-mediated targeted integration via homology-directed repair in mature post-mitotic neurons, enabling monitoring of endogenous proteins[41]. Nonetheless, infectivity and selectivity of natural AAVs for certain functionally relevant neuronal populations are suboptimal[34], and engineering approaches have been applied to the AAV capsid to create new vectors with improved neuronal tropism.

Photoreceptors have been the primary neuronal target in the retina, where intravitreal injections of AAV vectors have been widely pursued to avoid the adverse procedural effects of subretinal injections and engineering approaches have been applied to generate new AAVs that achieve robust pan-retinal distribution from this route[40]. For instance, surface-exposed tyrosines have been found to undergo intracellular phosphorylation and thereby mediate ubiquitylation and vector degradation, and tyrosine (Y) to phenylalanine (F) mutations have, in some cases, increased viral infectivity[42]. For example, intravitreal administration of AAV2-4YF, a quadruple Y-to-F mutant, mediates expression throughout the mouse retina, but only at high doses[43]. An additional tyrosine-to-valine mutation confers a three-fold improvement in photoreceptor transduction[44]. AAV2-4YF is also more effective than its parent AAV2 in reaching the outer layers of the canine retina but at lower efficiency than in the mouse eye[45]. Following these rational design efforts, the first application of *in vivo* selection in the CNS employed administration of large (~10^7) AAV capsid libraries, followed by isolation of rod photoreceptors and recovery of vectors that localized to these target cells[46]. This approach identified an AAV2 variant called 7m8, which mediates highly efficient gene delivery to photoreceptors and retinal pigment epithelium across the entire murine retina from a simple intravitreal injection[46]. While 7m8 was also found to be the most efficient intravitreal vector to date in non-human primate, additional engineering is necessary to provide efficacious vectors in large animal models, illustrating that optimal vectors in one species may not translate to others[46].

AAV has also been engineered for enhanced neuronal tropism and spread in the brain. For example, tyrosine mutations on the AAV2 capsid improve neuronal transduction in the rat brain[47]. Altering known binding footprints of natural AAVs can also affect their biodistribution. For instance, ablation of heparan sulfate (HS) binding enhances volumetric spread of AAV2[47], while the HS and galactose binding domains of an AAV2/9 chimeric vector named AAV2g9 enable it to preferentially infect neurons over glia after intracerebroventricular (ICV) administration[48]. Fortuitously, a vector devised to present a polyalanine peptide as a control in one study[49], termed AAV-AS, yielded 15-fold greater transduction of the mouse brain from systemic administration relative to AAV9, with extensive neuronal delivery to the motor cortex, striatum, and spinal cord[50]. Interestingly, AAV-AS demonstrated significantly reduced neuronal infectivity in certain brain regions (e.g. thalamus), suggesting that it utilizes cell surface receptor(s) with varying expression levels across the CNS. Finally, a randomly shuffled chimeric vector called AAV-B1 delivered 8.6-fold more vector genome copies to the brain than AAV9 upon intravenous administration, facilitating higher neuronal transduction in the motor cortex, thalamus, and spinal cord[51].

Astrocytes have also become increasingly attractive targets for functional investigation, due to their emerging roles in instructing synapse formation and maturation during neural circuit development[52]. Directed evolution was conducted to engineer AAVs that, in contrast to natural serotypes, preferentially infect glia over astrocytes, and the resulting chimeric variants ShH19 and ShH13 both exhibited a pronounced shift in tropism toward astrocytes in the rat striatum[21]. Specifically, ShH19 transduces 5.5-fold more astrocytes than parental AAV2[21]. ShH13 also robustly infects Muller cells after subretinal injection, suggesting a common mechanism for enhanced entry into glial cells[21]. Interestingly, an AAV6 variant called ShH10 demonstrates >90% selectivity for rat Muller glia upon intravitreal injection and >60% greater infectivity than AAV2[21, 53]. Oligodendrocytes have also garnered interest in light of recent findings that they dynamically shape the myelin profile of neural networks[54]. An interesting variant generated from random capsid

shuffling and directed evolution, Olig001, exhibits >95% selectivity for rat oligodendrocytes, which markedly distinguishes it from most natural AAVs[55].

In addition to genetic manipulation of post-mitotic cells in the CNS, interest in exploring and improving the regenerative capacity of the adult mammalian brain has fueled the generation of novel AAV vectors with improved *in vivo* delivery to adult neural stem cells (NSCs). In the subgranular zone (SGZ), a mutant of AAV2 that was evolved on cultured adult rat hippocampal NSCs named AAV r3.45[18] mediates up to five-fold more selective and three-fold more efficient transduction of NSCs than AAV2, 4, and 6 when administered directly into the rat dentate gyrus[56]. Similar properties were observed in the mouse hippocampus. To engineer an AAV variant for the subventricular zone (SVZ), a Cre-dependent strategy that enables recovery of clones that transduce a target cell population, in this case NSCs, was used to isolate a chimeric AAV vector called SCH9 that infects 60% of NSCs in both hemispheres after a unilateral ICV injection[57]. These engineered vectors can potentially be exploited to gain a deeper understanding of the contextual cues and regulatory mechanisms that govern adult NSC fate decisions, as well as their contributions to adult brain function.

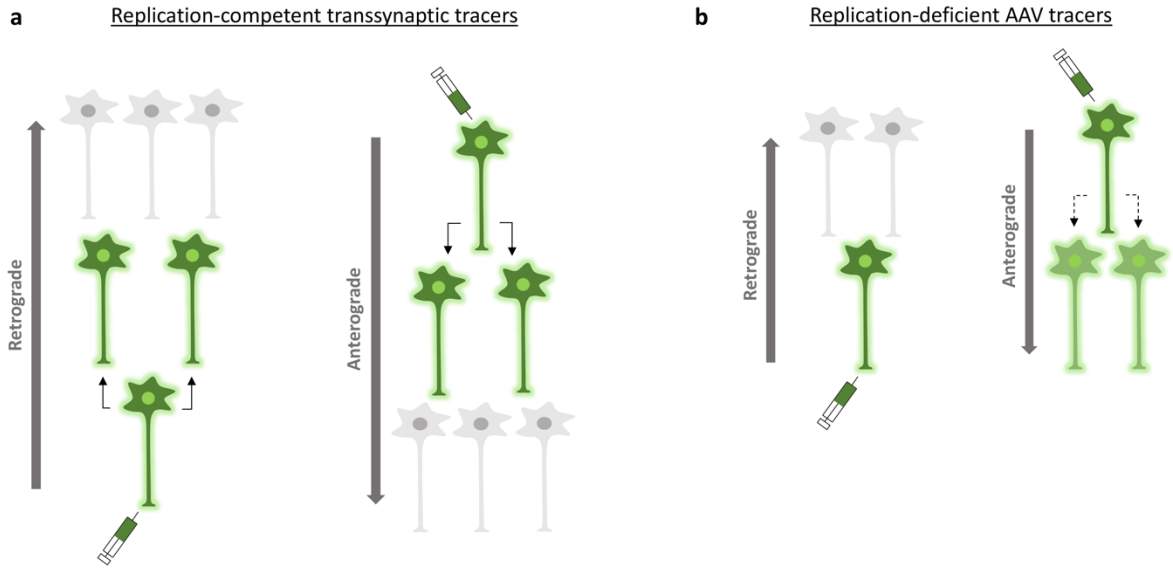


Figure 3.1. Retrograde and anterograde transport of viral tracers.

A) Trans-synaptic labeling using replication-competent viral vectors. Co-delivery of viral tracers and the glycoprotein necessary for replication labels infected cells and their direct presynaptic inputs or postsynaptic outputs through viral transmission.

B) Axonal transport of replication-deficient AAV vectors. AAV-mediated retrograde access can mediate long-term transgene expression in projection neurons. Putatively, anterograde transport of AAV vectors can also occur when some intact virions escape the initial infected neuron and transduce first-order downstream populations, but high titers and catalytically active cargos are necessary to detect trans-synaptic labeling.

3.4 Evolved AAV vectors for functional circuit interrogation

Delivery of sensors and effectors to projection neurons through retrograde transport enables investigation of long-range connections that link circuit modules and coordinate the interplay between large neural networks[58]. AAV was first shown to enter axonal terminals and shuttle its payload to cell nuclei in mice[59] and more recently, to a greater extent, in non-human primates[60-63]. However, the utility of natural AAVs for circuit interrogation in the murine brain is limited by the inefficiency of their transport[59]. AAV libraries were injected into downstream targets of projection neurons, and viral genomes that were retrogradely transported to the cell bodies were recovered. A resulting vector, AAV2-retro, offered up to two orders of magnitude greater retrograde access in the corticopontine tract[64], and this high efficiency extended to numerous other pathways.

AAV-mediated retrograde transport to lower motor neurons from intramuscular administration has also been explored to access specific regions of the spinal cord, but this approach is again hampered by a low transport efficiency of less than 1%[65]. Tet1, a peptide with high affinity for the tetanus toxin GT1b receptor that undergoes retrograde transport in the spinal cord[66], was grafted onto AAV1 to target the capsid for enhanced axonal terminal binding and uptake[67]. Indeed, this rationally engineered vector exhibited a four-fold enhancement over AAV1 in DRG explants.

AAV's natural ability to undergo anterograde transport primarily in non-human primates[61, 68-70], coupled with its minimal neurotoxicity, has sparked exploration of its use as a trans-synaptic tracer. AAV1 was recently established as a superior anterograde tracer compared to other serotypes in the corticocollicular pathways of the murine brain, but required high titers and Cre-mediated amplification to unlock robust transgene expression in up to 40% of postsynaptic V1-recipient neurons[71]. In contrast with replication-competent transsynaptic tracers that spread uncontrollably across multiple serial synapses, AAV1 only tags first-order downstream structures, providing more precise labeling of neural circuits. This attractive feature compels the development of modified AAV vectors with enhanced transneuronal transduction, and a more comprehensive mechanistic understanding of anterograde transport[72, 73] may benefit these engineering efforts.

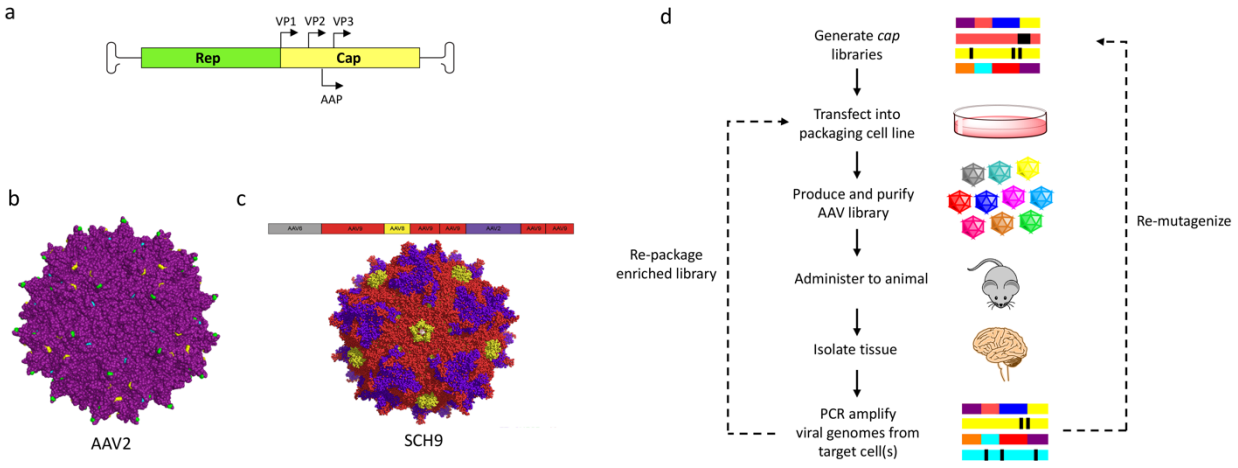


Figure 3.2. Directed evolution of AAV vectors.

A) Schematic of the native AAV genome. The *rep* ORF encodes non-structural proteins that are necessary for viral replication, transcriptional regulation, and virion assembly. The *cap* ORF encodes three structural proteins (VP1, VP2, and VP3) that assemble into a 60-mer viral capsid with the aid of the assembly-activating protein (AAP) encoded by an alternate ORF. The *cap* gene can be modified to manipulate the virus's tropism and infectious properties.

B) Representative target sites for mutagenesis on the capsid surface of AAV2. Peptide ligands can be incorporated at R588 (yellow), a residue near the three-fold protrusions on the capsid surface. Alternate insertion sites include the N-terminus of VP2 (not shown) and residue G453 (green) located on the highest three-fold spike. Similar positions on other AAV serotypes can also be targeted for mutagenesis. Mutation of surface exposed tyrosines, such as Y272 (cyan) on the AAV2 capsid, can yield more potent vectors. Image was rendered in PyMOL using the AAV2 crystal structure (PDB ID: 1LP3).

C) 3D structure of SCH9, an example of a chimeric AAV vector created by recombining *cap* genes from natural AAV serotypes. Each parent serotype is represented by a different color (AAV9 red, AAV8 yellow, AAV2 purple) in a schematic of the *cap* gene and mapped onto the AAV9 crystal structure (PDB ID: 3UX1) in PyMOL.

D) Paradigm for directed evolution of AAV vectors. The *cap* gene is diversified to generate large *cap* libraries that are transfected into a packaging cell line to produce viral particles, such that each capsid variant surrounds the *cap* gene encoding it. Purified capsid libraries are administered to an animal, and a functional selective pressure is imposed by harvesting the tissue of interest. Viral genomes are amplified by PCR to recover *cap* variants that localized to the target cell(s). Variants are re-packaged into virions and enriched through iterative rounds of selection. Alternatively, recovered *cap* genes can be further mutagenized.

3.5 Engineered AAV variants for sparse neuronal labeling

To further dissect the architecture of neural networks and the mechanisms that govern their development, function, and plasticity, deconvolution of complex circuits has primarily been pursued with sparse, stochastic labeling to distinguish individual cells[74]. AAV-mediated delivery of multicolor expression cassettes, such as Brainbow[75], has conferred ease of genetic manipulation, spatial and temporal control, and applicability to larger species, but non-uniform vector distribution from direct intracranial administration often renders cellular tracing and functional analysis of a target region challenging[76]. ICV administration may offer a route for stochastically labeling regions that border the cerebrospinal fluid[57]. Alternatively, systemic AAV administration to access the brain diffusely via its vasculature may reduce variability in vector copies delivered per cell, and vector dosage can be modulated to control labeling density[77].

The natural serotype AAV9 is capable of crossing the formidable blood-brain barrier (BBB) in the adult mammalian brain, resulting in widespread astrocytic transduction but sparse neuronal delivery[78]. AAVrh8 and AAVrh10 were subsequently shown to display similar biodistribution profiles in the CNS[79]. Moreover, mutation of surface-exposed tyrosine residues on AAV9 further increases infectivity[80]. Directed evolution has also recently been employed, in combination with a Cre-dependent selection strategy, to isolate AAVs capable of crossing the BBB[77, 81]. With this approach, AAV-PHP.B, an AAV9 variant containing a 7mer peptide insertion, was found to transduce the majority of astrocytes and neurons and a moderate fraction of oligodendrocytes in multiple regions after intravenous administration, resulting in at least 40-fold greater gene transfer throughout the CNS than AAV9 in mouse. AAV-PHP.A, another clone that

emerged from the same selection, exhibited weaker but more selective astrocytic transduction than AAV-PHP.B[81].

Further evolution on promising engineered vectors has produced new third-generation AAV capsids that retain their parental tropism but show even more enhanced infectivity of a target cell population, mimicking the natural evolutionary search for local maxima in the protein fitness landscape[82]. AAV PHP.eB was generated through additional diversification of the heptamer and flanking amino acids of AAV PHP.B[81], followed by two more rounds of *in vivo* selection in mice[77]. At a relatively low systemic dose of $1x10^{11}$ viral genomes per mouse, AAV PHP.eB facilitated delivery to 69% of cortical neurons, 55% of striatal neurons, and a similar fraction of glia in these two regions compared to AAV PHP.B. This work also identified AAV PHP.S, which provided a two-fold increase in transduction of dorsal root ganglion (DRG) neurons over AAV9 and strong localization to cardiac and enteric ganglia. These potent vectors, which can deliver multiple genomes per cell, can potentially be harnessed for sparse multicolor labeling with high transgene expression to enable morphological studies at the single cell level. Alternatively, they can be utilized to record or perturb the activity of individual cells.

Engineered AAV vectors and their tropism. PR, photoreceptors; RPE, retinal pigment epithelium; ICV, intracerebroventricular; SVZ, subventricular zone; NSC, neural stem cell; DRG, dorsal root ganglia; M, mouse; R, rat; C, canine; F, feline; NHP, non-human primate

Variant	Route of administration	Tropism	Species	Additional notes	Reference
AAV2(4YF)	Intravitreal	Inner retina > PR > RPE	M, C	High doses required	[45,43]
AAV2(4YF + T-V)	Intravitreal	Inner retina > PR > RPE	M	High doses required	[44]
7m8	Intravitreal	Pan-retinal	M, NHP		[46]
AAV2 T2 3Y+*dH*	Intracranial	Neurons	R	Improved spread	[47]
AAV2g9	ICV (neonates)	Neurons > astrocytes	M		[48]
ShH19	Intracranial	Enhanced astrocyte transduction	R		[21]
ShH13	Intracranial, subretinal	Enhanced astrocyte and Muller glia transduction	R		[21]
ShH10	Intravitreal	Muller glia >> inner retina	R		[53]
Olig001	Intracranial	Oligodendrocytes >> neurons	R		[55•]
AAV r3.45	Intracranial (hippocampus)	Adult NSCs > neurons	M, R		[56]
SCH9	Unilateral ICV	Adult NSCs in SVZ	M	Similar transduction profile in both hemispheres	[57•]
AAV2-retro	Intracranial	Retrograde access to neurons	M		[64••]
AAV1-Tet1	N/A	Retrograde access to DRG neurons	R		[67]
AAV-AS	Intravenous	Neurons, astrocytes, endothelia	M, F	Variable neuronal transduction profile across brain regions	[50]
AAV-B1	Intravenous	Neurons, astrocytes, endothelia	M, F	Variable neuronal transduction profile across brain regions	[51]
AAV-PHP.B	Intravenous	Neurons, astrocytes, oligodendrocytes, endothelia	M	Majority of neurons and astrocytes transduced	[81••]
AAV-PHP.A	Intravenous	Astrocytes	M		[81••]
AAV-PHP.eB	Intravenous	Neurons, astrocytes, oligodendrocytes, endothelia	M	Enhanced neuron and Purkinje cell transduction relative to AAV-PHP.B	[77•]
AAV-PHP.S	Intravenous	DRG, cardiac and enteric ganglia	M		[77•]

Table 3.1. Engineered AAV vectors and their tropism. PR, photoreceptors; RPE, retinal pigment epithelium; ICV, intracerebroventricular; SVZ, subventricular zone; NSC, neural stem cell; DRG, dorsal root ganglia**;** M, mouse; R, rat; C, canine; F, feline; NHP, non-human primate.

3.6 Conclusion

The engineered AAV viruses discussed here represent a diverse collection of vectors that have the potential to revolutionize the investigation of the mammalian connectome and the

underlying functionalities of embedded neural circuits. Their ability to mediate robust expression in functionally relevant cell types of the CNS may also facilitate population-wide gene ablation, repression, or upregulation to generate improved mouse models that mimic postnatal onset of certain diseases. To further expand the available toolkit of AAV vectors, directed evolution strategies described above can be applied to other cell types of interest (e.g. oligodendrocytes), as well as other species to address the ongoing challenge that vectors engineered in mouse rarely translate to large animals (including humans). Moreover, design of paradigms that select against transduction of antigen-presenting cells (APCs) in the brain can prevent cytotoxic immune responses and permit more extensive studies in non-human primates[83]. Advances in promoter engineering[84] will also aid in restricting expression to defined populations. Lastly, increasing AAV's packaging capacity, potentially via modulating residues on the capsid lumen that interact with its genome[85], will extend its applicability to larger transgenes and eliminate the need for less efficacious multi-vector systems. Such engineering approaches will further build momentum for AAV's application to both basic biological investigation and therapeutic translation.

3.7 Acknowledgements

SS was supported by the American Heart Association Predoctoral Fellowship, NIH R21 EB021572-01, and NIH R01 EY022975. We also thank Alla Y. Karpova for providing a critical read of the review.

3.8 References

1. Norgren, R.B., Jr. and M.N. Lehman, *Herpes simplex virus as a transneuronal tracer.* Neurosci Biobehav Rev, 1998. **22**(6): p. 695-708.
2. Nassi, J.J., et al., *Neuroanatomy goes viral!* Front Neuroanat, 2015. **9**: p. 80.
3. Lin, M.Z. and M.J. Schnitzer, *Genetically encoded indicators of neuronal activity.* Nat Neurosci, 2016. **19**(9): p. 1142-53.
4. Wietek, J. and M. Prigge, *Enhancing Channelrhodopsins: An Overview.* Methods Mol Biol, 2016. **1408**: p. 141-65.
5. Ojala, D.S., D.P. Amara, and D.V. Schaffer, *Adeno-associated virus vectors and neurological gene therapy.* Neuroscientist, 2015. **21**(1): p. 84-98.
6. Hocquemiller, M., et al., *Adeno-Associated Virus-Based Gene Therapy for CNS Diseases.* Hum Gene Ther, 2016. **27**(7): p. 478-96.
7. Kotterman, M.A. and D.V. Schaffer, *Engineering adeno-associated viruses for clinical gene therapy.* Nat Rev Genet, 2014. **15**(7): p. 445-51.
8. Mikals, K., et al., *The structure of AAVrh32.33, a novel gene delivery vector.* J Struct Biol, 2014. **186**(2): p. 308-17.
9. Walters, R.W., et al., *Structure of adeno-associated virus serotype 5.* J Virol, 2004. **78**(7): p. 3361-71.

10. Nam, H.J., et al., *Structure of adeno-associated virus serotype 8, a gene therapy vector.* J Virol, 2007. **81**(22): p. 12260-71.

11. Padron, E., et al., *Structure of adeno-associated virus type 4.* J Virol, 2005. **79**(8): p. 5047-58.

12. Huang, L.Y., et al., *Characterization of the Adeno-Associated Virus 1 and 6 Sialic Acid Binding Site.* J Virol, 2016. **90**(11): p. 5219-30.

13. Nam, H.J., et al., *Structural studies of adeno-associated virus serotype 8 capsid transitions associated with endosomal trafficking.* J Virol, 2011. **85**(22): p. 11791-9.

14. Venkatakrishnan, B., et al., *Structure and dynamics of adeno-associated virus serotype 1 VP1-unique N-terminal domain and its role in capsid trafficking.* J Virol, 2013. **87**(9): p. 4974-84.

15. Asuri, P., et al., *Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells.* Mol Ther, 2012. **20**(2): p. 329-38.

16. Maheshri, N., et al., *Directed evolution of adeno-associated virus yields enhanced gene delivery vectors.* Nat Biotechnol, 2006. **24**(2): p. 198-204.

17. Koerber, J.T., J.H. Jang, and D.V. Schaffer, *DNA shuffling of adeno-associated virus yields functionally diverse viral progeny.* Mol Ther, 2008. **16**(10): p. 1703-9.

18. Jang, J.H., et al., *An evolved adeno-associated viral variant enhances gene delivery and gene targeting in neural stem cells.* Mol Ther, 2011. **19**(4): p. 667-75.

19. Yang, L., et al., *A myocardium tropic adeno-associated virus (AAV) evolved by DNA shuffling and in vivo selection.* Proc Natl Acad Sci U S A, 2009. **106**(10): p. 3946-51.

20. Lisowski, L., et al., *Selection and evaluation of clinically relevant AAV variants in a xenograft liver model.* Nature, 2014. **506**(7488): p. 382-6.

21. Koerber, J.T., et al., *Molecular evolution of adeno-associated virus for enhanced glial gene delivery.* Mol Ther, 2009. **17**(12): p. 2088-95.

22. Koerber, J.T., et al., *Construction of diverse adeno-associated viral libraries for directed evolution of enhanced gene delivery vehicles.* Nat Protoc, 2006. **1**(2): p. 701-6.

23. Muller, O.J., et al., *Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors.* Nat Biotechnol, 2003. **21**(9): p. 1040-6.

24. Silberg, J.J., J.B. Endelman, and F.H. Arnold, *SCHEMA-guided protein recombination.* Methods Enzymol, 2004. **388**: p. 35-42.

25. Koerber, J.T. and D.V. Schaffer, *Transposon-based mutagenesis generates diverse adeno-associated viral libraries with novel gene delivery properties.* Methods Mol Biol, 2008. **434**: p. 161-70.

26. Santiago-Ortiz, J., et al., *AAV ancestral reconstruction library enables selection of broadly infectious viral variants.* Gene Ther, 2015. **22**(12): p. 934-46.

27. Callaway, E.M. and L. Luo, *Monosynaptic Circuit Tracing with Glycoprotein-Deleted Rabies Viruses.* J Neurosci, 2015. **35**(24): p. 8979-85.

28. Kim, E.J., et al., *Improved Monosynaptic Neural Circuit Tracing Using Engineered Rabies Virus Glycoproteins.* Cell Rep, 2016.

29. Reardon, T.R., et al., *Rabies Virus CVS-N2c(DeltaG) Strain Enhances Retrograde Synaptic Transfer and Neuronal Viability.* Neuron, 2016. **89**(4): p. 711-24.

30. Beier, K.T., et al., *Vesicular stomatitis virus with the rabies virus glycoprotein directs retrograde transsynaptic transport among neurons in vivo.* Front Neural Circuits, 2013. **7**: p. 11.

31. Zeng, W.B., et al., *Anterograde monosynaptic transneuronal tracers derived from herpes simplex virus 1 strain H129.* Mol Neurodegener, 2017. **12**(1): p. 38.

32. Ciabatti, E., et al., *Life-Long Genetic and Functional Access to Neural Circuits Using Self-Inactivating Rabies Virus.* Cell, 2017. **170**(2): p. 382-392 e14.

33. Parr-Brownlie, L.C., et al., *Lentiviral vectors as tools to understand central nervous system biology in mammalian model organisms.* Front Mol Neurosci, 2015. **8**: p. 14.

34. Betley, J.N. and S.M. Sternson, *Adeno-associated viral vectors for mapping, monitoring, and manipulating neural circuits.* Hum Gene Ther, 2011. **22**(6): p. 669-77.

35. Kobayashi, K., et al., *Altering Entry Site Preference of Lentiviral Vectors into Neuronal Cells by Pseudotyping with Envelope Glycoproteins.* Methods Mol Biol, 2016. **1382**: p. 175-86.

36. Davidson, B.L., et al., *Recombinant adeno-associated virus type 2, 4, and 5 vectors: transduction of variant cell types and regions in the mammalian central nervous system.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3428-32.

37. Cearley, C.N. and J.H. Wolfe, *Transduction characteristics of adeno-associated virus vectors expressing cap serotypes 7, 8, 9, and Rh10 in the mouse brain.* Mol Ther, 2006. **13**(3): p. 528-37.

38. Aschauer, D.F., S. Kreuz, and S. Rumpel, *Analysis of transduction efficiency, tropism and axonal transport of AAV serotypes 1, 2, 5, 6, 8 and 9 in the mouse brain.* PLoS One, 2013. **8**(9): p. e76310.

39. Markakis, E.A., et al., *Comparative transduction efficiency of AAV vector serotypes 1-6 in the substantia nigra and striatum of the primate brain.* Mol Ther, 2010. **18**(3): p. 588-93.

40. Schon, C., M. Biel, and S. Michalakis, *Retinal gene delivery by adeno-associated virus (AAV) vectors: Strategies and applications.* Eur J Pharm Biopharm, 2015. **95**(Pt B): p. 343-52.

41. Nishiyama, J., T. Mikuni, and R. Yasuda, *Virus-Mediated Genome Editing via Homology-Directed Repair in Mitotic and Postmitotic Cells in Mammalian Brain.* Neuron, 2017.

42. Qiao, C., et al., *Adeno-associated virus serotype 6 capsid tyrosine-to-phenylalanine mutations improve gene transfer to skeletal muscle.* Hum Gene Ther, 2010. **21**(10): p. 1343-8.

43. Petrs-Silva, H., et al., *Novel properties of tyrosine-mutant AAV2 vectors in the mouse retina.* Mol Ther, 2011. **19**(2): p. 293-301.

44. Kay, C.N., et al., *Targeting photoreceptors via intravitreal delivery using novel, capsid-mutated AAV vectors.* PLoS One, 2013. **8**(4): p. e62097.

45. Mowat, F.M., et al., *Tyrosine capsid-mutant AAV vectors for gene delivery to the canine retina from a subretinal or intravitreal approach.* Gene Ther, 2014. **21**(1): p. 96-105.

46. Dalkara, D., et al., *In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous.* Sci Transl Med, 2013. **5**(189): p. 189ra76.

47. Kanaan, N.M., et al., *Rationally Engineered AAV Capsids Improve Transduction and Volumetric Spread in the CNS.* Molecular Therapy - Nucleic Acids, 2017. **8**: p. 184-197.

48. Murlidharan, G., et al., *CNS-restricted Transduction and CRISPR/Cas9-mediated Gene Deletion with an Engineered AAV Vector.* Mol Ther Nucleic Acids, 2016. **5**(7): p. e338.

49. Pulicherla, N., et al., *Engineering liver-detargeted AAV9 vectors for cardiac and musculoskeletal gene transfer.* Mol Ther, 2011. **19**(6): p. 1070-8.

50. Choudhury, S.R., et al., *Widespread Central Nervous System Gene Transfer and Silencing After Systemic Delivery of Novel AAV-AS Vector.* Mol Ther, 2016. **24**(4): p. 726-35.

51. Choudhury, S.R., et al., *In Vivo Selection Yields AAV-B1 Capsid for Central Nervous System and Muscle Gene Therapy.* Mol Ther, 2016. **24**(7): p. 1247-57.

52. Clarke, L.E. and B.A. Barres, *Emerging roles of astrocytes in neural circuit development.* Nat Rev Neurosci, 2013. **14**(5): p. 311-21.

53. Klimczak, R.R., et al., *A novel adeno-associated viral variant for efficient and selective intravitreal transduction of rat Muller cells.* PLoS One, 2009. **4**(10): p. e7467.

54. Michalski, J.P. and R. Kothary, *Oligodendrocytes in a Nutshell.* Front Cell Neurosci, 2015. **9**: p. 340.

55. Powell, S.K., et al., *Characterization of a novel adeno-associated viral vector with preferential oligodendrocyte tropism.* Gene Ther, 2016. **23**(11): p. 807-814.

56. Kotterman, M.A., T. Vazin, and D.V. Schaffer, *Enhanced selective gene delivery to neural stem cells in vivo by an adeno-associated viral variant.* Development, 2015. **142**(10): p. 1885-92.

57. Ojala, D.S., et al., *In Vivo Selection of a Computationally Designed SCHEMA AAV Library Yields a Novel Variant for Infection of Adult Neural Stem Cells in the SVZ.* Mol Ther, 2017.

58. Ginger, M., et al., *Revealing the secrets of neuronal circuits with recombinant rabies virus technology.* Front Neural Circuits, 2013. **7**: p. 2.

59. Kaspar, B.K., et al., *Targeted retrograde gene delivery for neuronal protection.* Mol Ther, 2002. **5**(1): p. 50-6.

60. San Sebastian, W., et al., *Adeno-associated virus type 6 is retrogradely transported in the non-human primate brain.* Gene Ther, 2013. **20**(12): p. 1178-83.

61. Green, F., et al., *Axonal transport of AAV9 in nonhuman primate brain.* Gene Ther, 2016. **23**(6): p. 520-6.

62. Salegio, E.A., et al., *Axonal transport of adeno-associated viral vectors is serotype-dependent.* Gene Ther, 2013. **20**(3): p. 348-52.

63. Kells, A.P., et al., *Efficient gene therapy-based method for the delivery of therapeutics to primate cortex.* Proc Natl Acad Sci U S A, 2009. **106**(7): p. 2407-11.

64. Tervo, D.G., et al., *A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons.* Neuron, 2016. **92**(2): p. 372-382.

65. Kaspar, B.K., et al., *Retrograde viral delivery of IGF-1 prolongs survival in a mouse ALS model.* Science, 2003. **301**(5634): p. 839-42.

66. Federici, T., et al., *A means for targeting therapeutics to peripheral nervous system neurons with axonal damage.* Neurosurgery, 2007. **60**(5): p. 911-8; discussion 911-8.

67. Davis, A.S., et al., *Rational design and engineering of a modified adeno-associated virus (AAV1)-based vector system for enhanced retrograde gene delivery.* Neurosurgery, 2015. **76**(2): p. 216-25; discussion 225.

68. Ciesielska, A., et al., *Anterograde axonal transport of AAV2-GDNF in rat basal ganglia.* Mol Ther, 2011. **19**(5): p. 922-7.

69. Kells, A.P., J. Forsayeth, and K.S. Bankiewicz, *Glial-derived neurotrophic factor gene transfer for Parkinson's disease: anterograde distribution of AAV2 vectors in the primate brain.* Neurobiol Dis, 2012. **48**(2): p. 228-35.

70. Hutson, T.H., C. Kathe, and L.D. Moon, *Trans-neuronal transduction of spinal neurons following cortical injection and anterograde axonal transport of a bicistronic AAV1 vector.* Gene Ther, 2016. **23**(2): p. 231-6.

71. Zingg, B., et al., *AAV-Mediated Anterograde Transsynaptic Tagging: Mapping Corticocollicular Input-Defined Neural Pathways for Defense Behaviors.* Neuron, 2017. **93**(1): p. 33-47.

72. Castle, M.J., et al., *Long-distance axonal transport of AAV9 is driven by dynein and kinesin-2 and is trafficked in a highly motile Rab7-positive compartment.* Mol Ther, 2014. **22**(3): p. 554-66.

73. Castle, M.J., et al., *Adeno-associated virus serotypes 1, 8, and 9 share conserved mechanisms for anterograde and retrograde axonal transport.* Hum Gene Ther, 2014. **25**(8): p. 705-20.

74. Jefferis, G.S. and J. Livet, *Sparse and combinatorial neuron labelling.* Curr Opin Neurobiol, 2012. **22**(1): p. 101-10.

75. Richier, B. and I. Salecker, *Versatile genetic paintbrushes: Brainbow technologies.* Wiley Interdiscip Rev Dev Biol, 2015. **4**(2): p. 161-80.

76. Cai, D., et al., *Improved tools for the Brainbow toolbox.* Nat Methods, 2013. **10**(6): p. 540-7.

77. Chan, K.Y., et al., *Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems.* Nat Neurosci, 2017.

78. Foust, K.D., et al., *Intravascular AAV9 preferentially targets neonatal neurons and adult astrocytes.* Nat Biotechnol, 2009. **27**(1): p. 59-65.

79. Yang, B., et al., *Global CNS transduction of adult mice by intravenously delivered rAAVrh.8 and rAAVrh.10 and nonhuman primates by rAAVrh.10.* Mol Ther, 2014. **22**(7): p. 1299-309.

80. Iida, A., et al., *Systemic delivery of tyrosine-mutant AAV vectors results in robust transduction of neurons in adult mice.* Biomed Res Int, 2013. **2013**: p. 974819.

81. Deverman, B.E., et al., *Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain.* Nat Biotechnol, 2016. **34**(2): p. 204-9.

82. Romero, P.A. and F.H. Arnold, *Exploring protein fitness landscapes by directed evolution.* Nat Rev Mol Cell Biol, 2009. **10**(12): p. 866-76.

83. Samaranch, L., et al., *AAV9-mediated expression of a non-self protein in nonhuman primate central nervous system triggers widespread neuroinflammation driven by antigen-presenting cell transduction.* Mol Ther, 2014. **22**(2): p. 329-37.

84. de Leeuw, C.N., et al., *rAAV-compatible MiniPromoters for restricted expression in the brain and eye.* Mol Brain, 2016. **9**(1): p. 52.

85. Grieger, J.C. and R.J. Samulski, *Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps.* J Virol, 2005. **79**(15): p. 9933-44.

Chapter 4: Efficient Generation of Knock-In Mice by CRISPR RNP Electroporation and AAV Donor Infection (CRISPR-READI)

This chapter is adapted from a manuscript published as

S. Sun*, S. Chen*, D. Moonen, C. Lee, A.Y. Lee, D.V. Schaffer†, L. He†. CRISPR-READI: Efficient generation of knock-in mice by CRISPR RNP electroporation and AAV donor infection. *Cell Reports*, 2019. **27**(13): p.3780-3789.

*Indicates co-first authors
†Indicates co-corresponding authors

4.1 Introduction

Genetically modified mice are invaluable assets for investigating mammalian gene function, as well as for modeling human development, physiology, and disease. In particular, knock-in mice harboring large sequence insertions or substitutions are essential for a variety of applications including endogenous gene tagging, conditional gene knockout, site-specific transgene insertion, and gene replacement. While embryonic stem cells (ESCs) engineered by homologous recombination were classically used to generate these mouse models, the rapid adoption of CRISPR/Cas9 technology has provided an attractive alternative—direct microinjection of CRISPR components with a donor template to trigger homology-directed repair (HDR) and produce edited animals in one generation[1–8]. Multiple technical refinements have also been developed to increase the rate of HDR in mouse zygotes through the use of ssDNA, linearized dsDNA, or chemically modified dsDNA donor templates, co-injection of HDR-stimulating compounds, and timed microinjection in 2-cell stage embryos[9–14]. In spite of these improvements, complex editing in mice remains challenging due to the need for microinjection, a costly procedure with a high technical barrier and low throughput[15,16].

Recently, we and others developed electroporation-based methods to deliver CRISPR reagents into zygotes for highly efficient genome engineering; these strategies have garnered popularity due to improved throughput, technical ease, and cost-effectiveness across multiple mammalian species[17–26]. While electroporation-based approaches are highly effective in introducing indel mutations, large deletions, and small insertions or substitutions, the use of short (typically <200 nt) single-stranded oligodeoxynucleotides (ssODNs) as HDR donors renders these techniques unsuitable for editing schemes involving targeted insertion of multi-kilobase sequences. Alternatively, long single-stranded oligodeoxynucleotides (lssODNs) can be applied for more complex editing, and several lssODN synthesis methods, including reverse-transcription, plasmid nicking followed by gel extraction, and selective strand phosphorylation/degradation, have been reported[9–11,18,23,24,26–28]. However, these protocols carry several notable drawbacks, including limited length (typically ≤ 2000 nt), GC-content and other sequence constraints, suboptimal yield, and costly reagents[10,11,26,27,29]. Furthermore, electroporation of lssODNs

typically yields much lower knock-in efficiencies than that of short ssODNs, likely due to inefficient lssODN delivery into zygotes[9,23]. Owing to these restrictions, targeted knock-ins are predominantly performed using microinjection or ESC-based approaches with plasmid or linearized dsDNA as HDR donors.

Here, we developed a strategy for complex mouse genome engineering that leverages the simplicity and throughput of CRISPR electroporation while overcoming the current limitations of large HDR donor delivery. Recombinant adeno-associated virus (rAAV) has emerged as a safe and efficient gene delivery vector with the innate ability to transduce mammalian cells and stimulate gene targeting by promoting homologous recombination[30–34]. In particular, rAAV donors packaged with single-stranded DNA genomes have successfully served as repair templates for HDR upon Cas9-mediated cleavage in mammalian cell lines, enhancing the efficiency of site-specific gene integration by >10-fold relative to plasmid donor nucleofection[35]. More recently, three individual rAAV6 vectors were used to deliver Cas9, sgRNA, and donor components into mouse zygotes to mediate gene targeting, demonstrating targeted knock-in of up to ~700 bp[36]. Of note, AAV's naturally occurring single-stranded DNA genomes can be engineered into a duplexed form, termed self-complementary AAV (scAAV), which can also facilitate efficient gene targeting[37].

In this study, we identified AAV1 as an optimal naturally occurring serotype for highly efficient transduction of mouse zygotes. We then developed CRISPR-READI (CRISPR RNP Electroporation and AAV Donor Infection), a highly efficient method that combines AAV-mediated HDR donor delivery with Cas9/sgRNA RNP electroporation to generate complex genome modifications in mice. Using CRISPR-READI, we successfully inserted a 774 bp fluorescent reporter or a 2.1 kb CreERT2 cassette into the endogenous *Sox2* locus, as well as a 3.3 kb gene expression cassette into the *Rosa26* locus, in both preimplantation stage embryos and in mice. CRISPR-READI permits the use of HDR donors up to 4.9 kb in length[38] and thus enables a broad range of complex genome modifications, including site-specific integration of reporters, Cre-drivers, and expression cassettes. Altogether, CRISPR-READI produces genetically edited animals harboring multi-kilobase modifications with unparalleled efficiency and throughput.

4.2 Results

Specific natural AAV serotypes efficiently transduce mouse zygotes

To characterize the native capacity of rAAV vectors to penetrate the zona pellucida and deliver the viral genome into mouse zygotes, we infected C57BL/6J mouse zygotes with a panel of eight natural AAV serotypes packaged with a self-complementary CMV-eGFP reporter (scAAV-CMV-eGFP) at a dose of 2×10^8 genome copies (GCs) per culture droplet (Fig. 1A; Fig. S1A). Strong fluorescent signal was only detected for serotypes 1 and 6 (scAAV1 and scAAV6), which yielded comparable signal intensity by the 4 to 8-cell stage, suggesting that efficient rAAV-mediated DNA delivery can be achieved in mouse zygotes (Fig. 1A; Fig. S1A). While ~20% of scAAV1- or scAAV6-tranduced embryos exhibited bright eGFP fluorescence, all treated embryos displayed signal above background (Fig. S1A). Although a previous study utilized rAAV6 for mouse editing[36], we selected rAAV1 for knock-in experiments with CRISPR-READI due to the superior viral yield of rAAV1 compared to rAAV6 when produced in HEK293T cells[39].

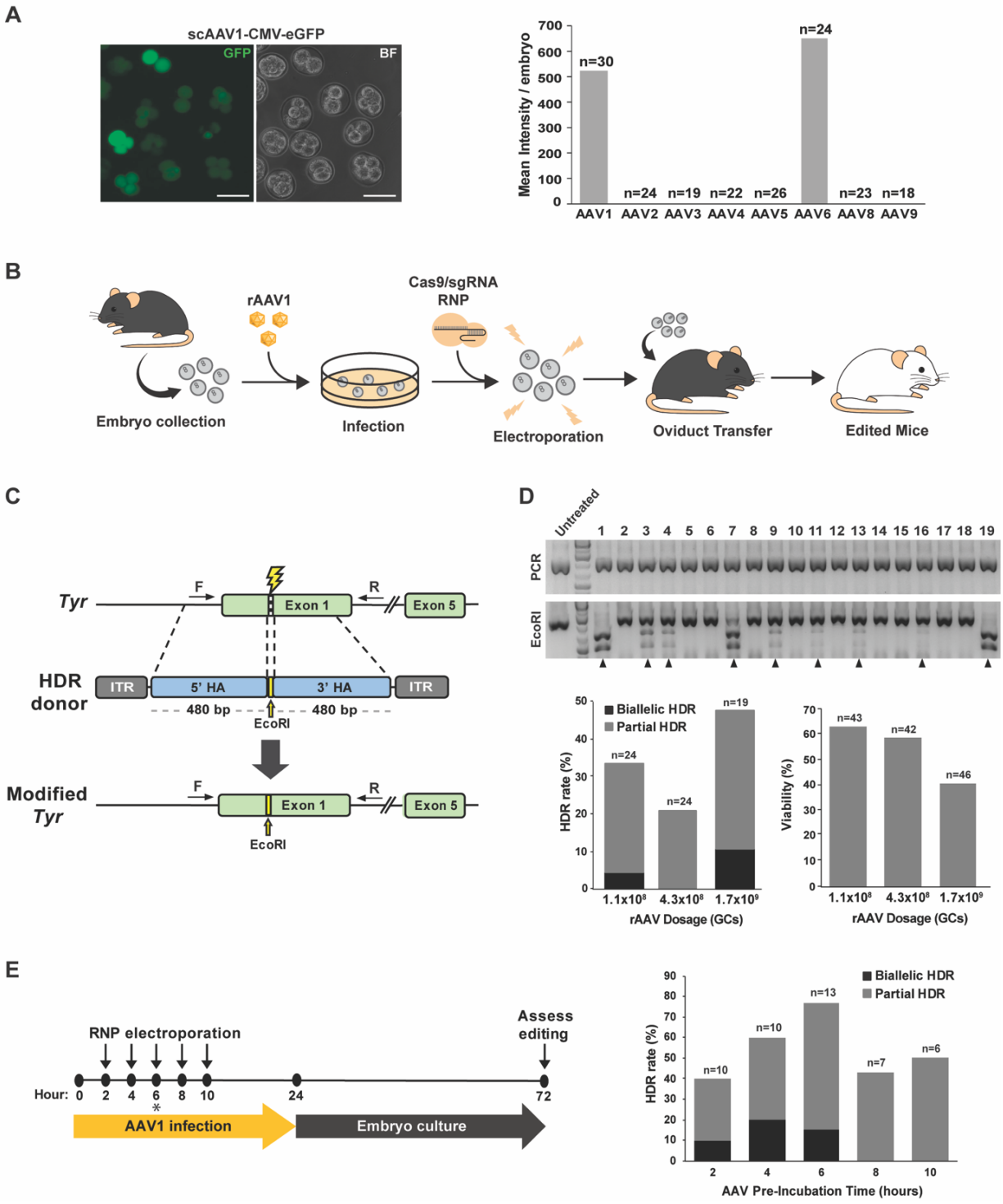

A
scAAV1-CMV-eGFP
GFP
BF
Mean Intensity / embryo
n=30
n=24
n=24
n=19
n=22
n=26
n=23
n=18
AAV1
AAV2
AAV3
AAV4
AAV5
AAV6
AAV8
AAV9
B
rAAV1
Cas9/sgRNA RNP
Embryo collection
Infection
Electroporation
Oviduct Transfer
Edited Mice
C
Tyr
F
R
Exon 1
Exon 5
HDR donor
ITR
5' HA
3' HA
480 bp
480 bp
EcoRI
Modified Tyr
D
Untreated
PCR
EcoRI
Biallelic HDR
Partial HDR
HDR rate (%)
Viability (%)
rAAV Dosage (GCs)
E
RNP electroporation
Assess editing
Hour:
AAV1 infection
Embryo culture
AAV Pre-Incubation Time (hours)

Figure 4.1. CRISPR-READI optimization for efficient HDR editing in mouse embryos.

A) Zygotes were transduced with a panel of AAV serotypes harboring a CMV-eGFP reporter and imaged by fluorescent microscopy 48 hours post-transduction. Representative embryos transduced with scAAV1-CMV-eGFP are shown (left), and mean fluorescence intensity per embryo was quantified for each serotype (right). Scale bars = 50 μm.

B) Cartoon depiction of CRISPR-READI workflow. Embryos are collected from superovulated female mice, transduced with rAAV1 harboring the donor template, electroporated with Cas9/sgRNA RNPs, and implanted into pseudopregnant females to generate edited mice.

C) Schematic of *Tyr* targeting strategy. The scAAV1-Tyr donor creates an EcoRI restriction site in exon 1 of the *Tyr* locus upon HDR editing. ITR: inverted terminal repeat, HA: homology arm, F/R: forward/reverse primers for RFLP analysis.

D) Optimization of rAAV1 dosage for HDR editing. Zygotes were transduced with scAAV1-Tyr at a dose of 1.1×10^8, 4.2×10^8, or 1.7×10^9 GCs, electroporated with RNPs 5 hours post-transduction, and then returned to rAAV1 incubation for another 19 hours. Treated embryos were cultured to the morula stage and genotyped by restriction fragment length polymorphism (RFLP) analysis (shown for dose of 1.7×10^9 GCs). Edited embryos yield 650 bp and 420 bp bands upon EcoRI digestion of the PCR amplicon (top, black arrows). HDR rate was quantified by RFLP analysis for each dose (bottom left), and embryo viability was scored as percentage of cultured embryos that reached the morula stage (bottom right).

E) Optimization of RNP electroporation timing relative to rAAV transduction. Zygotes were transduced with scAAV1-Tyr, electroporated at varying time points post-transduction (2, 4, 6, 8, or 10 hours), and returned to rAAV incubation for a total of 24 hours. Treated embryos were cultured to the morula stage, lysed, and assessed by RFLP analysis (right). 6 hours (*) was identified as the optimal time of RNP electroporation for maximal editing efficiency.

Optimization of CRISPR-READI enables efficient HDR editing in embryos

To generate edited animals with CRISPR-READI, we devised a workflow wherein mouse zygotes are harvested from superovulated females, infected with rAAV1 donors, electroporated with pre-assembled Cas9/sgRNA RNPs, and cultured to the 2-cell stage before transferring to the oviducts of pseudopregnant females (Fig. 1B). In a proof-of-principle CRISPR-READI experiment, we designed a strategy to insert an EcoRI restriction site into *Tyrosinase* (*Tyr*) exon 1 using a ~960 bp donor (including ~480 bp homology arms) that was packaged into an scAAV1 vector (scAAV1-Tyr) (Fig. 1C). As intracellular trafficking, nuclear localization, and capsid uncoating precede AAV-mediated HDR editing[40], we pre-incubated the zygotes for 5 hours with three doses of scAAV1-Tyr (1.1×10^8, 4.3×10^8, and 1.7×10^9 GCs), introduced Cas9/sgRNA RNP by electroporation, and then continued viral incubation for a total of 24 hours (Fig. 1D). Treated embryos were grown for another 48 hours post-transduction until the morula / early blastocyst stage, and editing efficiency was assessed by screening for the engineered EcoRI site using restriction fragment length polymorphism (RFLP) analysis[19]. As anticipated, HDR editing efficiency was dependent on viral titer. The highest AAV dose, 1.7×10^9 GCs, resulted in 48% (9 out of 19) embryos harboring the precise sequence substitution, with 11% (2 out of 19) showing bi-allelic HDR-mediated editing, while the lower doses of 1.1×10^8 and 4.3×10^8 GCs yielded HDR rates of 33% and 21%, respectively (Fig. 1D; Fig. S1B; Table 1). Although increased AAV dosage elicited a higher HDR rate, it was also correlated with a moderate reduction in embryo viability,

as 41% (19 out of 46) embryos developed to the morula stage in culture when treated with 1.7×10^9 GCs, while 63% (25 out of 43) embryos did so with 1.1×10^8 GCs (Fig. 1D; Table 1). Altogether, these results suggest that the optimal AAV dose for CRISPR-READI ranges from 10^8 to 10^9 GCs, and a balance between HDR rate and embryo viability should be considered when determining AAV dosage for each CRISPR-READI experiment.

HDR requires nuclear co-localization of Cas9, sgRNA, and the donor template. As multiple intracellular trafficking steps must occur prior to nuclear import of the AAV genome[40], we reasoned that optimizing the timing of RNP delivery relative to rAAV transduction would enhance HDR editing efficiency. Using the *Tyr* editing scheme described above (Fig. 1C), we pre-incubated zygotes with the rAAV donor vector for 0, 2, 4, 6, 8, or 10 hours prior to RNP electroporation (Fig. 1E). In all conditions, we achieved a minimum of 40% HDR-mediated editing (Fig. 1E; Fig. S1C; Table 1). We found that RNP electroporation at 6 hours after rAAV transduction resulted in maximal editing, with 77% (10 out of 13) assayed embryos harboring the precise sequence modification and 15% (2 out of 13) showing bi-allelic editing. Notably, this exceeds the HDR frequencies we previously achieved by CRISPR-EZ, using electroporation of short ssODNs (46%)[19]. Henceforth, we settled on 6 hours of rAAV pre-incubation prior to RNP electroporation for subsequent CRISPR-READI experiments.

As double-strand breaks induced by Cas9 can be repaired either by non-homologous end joining (NHEJ) and HDR, we next sought to characterize the extent of NHEJ in this system. To quantitatively measure the frequency of HDR versus NHEJ-mediated editing, we treated embryos using scAAV1-Tyr at a dosage of 1.7×10^9 GCs, amplified the edited region by PCR, and clonally analyzed editing events by Sanger sequencing (10 clones from each of 10 embryos). As expected, NHEJ occurs frequently, as 90% (9 out of 10) of embryos carried indels. In comparison, 70% (7 out of 10) harbored the desired HDR-mediated point mutation, most of which exhibited both HDR and NHEJ-mediated editing events (Fig. S1D). HDR frequencies of each edited embryo ranged from 10% to 67%, revealing a degree of editing mosaicism.

AAV vectors can be packaged in either single-stranded (ssAAV) or self-complementary (scAAV) forms, the latter of which bypasses the necessity of second-strand synthesis and in turn enhances transduction efficiency[41]. In a previous study, scAAV vectors were also shown to modestly enhance gene correction in mammalian cells *in vitro* relative to their single-stranded equivalents, possibly due to their increased stability, their two copies of the donor template in opposite polarities, or their ability to undergo double crossover events[37]. Using the same *Tyr* editing scheme, we found that scAAV1-Tyr moderately outperformed ssAAV1-Tyr by ~17% in HDR efficiency, demonstrating that scAAV vectors promote HDR-mediated editing in mouse embryos (Fig. S1E; Table 1). However, it is important to note that scAAV vectors possess half the packaging capacity of ssAAV and thus limit the length of HDR donors to ~2.4 kb[41].

CRISPR-READI enables efficient targeted integration of fluorescent reporters

A common strategy to characterize the expression pattern of a gene is to insert a fluorescent reporter under the control of its endogenous promoter. We designed a 1.7 kb scAAV1 donor vector consisting of a 774 bp P2A-mStrawberry insert flanked by ~480 bp homology arms (scAAV1-Sox2-mStr) to knock in the P2A-mStrawberry cassette immediately downstream of the 3' terminus of the *Sox2* open reading frame (ORF) (Fig. 2A). Mouse zygotes were treated with CRISPR-

READI using scAAV1-Sox2-mStr at three doses (3.2×10^7, 1.3×10^8, and 5.0×10^8 GCs) and cultured to the blastocyst stage. Robust red fluorescence was detected in the inner cell mass (ICM) of a subset of the resulting blastocysts, recapitulating endogenous *Sox2* expression (Fig. 2B). scAAV1 donor at a dose of 1.3×10^8 GCs produced the highest number of correctly targeted embryos, with 33% (6 out of 18) of blastocysts showing ICM-specific fluorescent signal (Fig. 2B; Fig. S2A; Table 1). Since many widely used reporters, such as fluorescent proteins, luciferase, and HaloTag, can be targeted with an HDR donor under 2.4 kb in length, CRISPR-READI using scAAV1 donors can be applied to tag endogenous loci with many common reporter cassettes.

To assess the ability of our method to generate live knock-in mice, we treated zygotes with CRISPR-READI using scAAV1-Sox2-mStr at a dose of 1.3×10^8 GCs and transferred them to the oviducts of pseudopregnant females. Of the 11 live pups we obtained, 18% harbored the desired P2A-mStrawberry insertion, as validated by 5' and 3' integration of the targeted cassette using PCR analysis and sequencing confirmation (Fig. 2C-D; Table 1). Since *Sox2* is an essential gene, we also assayed for the presence of indels in the *Sox2* locus using RFLP analysis (Fig. S2B). 9 out of 11 pups (82%) carried indels in the *Sox2* allele (Fig. S2C), confirming our previous observation that NHEJ is more prevalent than HDR (Fig. S1D). Notably, the high rate of NHEJ is not specific to CRISPR-READI, but is an inherent property of CRISPR/Cas9-based approaches for mouse zygotic editing[42,43], and thus may pose a challenge when editing essential genes. In fact, we encountered challenges with breeding our founder animals that are likely due to NHEJ-mediated *Sox2* deficiency[44,45]. Nevertheless, our results show that CRISPR-READI facilitates efficient knock-in of a fluorescent reporter in live animals, demonstrating its utility in generating reporter mouse models.

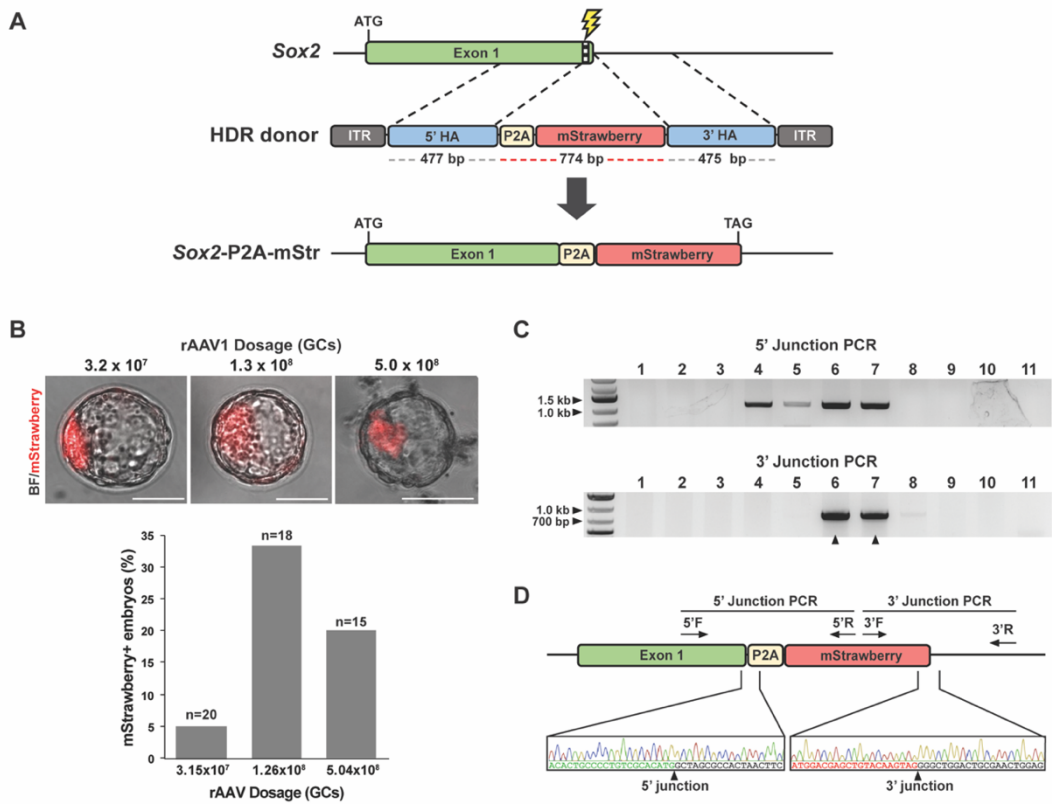


Figure 4.2. CRISPR-READI enables efficient knock-in of fluorescent reporters in embryos and animals.

A) Schematic of strategy to engineer an mStrawberry fluorescent reporter downstream of the endogenous *Sox2* locus. The scAAV1-Sox2-mStr vector contains a 774 bp P2A-mStrawberry cassette flanked by ~480 bp homology arms that mediate insertion at the 3' terminus of the *Sox2* ORF upon HDR editing. ITR: inverted terminal repeat, HA: homology arm.

B) CRISPR-READI efficiently generates embryos with an mStrawberry reporter driven by the endogenous *Sox2* promoter. Embryos were treated with scAAV1-Sox2-mStr at a dose of 3.2×10^7, 1.3×10^8, or 5.0×10^8 GCs. Treated embryos were cultured to the late blastocyst stage and imaged by fluorescent microscopy. In edited blastocysts, merged brightfield and fluorescent images show localization of mStrawberry fluorescence to the inner cell mass, recapitulating the endogenous *Sox2* expression pattern (top). HDR rate was quantified for each dose by the percentage of mStrawberry-positive embryos (bottom). Scale bars: 50 μm.

C) CRISPR-READI efficiently generates Sox2-P2A-mStrawberry knock-in mice. PCR genotyping confirmed correctly edited 5' and 3' junctions of the modified *Sox2* locus in 2 out of 11 mice generated by CRISPR-READI, as indicated by black arrows. We also identified two animals harboring the 5' but not the 3' end of the donor sequence, which is likely due to incomplete HDR (lanes 4 and 5). Primers were designed such that one primer binds outside the homology arm and the other primer binds within the P2A-mStrawberry cassette. 5' expected band size: 1,088 bp, 3' expected band size: 836 bp.

D) Representative Sanger sequencing and chromatograms for correctly edited mice. 5' F/R: forward/reverse primers for 5' junction genotyping, 3' F/R: forward/reverse primers for 3' junction genotyping.

CRISPR-READI achieves site-specific knock-in of large gene cassettes

While CRISPR-READI using scAAV1 donors mediates efficient gene targeting, its 2.4 kb packaging capacity limits its usage for more complex editing schemes that require larger HDR donors. In contrast, ssAAV1 doubles the packaging capacity and can accommodate a donor construct of up to 4.9 kb[38]. To demonstrate the ability of ssAAV1 to mediate knock-in of multi-kilobase cassettes, we applied CRISPR-READI to insert an inducible CreERT2[46] into the *Sox2* locus. We designed a 3 kb ssAAV1 donor vector consisting of a 2112 bp Sox2-P2A-CreERT2 cassette flanked by two ~480 bp homology arms (ssAAV1-Sox2-CreERT2) (Fig. 3A). Treated embryos were cultured to the blastocyst stage for PCR genotyping analysis. Remarkably, 69% (22 out of 32) of treated blastocysts exhibited correct targeting of the P2A-CreERT2 cassette, highlighting the considerable HDR efficiency achieved by CRISPR-READI (Fig. 3B; Table 1). We then replicated this experiment and transferred the embryos to pseudopregnant females to generate live mice. 40% (2 out of 5) of the resulting pups were successfully edited, as confirmed by PCR genotyping and Sanger sequencing (Fig. 3C-D; Table 1). Similar to results obtained from the scAAV1-Sox2-mStr experiments, 60% (3 out of 5) of the mice carried indels in the *Sox2* locus, with both HDR-positive animals also harboring indels (Fig. S2D). Since HDR donors containing recombinase platforms (Cre and Flp) and tetracycline-inducible systems can be easily accommodated by the multi-kilobase packaging capacities of scAAV1 (~2.4 kb) and ssAAV1 (~4.9 kb), CRISPR-READI is well-suited for engineering a wide range of complex mouse models with temporal and spatial gene regulation.

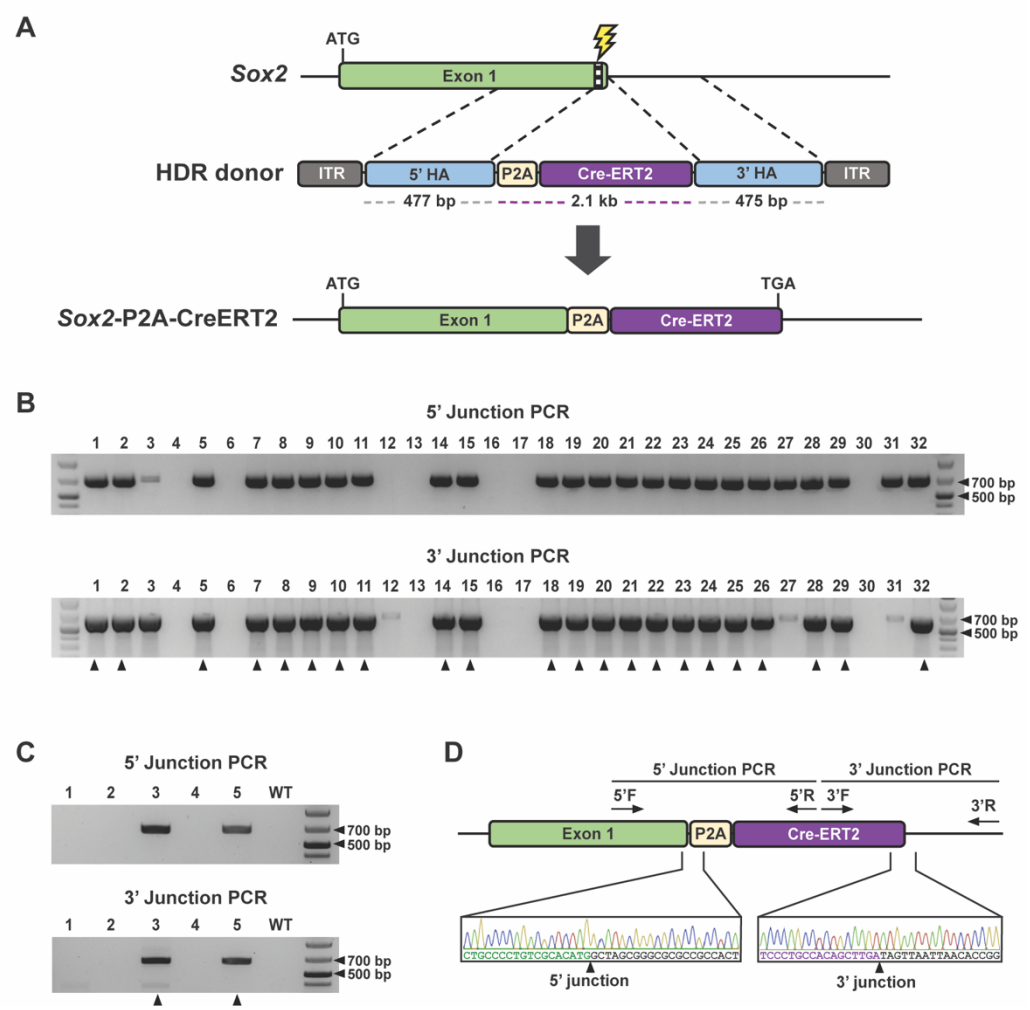


Figure 4.3. CRISPR-READI efficiently engineers an inducible CreERT2 reporter driven by its endogenous promoter.

A) Schematic of strategy to create a *Sox2*-driven inducible CreERT2 reporter. The ssAAV1-Sox2-CreERT2 vector contains a 2112 bp P2A-CreERT2 cassette flanked by ~480 bp homology arms that mediate P2A-CreERT2 insertion at the 3' terminus of the *Sox2* ORF upon successful HDR editing. ITR: inverted terminal repeat, HA: homology arm.

B) CRISPR-READI efficiently produces embryos with an inducible CreERT2 reporter driven by the endogenous *Sox2* promoter. PCR genotyping confirmed the correctly edited 5' and 3' junctions of the modified *Sox2* locus in 22 out of 32 embryos, as indicated by black arrows. Primers were designed such that one primer binds outside the homology arm and the other primer binds within the CreERT2 cassette. 5' expected band size: 658 bp, 3' expected band size: 666 bp.

C) CRISPR-READI efficiently generates Sox2-P2A-CreERT2 mice. PCR genotyping confirmed correctly edited 5' and 3' junctions of the modified *Sox2* locus in 2 out of 5 mice with previously described primers.

D) Representative Sanger sequencing and chromatograms for correctly edited mice. 5' F/R: forward/reverse primers for 5' junction genotyping, 3' F/R: forward/reverse primers for 3' junction genotyping.

To evaluate the capacity of CRISPR-READI to engineer even larger knock-ins, we designed an editing scheme to insert a 3.3 kb FLEX-mStrawberry expression cassette into the *Rosa26* safe harbor locus. The resulting 4.3 kb donor consists of a splice acceptor and a polyA element that terminate the endogenous *Rosa26* transcript, a CAG promoter driving expression of a FLEX (flip-excision) mStrawberry switch, a chimeric SV40/bGH polyA signal, and two ~500 bp homology arms (ssAAV1-R26-FLEX-mStr) (Fig. 4A). We applied CRISPR-READI using ssAAV1-R26-FLEX-mStr to target the FLEX-mStrawberry cassette to the *Rosa26* locus in CMV-Cre embryos [47]. In successfully edited embryos, mStrawberry expression is irreversibly activated through consecutive Cre-mediated recombination events (Fig. 4A). We detected red fluorescent signal in 34% (11 out of 32) 8-cell stage embryos, with an increase in signal intensity over the following 24 hours, indicating functional recombination of the FLEX construct in developing embryos (Fig. 4B). In a separate experiment, we confirmed HDR-mediated editing in 34% (9 out of 26) of embryos by validating 5' and 3' integration of the targeted cassette using PCR (Fig. 4C; Table 1). Finally, we generated live animals using this editing scheme in a C57BL/6J background, obtaining 27% (3 out of 11) successfully edited pups, as confirmed by PCR and Sanger sequencing (Fig. 4D-E; Table 1). Our findings affirm the utility of CRISPR-READI for highly efficient, site-specific knock-in of large expression cassettes at endogenous loci.

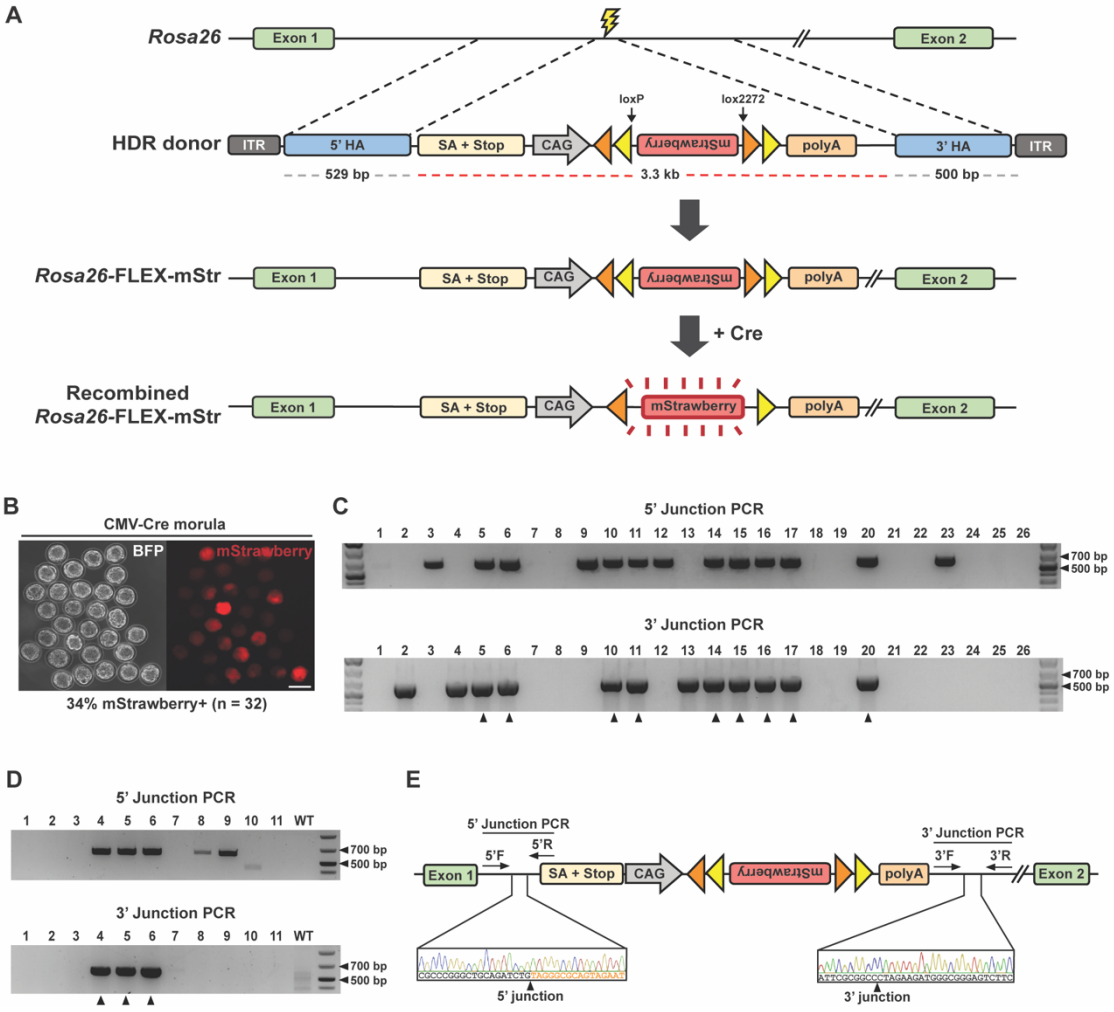


Figure 4.4. CRISPR-READI enables efficient site-specific insertion of a large expression cassette.

A) Schematic of strategy to knock a CAG-FLEX-mStrawberry cassette into the *Rosa26* locus. The ssAAV1-R26-FLEX-mStr vector contains a splice acceptor + polyA element, a CAG promoter, an inverted mStrawberry ORF flanked by loxP and lox2272 sites, chimeric SV40/bGH polyA signal, and ~500 bp homology arms that direct targeting to the first *Rosa26* intron. Upon Cre-mediated recombination, the mStrawberry ORF is irreversibly inverted, leading to robust mStrawberry expression. ITR: inverted terminal repeat, HA: homology arm, SA: splice acceptor.

B) Representative fluorescent image of CMV-Cre embryos treated with ssAAV1-R26-FLEX-mStr. Successful editing and Cre-mediated recombination lead to robust mStrawberry expression by the morula stage. Scale bar: 50 μm.

C) PCR genotyping confirmed the correctly edited 5' and 3' junctions of the modified *Rosa26* locus in 9 out of 26 embryos, as indicated by black arrows. Primers were designed such that one primer binds outside the homology arm and one primer binds within the ssAAV1-R26-FLEX-mStr cassette. 5' expected band size: 608 bp, 3' expected band size: 626 bp.

D) CRISPR-READI efficiently generates Rosa-FLEX-mStr mice. PCR genotyping confirmed correctly edited 5' and 3' junctions of the modified *Rosa26* locus in 3 out of 11 mice with previously described primers. Notably, we also detected two animals exhibiting partial sequence insertion, similar to our Sox2-P2A-mStr animals (lanes 8 and 9).

E) Representative Sanger sequencing and chromatograms for correctly edited mice. 5' F/R: forward/reverse primers used for 5' junction genotyping, 3' F/R: forward/reverse primers used for 3' junction genotyping.

rAAV donor	rAAV titer (GCs)	rAAV preincubation	Embryos Treated	Embryos developed to morula	Embryos transferred	Pups born	Embryos/Pups Assayed	HDR edited embryos/pups
scAAV1-Tyr	1.1 x 10^8	5 hours	43	27 (63%)	N/A	N/A	24	8 (33%)
scAAV1-Tyr	4.3 x 10^8	5 hours	42	25 (60%)	N/A	N/A	24	5 (21%)
scAAV1-Tyr	1.7 x 10^9	5 hours	46	19 (41%)	N/A	N/A	19	9 (48%)
scAAV1-Tyr	1.7 x 10^9	2 hours	50	12 (24%)	N/A	N/A	10	4 (40%)
scAAV1-Tyr	1.7 x 10^9	4 hours	47	10 (21%)	N/A	N/A	10	6 (60%)
scAAV1-Tyr	1.7 x 10^9	6 hours	44	14 (32%)	N/A	N/A	13	10 (77%)
scAAV1-Tyr	1.7 x 10^9	8 hours	49	12 (24%)	N/A	N/A	7	3 (43%)
scAAV1-Tyr	1.7 x 10^9	10 hours	47	8 (17%)	N/A	N/A	6	3 (50%)
scAAV1-Tyr	1.7 x 10^9	6 hours	51	16 (31%)	N/A	N/A	16	12 (75%)
ssAAV1-Tyr	1.7 x 10^9	6 hours	45	15 (33%)	N/A	N/A	14	9 (64%)
scAAV1-Tyr	1.0 x 10^8	6 hours	76	29 (38%)	N/A	N/A	22	7 (32%)
scAAV1-Tyr, ssAAV1-SpCas9, scAAV1-Tyr-gRNA	1.0 x 10^8	6 hours	100	36 (36%)	N/A	N/A	22	8 (36%)
scAAV1-Sox2-mStr	3.2 x 10^7	6 hours	67	36 (54%)	N/A	N/A	20	1 (5%)
scAAV1-Sox2-mStr	1.3 x 10^8	6 hours	90	38 (42%)	N/A	N/A	18	6 (33%)
scAAV1-Sox2-mStr	5.0 x 10^8	6 hours	109	15 (14%)	N/A	N/A	15	3 (20%)
ssAAV1-Sox2-CreERT2	1.3 x 10^8	6 hours	52	36 (69%)	N/A	N/A	32	22 (69%)
ssAAV1-R26-FLEX-mStr	2.9 x 10^7	6 hours	36	26 (72%)	N/A	N/A	26	9 (34%)
scAAV1-Sox2-mStr	1.3 x 10^8	6 hours	135	N/A	59	11	11	2 (18%)
ssAAV1-Sox2-CreERT2	1.3 x 10^8	6 hours	77	N/A	60	5	5	2 (40%)
ssAAV1-R26-FLEX-mStr	1.3 x 10^8	6 hours	250	N/A	100	11	11	3 (27%)

Table 4.1. Dose, viability, and editing efficiency for CRISPR-READI experiments

Oligos for sgRNA synthesis	
T7-sgTyr	GGATCCTAATACGACTCACTATAGGGGTGGATGACCGTGAGTCCGTTTTAGAGCTAGAA
T7-sgSox2	GGATCCTAATACGACTCACTATAGTGCCCCTGTCGCACATGTGAGTTTTAGAGCTAGAA
T7-sgSox2	GGATCCTAATACGACTCACTATAGACTCCAGTCTTTCTAGAAGAGTTTTAGAGCTAGAA
T7RevLong	AAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTT AACTTGCTATTTCTAGCTCTAAAAC
T7FwdAmp	GGATCCTAATACGACTCACTATAG
T7RevAmp	AAAAAAGCACCGACTCGG
Primers for RFLP and genotyping	
Tyr RFLP F1	GCACATTTTCTGCCCTGAGA
Tyr RFLP R1	AGTTACCTCACTATGGGCTATGT
Tyr RFLP F2	TGTGGCTGCTGAAGTACCAG
Tyr RFLP R2	AGTCATGTGCTTTGCAGAAGA
Sox2-P2A-Str 5' F	TAAGTACACGCTTCCCGGAG
Sox2-P2A-Str 5' R	AGCCCATGGTCTTCTTCTGC
Sox2-P2A-Str 3' F	GCAGAAGAAGACCATGGGCT
Sox2-P2A-Str 3' R	CCCAGCAAGAACCCTTTCCT
Sox2-P2A-Cre 5' F1	TTATAAATACCGGCCGCGGC
Sox2-P2A-Cre 5' R1	GTTGCATCGACCGGTAATGC
Sox2-P2A-Cre 5' F2	TAAGTACACGCTTCCCGGAG
Sox2-P2A-Cre 5' R2	CATTGGCCCGGGATTCTCTT
Sox2-P2A-Cre 3' F1	CCCACATCAGGCACATGAGT
Sox2-P2A-Cre 3' R1	CACGAAAACGGTCTTGCCAG
Sox2-P2A-Cre 3' F2	CGGGCTCTACTTCATCGCAT
Sox2-P2A-Cre 3' R2	CCCAGCAAGAACCCTTTCCT
Sox2 RFLP F	TAAGTACACGCTTCCCGGAG
Sox2 RFLP R	CCCAGCAAGAACCCTTTCCT
Rosa26 5' F1	CTGGGGGAGTCGTTTTACCC
Rosa26 5' R1	ACTGGAAAGACCGCGAAGAG
Rosa26 5' F2	TCGTCGTCTGATTGGCTCTC
Rosa26 5' R2	CTACTGCGCCCTACAGATCTG
Rosa26 3' F1	GGTGGGCTCTATGGCTTCTG
Rosa26 3' R1	AGAACTGCAGTGTTGAGGCC
Rosa26 3' F2	GTTTCAGGTTCAGGGGGAGG
Rosa26 3' R2	ACAGCCTCGATTTGTGGTGT
Primers for plasmid construction	
Tyr EcoRI 1kb XbaI F	TCTAGATTAGTTACCTCACTATGGGCTATGT
Tyr EcoRI 1kb XbaI R	TCTAGATTCTGCATCTCTCCAATCCCAG
pGEM stuffer HindIII F	AAGCTTGGGATAATACCGCGCCACAT
pGEM stuffer HindIII R	AAGCTTCCCTGGCGTTACCCAACTTA
Sox2 P2A Straw XbaI F	TCTAGATGCATGCTCGAATAAACAATGGA
Sox2 P2A Straw XbaI R	TCTAGAAGCTCGGATCACATGAACGG
mStr Fwd	GCATGGTACCGCCACCATGGTGAGCAAGGGCGAG
mStr Rev	ACGTCTCGAGCTACTTGTACAGCTCGTCCATGC
Rosa FLEX mStr Fwd	GCATATGCATCATTGGCTCGTGTTCGTG
Rosa FLEX mStr Rev	GCATGCGGCCGCTGCCAATGCTCTGTCTAGGG
CAG mStr Fwd	GCATGCTAGCGCTCTAGGAAGAGTACCATTGACG
CAG mStr Rev	GCATACGCGTGAATTCGAACCTCTTCGAGG
NheI AscI CreERT2 fwd	GCATGCTAGCGGGCGCGCC GCCACTAACTTCTCCCTGTT
SgrAI PacI CreERT2 rev	GCATCGCCGGTGTTAATTAACTATCAAGCTGTGGCAGG
FseI Sox2 donor fwd	GCATGGCCGGCCAGCTCGGATCACATGAAC
NotI Sox2 donor rev	AGCTGCGGCCGCGACTACCTTGGTGATCTCG
KpnI-U6-sgTyr	GCATGGTACCGAGGGCCTATTTCCCATG
NotI-U6-sgTyr	GCATGCGGCCGCCGCTAAAAACGGACTAGCC

Table 4.2. sgRNA sequences and PCR primers for cloning, RFLP, and genotyping

4.3 Discussion

Electroporation-based CRISPR technologies offer significant advantages over microinjection in throughput, affordability, and ease of use. Previously, we developed CRISPR-EZ, an approach that utilizes Cas9/sgRNA RNP and short ssODN (<200 nt) electroporation to achieve highly efficient mouse genome editing[19,43]. Many groups have reported success using similar electroporation strategies, including both *ex vivo* and *in utero* approaches[17–23]. However, inefficient delivery of long HDR donors into pronuclear embryos by electroporation, possibly due to their highly anionic properties, limits its application for complex genome engineering. While lssODN electroporation has mediated large knock-ins in mice with some success, targeting efficiencies are typically suboptimal[9,23]. Furthermore, current lssODN synthesis protocols are hampered by low yield, costly reagents, and limited length[10,11,26,27,29]. Hence, lssODNs remain unsuitable for many mouse engineering schemes involving large genomic modifications.

In contrast, CRISPR-READI harnesses rAAV vectors to deliver large donor constructs into mouse embryos for complex editing[38]. With multi-kilobase packaging capacities, scAAVs (~2.4 kb) and ssAAVs (~4.9 kb) can accommodate many commonly used HDR donors for engineering large knock-ins in mice. The enhanced genome stability and transduction efficiency of scAAVs may further improve gene correction frequency in pronuclear embryos[37,48–50]. With rAAV donors, CRISPR-READI enables a variety of sophisticated HDR editing schemes that are used for disease modeling and functional studies; these include precise sequence modifications (Fig. 1), as well as site-specific integration of reporters (Fig. 2), recombinases (Fig. 3), and large expression cassettes (Fig. 4). Our proof-of-concept studies demonstrate the potential of CRISPR-READI to efficiently produce a vast range of popular knock-in models, such as those harboring fluorescent proteins, luciferase reporters, HaloTag, Cre and Flp recombinase, Tet-on/off inducible systems, and many others. Moreover, we show that CRISPR-READI is an electroporation-based genome editing methodology capable of delivering a 4.3 kb donor template and inserting a 3.3 kb reporter cassette, which was previously only feasible with dsDNA microinjection or ESC-based approaches.

While our manuscript was in preparation, Yoon *et al* reported AAV6 as the most efficient natural AAV serotype for transduction of mouse zygotes[36]. In our studies, AAV1 and AAV6, which share 99.2% amino acid sequence identity[51], exhibited comparable infectivity, with AAV1 being superior in viral particle assembly. Yoon *et al* also described a triple AAV co-infection system to deliver CRISPR-Cas9 components (AAV-Cas9, AAV-sgRNA) and an HDR donor (AAV-donor). Using the *Tyr* editing scheme, we directly compared CRISPR-READI (scAAV1-Tyr and Cas9/sgRNA RNP) with a triple AAV1 co-infection strategy (scAAV1-Tyr, ssAAV1-SpCas9 and scAAV1-Tyr-gRNA) at the same AAV dosage (1×10^8 GCs). The two methods yielded similar HDR rates and embryo viability (Fig. S1F; Table 1); however, the triple AAV infection strategy resulted in multiple embryos harboring large deletions, while CRISPR-READI did not (Fig. S1F).

CRISPR-READI likely carries several notable advantages compared to the triple AAV co-infection method. First, CRISPR-READI harnesses pre-assembled Cas9/sgRNA complexes that are electroporated with 100% delivery efficiency[19] and are active immediately upon delivery, whereas virally delivered Cas9 may exhibit delayed activity, as it must undergo transcription, translation, and RNP assembly prior to mediating cleavage. Second, it was recently shown that

Cas9 can induce large-scale genomic deletions[52], which may be facilitated by the persistence of AAV episomes expressing Cas9 and sgRNA within the developing mouse embryo. In comparison, RNPs are less prone to these effects due to their rapid turnover[53]. Finally, Yoon *et al* observed rare but detectable integration of both the AAV-Cas9 and AAV-sgRNA vectors at their target site in the *Tyr* locus[36], leading to long-term expression of Cas9 and/or sgRNAs, whereas CRISPR-READI circumvents this undesirable knock-in through the use of RNPs.

We demonstrate that CRISPR-READI can be broadly applied to various gene targeting schemes to rapidly generate mouse lines for functional studies and disease modeling. The unprecedented efficiency of CRISPR-READI for knocking in gene cassettes up to 3.3 kb makes this approach highly appealing for creating multi-kilobase genome modifications. With such high editing efficiencies, CRISPR-READI could enable multiplexed gene targeting to simultaneously modify two different target loci. Furthermore, a dual rAAV donor system carrying a split gene has recently been shown to mediate insertion of large gene cassettes through sequential homologous recombination events in primary human cells[54], and it would be interesting to explore this strategy in the context of mouse embryo editing. Taken together, CRISPR-READI provides a simple, efficient, and high-throughput alternative to microinjection and ESC-based methods for sophisticated mouse genome engineering.

4.4 Methods

Plasmid construction

For construction of the scAAV-Tyr donor plasmid, a *Tyr* HDR donor fragment and a 1 kb stuffer sequence were assembled to optimize the genome size for vector packaging[38]. Briefly, a 968 bp *Tyr* fragment with the engineered EcoRI site was amplified from a 1 kb megamer ssDNA fragment (IDT custom synthesis) using Tyr-XbaI-F and Tyr-XbaI-R primers (Table 2) and then digested with XbaI for subsequent cloning into rAAV destination vectors. To prepare the destination scAAV vector, a plasmid derivative of scAAV-CMV-GFP[55] containing a nanoluciferase cassette was digested with XbaI to remove the majority of the nanoluciferase ORF, leaving a 1 kb stuffer sequence. The digested *Tyr* insert was then ligated to the scAAV destination vector containing the stuffer sequence to generate the final scAAV-Tyr plasmid. To build the ssAAV-Tyr donor plasmid, an additional stuffer sequence was first amplified from pGEM-T easy vector (Promega, A1360) using pGEM-stuffer-HindIII-F and pGEM-stuffer-HindIII-R primers and digested with HindIII. A plasmid derivative of pX601 (Addgene #61591) containing a nanoluciferase cassette was digested with HindIII and ligated to the pGEM stuffer sequence to generate pX601-long. Subsequently, pX601-long was digested with XbaI and ligated to the previously described XbaI digested *Tyr* insert to create the ssAAV-Tyr plasmid.

For construction of the scAAV-Sox2-Str donor plasmid, a *Sox2* targeting cassette was subcloned into the XbaI digested scAAV destination vector described above. Briefly, a 1726 bp fragment containing a P2A-mStrawberry cassette flanked by homology arms from the *Sox2* locus was amplified from a Sox2-P2A-mStrawberry plasmid (gift from D Stafford) using Sox2-P2A-Straw-XbaI-F and Sox2-P2A-Straw-XbaI-R primers and then digested with XbaI. This insert was ligated to the XbaI digested scAAV vector described above to generate the scAAV-Sox2-Str plasmid. For the ssAAV-Sox2-Cre-ERT2 donor plasmid, the previously described *Sox2* targeting

cassette and a 981 bp stuffer sequence were assembled and subcloned into a ssAAV destination vector, and the P2A-mStrawberry cassette was then replaced with a P2A-Cre-ERT2 cassette. Briefly, the *Sox2* targeting construct with an additional 981 bp stuffer sequence downstream of the donor was PCR amplified from an ssAAV-Sox2-Str plasmid (data not shown) using FseI-Sox2-donor fwd and NotI-Sox2-donor rev primers and subcloned into a plasmid derivative of pX601 using FseI and NotI to generate pX601-Sox2. The P2A-Cre-ERT2 donor was then PCR amplified from a plasmid derivative of WT1-2A-eGFP (Addgene #82333) containing a P2A-Cre-ERT2 cassette using NheI-AscI-CreERT2-fwd and SgrAI-PacI-CreERT2-rev primers, digested with NheI and SgrAI, and cloned into pX601-Sox2 to produce the final ssAAV-Sox2-Cre-ERT2 plasmid.

For construction of the ssAAV-R26-FLEX donor plasmid, a *Rosa26* targeting construct was first subcloned into a ssAAV destination vector, and the original donor cassette was then replaced with a CAG-FLEX-mStrawberry cassette. Briefly, a *Rosa26* donor construct containing ~500 bp homology arms was PCR amplified from pR26-CAG-AsiSI/MluI (Addgene #74286) with Rosa-FLEX-mStr-Fwd and Rosa-FLEX-mStr-Rev primers, digested with NotI and NsiI, and ligated into a plasmid derivative of pX601 (Addgene #61591) to produce ssAAV-Rosa. To generate the CAG-FLEX-mStrawberry cassette, mStrawberry was PCR amplified from the scAAV-Sox2-Str plasmid with mStr-Fwd and mStr-Rev primers, digested with KpnI and XhoI, and cloned into pAAV-FLEX-GFP (Addgene #28304), generating an intermediate pAAV-FLEX-mStr plasmid. The assembled CAG-FLEX-mStrawberry cassette was then PCR amplified from pAAV-FLEX-mStr with CAG-mStr-Fwd and CAG-mStr-Rev primers, digested with NheI and MluI, and ligated into ssAAV-Rosa to generate the final ssAAV-R26-FLEX plasmid.

For construction of the AAV1-SpCas9 vector, the pAAV-nEFCas9 plasmid (Addgene #87115) was digested with HindIII and AgeI to remove the truncated 5'LTR of the nEF promoter, and a 9 bp stuffer sequence was ligated in to generate a shortened EF1α (EFS) promoter. For construction of the AAV1-Tyr-gRNA vector, a plasmid derivative of scAAV-CMV-GFP[55] with a CAG promoter in place of the original CMV promoter and a short 48 bp polyA in place of the original SV40 polyA was digested with KpnI and NotI to produce scAAV-CAG-GFP. The sgTyr sequence was cloned into pX458 (Addgene #48138) with BbsI, and the U6-Tyr-gRNA expression cassette was PCR amplified using KpnI-U6-sgTyr and NotI-U6-sgTyr primers. The PCR amplicon was then digested with KpnI and NotI and ligated into scAAV-CAG-GFP to generate the final scAAV-U6-sgTyr-CAG-GFP plasmid.

Production and purification of recombinant AAV vectors

HEK293T cells were obtained from the American Type Culture Collection (Manassas, CRL-3216) and cultured in DMEM (GIBCO, 12800082) with 10% fetal bovine serum (Invitrogen, 10437028) and 1% Antibiotic-Antimycotic (GIBCO, 15240062) at 37°C and 5% CO_2. All recombinant AAV vectors were packaged in HEK293T cells, as previously described[56,57]. In summary, recombinant AAV vectors were produced by triple transient transfection of a helper plasmid encoding adenoviral helper genes, an AAV helper plasmid encoding AAV *rep* and *cap* genes, and a transfer plasmid containing donor constructs flanked by AAV ITRs, into HEK293T cells using polyethylenimine (PEI) (Polysciences, 23966-1). Culture media was changed 14-16 hours post-transfection to reduce PEI-associated toxicity. At 72 hours post-transfection, cells were dissociated from culture dishes using a cell scraper and pelleted by centrifugation at 1500 *g* for 2.5

min. Cells were then resuspended in AAV lysis buffer (50 mM Tris base, 150 mM NaCl, pH 8.2-8.5) and lysed with 3 freeze/thaw cycles using a dry ice / ethanol bath and a 37°C water bath. Supernatant was collected following centrifugation at 10,000 rpm for 10 min in a tabletop centrifuge. For additional purification of vector preparations (optional), clarified cell lysate and culture supernatant can be combined with a solution of 40% polyethylene glycol (PEG) 8000 (Sigma-Aldrich, P2139) and 2.5 M NaCl at a 4:1 ratio, incubated at 4°C overnight, centrifuged at 2500 *g* for 30 min to harvest precipitated AAV particles, and resuspended in AAV lysis buffer[58]. Crude lysates were treated with 10 U benzonase (Sigma-Aldrich, E8263) per mL of lysate for 30 min at 37°C.

For AAV purification by iodixanol density centrifugation, discontinuous gradients comprised of 15%, 25%, 40%, and 54% iodixanol layers were set up in Optiseal tubes (Beckman Coulter, 362185) using OptiPrep (Axis-Shield, AVS-1114542). Benzonase-treated lysates were then loaded onto iodixanol gradients and centrifuged at 174,000 *g* for 2 hr at 18°C. Tubes were punctured using 21-gauge needles attached to a 1 mL syringe at the 40% / 54% interface, and the 40% iodixanol fraction containing AAV particles was collected. Vector preps were then buffer exchanged and concentrated into PBS with 0.001% or 0.00025% Tween 20 (Sigma-Aldrich, 9416) using Amicon filtration (EMD Millipore, UFC910024) at 3000 *g*. DNase-resistant (Sigma-Aldrich, 04536282001) viral genomic titers were measured by real-time qPCR[57] using SYBR Green I (Invitrogen, S7567) and the CFX96 Real-Time PCR cycler (Bio-Rad, 1855195).

Cas9/sgRNA RNP assembly

To synthesize sgRNAs *in vitro*, a DNA template that contained a T7 promoter, a 20-nt guide sequence, and a sgRNA scaffold was generated by overlapping PCR. Specifically, we performed PCRs using Phusion high fidelity DNA polymerase (New England Biolabs, catalog no. M0530), with the annealed product from a uniquely designed oligo (5'-GGA TCC TAA TAC GAC TCA CTA TAG‑guide sequence‑GTT TTA GAG CTA GAA-3', 0.02 μM) and a common oligo T7RevLong (5'-AAA AAA GCA CCG ACT CGG TGC CAC TTT TTC AAG TTG ATA ACG GAC TAG CCT TAT TTT AAC TTG CTA TTT CTA GCT CTA AAA C-3', 0.02 μM) as the template, and T7FwdAmp (5'-GGA TCC TAA TAC GAC TCA CTA TAG-3', 1 μM) and T7RevAmp (5'-AAA AAA GCA CCG ACT CGG-3', 1 μM) as two common primers. All oligo sequences are listed in Table 2. The thermocycler setting consisted of 30 cycles of 95°C for 10 s, 57°C for 10 s, and 72°C for 10 s. A 20-μl *in vitro* transcription reaction consisting of 25 ng/μl PCR-amplified DNA template, 10 mM NTPs, and 1 unit of T7 RNA polymerase (New England Biolabs, E2040S) was incubated at 37°C for more than 18 h, followed by a brief treatment of RNase-Free DNase I (New England Biolabs, M0303S, 2 units) at room temperature for 20 min. The *in vitro* synthesized sgRNAs were purified by magnetic SeraMeg Speedbeads (GE Healthcare, 65152105050250). The *in vitro* transcription reaction was first brought to 100 μl in volume with 100% ethanol, followed by gentle mixing of 100 μl of SeraMeg Speedbeads 10 times before a 5-min room temperature incubation. The reaction was subsequently placed on a magnetic stand (Invitrogen, 12321D) for 5 min under room temperature to allow the formation of compact RNA/bead pellets. After the supernatant was carefully aspirated by pipette, we washed the pellets gently with 80% ethanol three times (2 min wash each time, without pipetting) and air-dried the

pellets for 10 min. sgRNAs bound to the beads were eluted by incubating with 20 μl of RNase-free H_2O (Ambion, AM9937) and stored at −80°C.

To assemble the Cas9/sgRNA RNPs, we incubated purified Cas9 protein (QB3 Macrolab, University of California at Berkeley) in a 1:1.5 molar ratio with sgRNAs to obtain a final concentration of 8 μM Cas9/sgRNA RNPs in a 10-μl solution containing 20 mM HEPES, pH 7.5 (Sigma, H3375), 150 mM KCl (Sigma, P9333), 1 mM $MgCl_2$ (Sigma, M8266), 10% glycerol (Thermo Fisher, BP229), and 1 mM reducing agent tris(2-carboxyethyl)phosphine (TCEP, Sigma, C4706). This mixture was incubated at 37°C for 10 min immediately before use.

CRISPR-READI

Three-to-five-week old female C57BL/6J mice (000664; Jackson Laboratory), or CMV-Cre mice (006054; Jackson Laboratory), were superovulated by intraperitoneal administration of 5 IU of pregnant mare serum gonadotropin (Calbiochem, Millipore, 367222), and 46–48 h later, 5 IU human chorion gonadotropin (Calbiochem, Millipore, 230734). Superovulated females were mated at a 1:1 ratio with 3–8-month-old C57BL/6J males to generate one-cell zygotes at 0.5 days post-coitum. Fertilized zygotes were harvested and washed, as previously described[19]. Briefly, zygotes were harvested from the ampulla of euthanized females, washed in hyaluronidase/M2 solution (Millipore, MR-051-F) to remove cumulus cells, washed four times in M2 media (Zenith, ZFM2–100) supplemented with 4 mg/ml bovine serum albumin (BSA, Sigma, A3311), washed briefly in acid Tyrode's solution (Sigma, T1788) to thin the zona pellucida, and washed again four times in M2 + BSA media. Embryos were then cultured in 20 μL droplets of KSOM + AA media (KCl-enriched simplex optimization medium with amino acid supplement, Zenith Biotech, ZEKS-050) containing the specified AAV vector dosage in 35 × 10 mm culture dishes (CellStar Greiner Bio-One, 627160) at 37°C with 95% humidity and 5% CO_2 for 6 hours prior to RNP electroporation.

AAV-transduced embryos were electroporated with assembled RNPs as previously described[19]. In summary, embryos were transferred to 40 uL of Opti-MEM reduced serum media (Thermo Fisher, 31985062) to dilute the M2 + BSA media, and 10 uL of Opti-MEM containing the embryos were mixed with 10 uL assembled Cas9/sgRNA RNPs. The 20-μl embryo and RNP mixture was transferred to a 1-mm electroporation cuvette (Bio-Rad, 1652089) and electroporated (square wave, 6 pulses, 30 V, 3 ms duration, 100 ms interval) using a Gene Pulser XCell electroporator (Bio-Rad, 1652660). Zygotes were recovered from the cuvette by flushing three times with 100 μl of KSOM + AA media, then transferred into the culture droplets containing AAV for a total incubation length of 24 hours. The following day, embryos were transferred to fresh KSOM + AA media overlayed with mineral oil until analysis or oviduct transfer. For generating live mice, treated zygotes that successfully developed into two-cell embryos were surgically transferred into the oviducts of pseudopregnant CD1 females (Charles River, Strain 022), using 15 embryos per oviduct.

Fluorescence imaging and analysis of edited embryos

Embryos treated with rAAV CMV-GFP and scAAV1-Sox2-Str were imaged using a Zeiss Observer A1 fluorescent microscope 48 hours and 72 hours post-transduction, respectively. Images were processed using ZEN software (Zeiss), and fluorescence intensity was quantified for

embryos transduced with rAAV-CMV-GFP using ImageJ software. Fluorescence intensity was calculated by the following equation: (total signal intensity – negative control background signal) / number of embryos.

RFLP and genotyping analyses of edited embryos or pups

Crude DNA extract was obtained from morula and blastocyst stage embryos by placing individual embryos in 10 uL of embryo lysis buffer (0.2 mg/ml Proteinase K, 50 mM KCl, 10 mM Tris-HCl pH 8.3, 2.5 mM MgCl2, 0.1 mg/ml gelatin, 0.45% NP40, 0.45% Tween 20), followed by heating at 55° for 4 hours to lyse the embryos, as previously described[19]. Tail DNA was extracted by standard chloroform extraction. For all embryo RFLP and genotyping, two nested PCR reactions were performed using an external and internal pair of primers in order to obtain sufficient signal from crude embryo lysate, while only the internal set of the primers were used when amplifying from mouse tail samples. All PCRs were conducted using GoTaq (Promega, M712), except for Sox2-CreERT2 5' junction and Rosa26-FLEX 5' junction genotyping, which were conducted using Q5 polymerase with GC enhancer due to high GC content (New England Biolabs, M0491S). For *Tyr* editing experiments, successful editing replaces the endogenous HinfI restriction site with an EcoRI site, while for *Sox2* editing experiments, the presence of indels over the sgRNA target sequence disrupts a PciI restriction site. RFLP was performed to screen for these editing events by digesting the PCR products with EcoRI (New England Biolabs, R0101S) or PciI (New England Biolabs, R0655S), respectively, for 4 hours at 37°C. For Sox2-P2A-mStrawberry, Sox2-P2A-CreERT2 and Rosa26-FLEX-mStrawberry knock-in experiments, primer pairs used for PCR genotyping of the 5' and 3' junctions consisted of one primer designed from the genomic region flanking the homology arm and another from within the inserted sequence. All primers used for RFLP and genotyping analysis are listed in Table 2.

4.5 Acknowledgements

We thank MA Dewitt, J Corn, AJ Modzelewski, S Jin, C Jeans and T Gaj, for stimulating discussion and input on experimental design. We also thank D Stafford for providing reagents.

4.6 Funding

LH is a Thomas and Stacey Siebel Distinguished Chair Professor, supported by a Howard Hughes Medical Institute (HHMI) Faculty Scholar award, a Bakar Fellow award at UC Berkeley, and several grants from the National Institutes of Health (NIH; 1R01GM114414, 1R21HD088885, 2R01CA139067, GRANT12095758). SC was supported by the Cancer Research Coordinating Committee Fellowship. SS was supported by the American Heart Association Predoctoral Fellowship, NIH R21 EB021572-01, and NIH R01 EY022975.

Conflict of interest statement. DVS is an inventor on patents involving AAV directed evolution and a co-founder and shareholder of a company developing AAV vectors for clinical gene therapy. In addition, he is on the board of directors of and holds shares in a second company developing AAV vectors for clinical gene therapy. SC and LH are inventors on patents involving an

electroporation-based CRISPR technology for mouse genome engineering and are founders of a company to further develop this technology for mammalian genome editing.

4.7 References

1. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–21 (2012).
2. Yang, H. *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370–9 (2013).
3. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
4. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–23 (2013).
5. Li, D. *et al.* Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 681–683 (2013).
6. Fujii, W., Kawasaki, K., Sugiura, K. & Naito, K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res.* **41**, e187–e187 (2013).
7. Aida, T. *et al.* Cloning-free CRISPR/Cas system facilitates functional cassette knock-in in mice. *Genome Biol.* **16**, 1–11 (2015).
8. Nakagawa, Y. *et al.* Ultra-superovulation for the CRISPR-Cas9-mediated production of gene-knockout, single-amino-acid-substituted, and floxed mice. *Biol. Open* **5**, 1142–8 (2016).
9. Yoshimi, K. *et al.* ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. *Nat. Commun.* **7**, 10431 (2016).
10. Quadros, R. M. *et al.* Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. *Genome Biol.* **18**, 92 (2017).
11. Miura, H., Quadros, R. M., Gurumurthy, C. B. & Ohtsuka, M. Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. *Nat. Protoc.* **13**, 195–215 (2017).
12. Yao, X. *et al.* Tild-CRISPR Allows for Efficient and Precise Gene Knockin in Mouse and Human Cells. *Dev. Cell* **45**, 526-536.e5 (2018).
13. Gu, B., Posfai, E. & Rossant, J. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat. Biotechnol.* **36**, 632–637 (2018).
14. Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542 (2015).
15. Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Yagle, M. K. & Palmiter, R. D. Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. *Proc.*

Natl. Acad. Sci. U. S. A. **82**, 4438–42 (1985).

16. Nagy, A., Gertenstein, M., Vintersten, K. & Behringer, R. *Manipulating the Mouse Embryo: A Laboratory Manual . Third Edition.* (Cold Spring Harbor Laboratory Press, 2003). doi:10.1086/380032

17. Hashimoto, M. & Takemoto, T. Electroporation enables the efficient mRNA delivery into the mouse zygotes and facilitates CRISPR/Cas<inf>9</inf>-based genome editing. *Sci. Rep.* **5**, 1–3 (2015).

18. Takahashi, G. *et al.* GONAD: Genome-editing via Oviductal Nucleic Acids Delivery system: a novel microinjection independent genome engineering method in mice. *Sci. Rep.* **5**, 11406 (2015).

19. Chen, S., Lee, B., Lee, A. Y. F., Modzelewski, A. J. & He, L. Highly efficient mouse genome editing by CRISPR ribonucleoprotein electroporation of zygotes. *J. Biol. Chem.* **291**, 14457–14467 (2016).

20. Hashimoto, M., Yamashita, Y. & Takemoto, T. Electroporation of Cas9 protein/sgRNA into early pronuclear zygotes generates non-mosaic mutants in the mouse. *Dev. Biol.* **418**, 1–9 (2016).

21. Hur, J. K. *et al.* Targeted mutagenesis in mice by electroporation of Cpf1 ribonucleoproteins. *Nat. Biotechnol.* **34**, 807–808 (2016).

22. Wang, W. *et al.* Delivery of Cas9 protein into mouse zygotes through a series of electroporation dramatically increased the efficiency of model creation. *J. Genet. Genomics* (2016). doi:10.1016/j.jgg.2016.02.004

23. Ohtsuka, M. *et al.* I-GONAD: A robust method for in situ germline genome engineering using CRISPR nucleases. *Genome Biol.* **19**, 1–15 (2018).

24. Remy, S. *et al.* Generation of gene-edited rats by delivery of CRISPR/Cas9 protein and donor DNA into intact zygotes using electroporation. *Sci. Rep.* **7**, 16554 (2017).

25. Tanihara, F. *et al.* Somatic cell reprogramming-free generation of genetically modified pigs. *Sci. Adv.* **2**, e1600803–e1600803 (2016).

26. Miyasaka, Y. *et al.* CLICK: one-step generation of conditional knockout mice. *BMC Genomics* **19**, 318 (2018).

27. Li, H. *et al.* Design and specificity of long ssDNA donors for CRISPR-based knock-in. *bioRxiv* 178905 (2017). doi:10.1101/178905

28. Chen, F. *et al.* High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* **8**, 753–755 (2011).

29. Chen, F. *et al.* High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* **8**, 753–755 (2011).

30. Russell, D. W. & Hirata, R. K. Human gene targeting by viral vectors. *Nat. Genet.* **18**, 325–330 (1998).

31. Hirata, R., Chamberlain, J., Dong, R. & Russell, D. W. Targeted transgene insertion into

human chromosomes by adeno-associated virus vectors. *Nat. Biotechnol.* **20**, 735–738 (2002).

32. Khan, I. F., Hirata, R. K. & Russell, D. W. AAV-mediated gene targeting methods for human cells. *Nat. Protoc.* **6**, 482–501 (2011).

33. Hiramoto, T., Li, L. B., Funk, S. E., Hirata, R. K. & Russell, D. W. Nuclease-free Adeno-Associated Virus-Mediated Il2rg Gene Editing in X-SCID Mice. *Mol. Ther.* (2018). doi:10.1016/j.ymthe.2018.02.028

34. Kotterman, M. A., Chalberg, T. W. & Schaffer, D. V. Viral Vectors for Gene Therapy: Translational and Clinical Outlook. *Annu. Rev. Biomed. Eng.* **17**, 63–89 (2015).

35. Gaj, T. *et al.* Targeted gene knock-in by homology-directed genome editing using Cas9 ribonucleoprotein and AAV donor delivery. *Nucleic Acids Res.* 1–11 (2017). doi:10.1093/nar/gkx154

36. Yoon, Y. *et al.* Streamlined ex vivo and in vivo genome editing in mouse embryos using recombinant adeno-associated viruses. *Nat. Commun.* **9**, 412 (2018).

37. Hirsch, M. L., Green, L., Porteus, M. H. & Samulski, R. J. Self-complementary AAV mediates gene targeting and enhances endonuclease delivery for double-strand break repair. *Gene Ther.* **17**, 1175–1180 (2010).

38. Dong, J.-Y., Fan, P.-D. & Frizzell, R. A. Quantitative Analysis of the Packaging Capacity of Recombinant Adeno-Associated Virus. *Hum. Gene Ther.* **7**, 2101–2112 (1996).

39. Vandenberghe, L. H. *et al.* Efficient Serotype-Dependent Release of Functional Vector into the Culture Medium During Adeno-Associated Virus Manufacturing. *Hum. Gene Ther.* **21**, 1251–1257 (2010).

40. Nonnenmacher, M. & Weber, T. Intracellular transport of recombinant adeno-associated virus vectors. *Gene Ther* **19**, 649–658 (2012).

41. McCarty, D. M. *et al.* Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther.* **10**, 2112–2118 (2003).

42. Chen, S., Lee, B., Lee, A. Y.-F., Modzelewski, A. J. & He, L. Highly Efficient Mouse Genome Editing by CRISPR Ribonucleoprotein Electroporation of Zygotes. *J. Biol. Chem.* **291**, 14457–67 (2016).

43. Modzelewski, A. J. *et al.* Efficient mouse genome engineering by CRISPR-EZ technology. *Nat. Protoc.* **13**, 1253–1274 (2018).

44. Ferri, A. L. M. *et al.* Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. *Development* **131**, 3805–19 (2004).

45. Avilion, A. A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126–40 (2003).

46. Seibler, J. *et al.* Rapid generation of inducible mouse mutants. *Nucleic Acids Res.* **31**, e12 (2003).

47. Schwenk, F., Baron, U. & Rajewsky, K. A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. *Nucleic Acids Res.* **23**, 5080–1 (1995).

48. Ferrari, F. K., Samulski, T., Shenk, T. & Samulski, R. J. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J. Virol.* **70**, 3227–34 (1996).

49. Fisher, K. J. *et al.* Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *J. Virol.* **70**, 520–532 (1996).

50. McCarty, D. M. Self-complementary AAV vectors; advances and applications. *Mol. Ther.* **16**, 1648–1656 (2008).

51. Huang, L. *et al.* Characterization of the Adeno-Associated Virus 1 and 6 Sialic Acid Binding Site. **90**, 5219–5230 (2016).

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

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

54. Bak, R. O. & Porteus, M. H. CRISPR-Mediated Integration of Large Gene Cassettes Using AAV Donor Vectors. *Cell Rep.* **20**, 750–756 (2017).

55. Fu, H. *et al.* Self-complementary adeno-associated virus serotype 2 vector: Global distribution and broad dispersion of AAV-mediated transgene expression in mouse brain. *Mol. Ther.* (2003). doi:10.1016/j.ymthe.2003.08.021

56. Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nat. Protoc.* **1**, 1412–1428 (2006).

57. Gaj, T. & Schaffer, D. V. Adeno-associated virus-mediated delivery of CRISPR-cas systems for genome engineering in mammalian cells. *Cold Spring Harb. Protoc.* **2016**, 941–952 (2016).

58. Ayuso, E. *et al.* High AAV vector purity results in serotype- and tissue-independent enhancement of transduction efficiency. *Gene Ther.* **17**, 503–510 (2010).

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

60. Bao, X. *et al.* Long-term self-renewing human epicardial cells generated from pluripotent stem cells under defined xeno-free conditions. *Nat. Biomed. Eng.* **1**, (2016).

61. Chu, V. T. *et al.* Efficient generation of Rosa26 knock-in mice using CRISPR/Cas9 in C57BL/6 zygotes. *BMC Biotechnol.* **16**, 4 (2016).

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

63. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–308 (2013).

Chapter 5: HITI-Mediated Correction of Human iPSC-Derived Cardiomyocyte Model of Duchenne muscular dystrophy using AAV Vectors

5.1 Introduction

Duchenne muscular dystrophy (DMD) is a progressive muscle wasting disorder caused by mutations in the X-linked dystrophin (*DMD*) gene that lead to premature death for 1 in 5000 males[1]. Comprised of 79 exons and primarily expressed in muscle cells, the longest *DMD* isoform encodes the 427 kDa rod-shaped dystrophin protein that links the cytoskeleton to a complex of membrane proteins and regulates important signaling pathways[2]. In DMD patients, absence of dystrophin results in structural and signaling aberrations in myocytes that manifest as severe degeneration of both skeletal and cardiac muscle and culminate in fatal respiratory and cardiac complications[2]. As the massive size (~11 kb) of *DMD* exceeds the packaging capacity of existing vectors, a shortened version of the gene called microdystrophin has been engineered for AAV-mediated cDNA delivery to alleviate muscle damage[3]. Recently, robust exogenous expression of microdystrophin in myocytes yielded early signs of improved muscle function in DMD patients (clinicaltrials.gov identifier: NCT03375164), though its therapeutic efficacy remains curbed by its highly truncated form.

To overcome the functional limitations of microdystrophin, gene editing approaches have been developed with the aim of restoring endogenous dystrophin expression. In particular, the technical ease and programmability of CRISPR/Cas9 technologies have precipitated the adoption of these editing tools for treating DMD[4]. While CRISPR/Cas9-induced homology-directed repair (HDR) offers the potential for accurate repair of the *DMD* gene, downregulation of homologous recombination in post-mitotic myocytes[5, 6] poses a major challenge. To enhance homology-directed genome engineering, modulation of NHEJ and HDR activity using small molecule drugs has been widely studied; however, increases in HDR-mediated gene targeting have been inconsistent and modest at best[7]. Recently, Cpf1, a Class 2 CRISPR endonuclease with a unique staggered cleavage pattern that forms sticky ends, has emerged as a promising tool that may be able to stimulate HDR-independent directional repair[8], but its utility for gene knock-in has been impeded by its low cutting efficiency compared to Cas9-based enzymes[9]. These limitations have driven the pursuit of editing strategies that employ non-homologous end joining (NHEJ) instead. By generating indels and genomic deletions, NHEJ-based approaches utilize reframing or exon skipping to bypass mutations that disrupt the *DMD* open reading frame (ORF) and recover a truncated but partially functional dystrophin protein[10-19]. These genomic modifications thereby achieve the clinically milder phenotype of Becker muscular dystrophy (BMD), in which voluntary muscle function is considerably ameliorated and life expectancy is prolonged[20]. Nonetheless, cardiomyopathy remains the primary cause of death in BMD patients[21], suggesting that full-length dystrophin restoration is necessary for complete rescue of cardiomyocyte function.

Recently, homology-independent targeted integration (HITI) was shown to enable NHEJ-based site-specific donor insertion in mouse cells *in vitro* and *in vivo*[22]. This technique was also

successfully applied to human cells, producing higher knock-in efficiencies than HDR-mediated gene targeting[23, 24]. Through the NHEJ pathway, HITI utilizes guide RNA target sequences that flank the exogenous repair substrate to mediate donor integration into the genomic target site[22]. To address the shortcomings of *DMD* reframing and exon skipping, we harness HITI-based targeted knock-in to restore full-length dystrophin expression. In this work, we propose and characterize a novel AAV-based excise-and-replace strategy that employs dual guide RNA directed deletion and repair of a target exon to achieve full correction of the *DMD* gene in a human cardiomyocyte model of DMD. We demonstrate that our approach results in targeted donor insertion and produces complete *DMD* transcripts in our model system. We also optimize our method by leveraging a third guide RNA to drive higher gene targeting efficiencies. Thus, our HITI-based gene replacement strategy is well-suited for correction of various deletion and small mutation genotypes of DMD and can potentially be used to repair mutations of other recessive disorders.

5.2 Results

Efficient gene targeting in hESC-CMs by AAV-HITI

To evaluate the capacity of HITI to mediate gene targeting in human cardiomyocytes, we devised an AAV-based editing scheme to tag the cardiac Troponin T (*TNNT2*) locus with a fluorescent reporter in cardiomyocytes derived from human embryonic stem cells, or hESC-CMs, (Figure 1A). We designed a self-complementary AAV HITI donor vector (AAV-cTnT-HITI) containing a guide RNA that targets the last exon of the *TNNT2* gene (cTnT-gRNA) under a U6 promoter and a promoterless P2A-GFP construct flanked by target sequences complementary to the cTnT-gRNA. Upon co-delivery of AAV-cTnT-HITI and an AAV vector encoding SpCas9 driven by a constitutive hybrid EF1α/HTLV (nEF) promoter (AAV-SpCas9) into hESC-CMs, expression from both vectors results in SpCas9-mediated cleavage of the genomic target site and the donor target sequences, thereby enabling NHEJ-based insertion of the P2A-GFP donor under the control of the endogenous cardiac muscle specific *TNNT2* promoter. Correct gene targeting thus produces GFP expression only in cardiomyocytes.

To optimize the knock-in efficiency of our strategy, we screened several candidate cTnT-gRNAs for on-target cleavage by Surveyor nuclease assay. All gRNAs were cloned into the U6-gRNA expression cassette of a plasmid that co-expresses SpCas9 and GFP and transfected into HEK293T cells. PCR was used to amplify the target region from genomic DNA 72 hours post-transfection, and heteroduplexes formed from PCR products were treated with Surveyor nuclease. Cleaved DNA fragments of expected size were detected for all cTnT-gRNAs (Figure S1A). We also characterized the infectivity of several natural AAVs for hESC-CMs to identify a serotype with maximal transduction efficiency. We infected hESC-CMs with AAV1, 2, 6, and 9 packaged with a self-complementary CAG-eGFP reporter at a multiplicity of infection (MOI) of 10,000 and quantified eGFP-positive cells 6 days post-infection by flow cytometry (Figure S1B). Remarkably, AAV6 resulted in eGFP expression in 88% treated cells, reaffirming its robust transduction of human cardiomyocytes as previously described[25, 26]. We therefore utilized recombinant AAV6 vectors to deliver SpCas9 and HITI donors for all gene knock-in and correction experiments.

Using cTnT-g4, we next applied AAV-SpCas9 and AAV-cTnT-HITI to knock the P2A-GFP cassette into the human *TNNT2* locus. We first validated the functionality of our vectors by co-transfection of AAV-SpCas9 and AAV-cTnT-HITI plasmids into HEK293T cells. Targeted donor insertion was detected by 5' and 3' junction PCR on genomic DNA 72 hours post-transfection (Figure 1B), recapitulating NHEJ-based gene targeting as previously reported[22]. AAV6-mediated co-delivery of SpCas9 and self-complementary cTnT-HITI vectors also enabled 5' and 3' integration of the targeted cassette in hESC-CMs, as confirmed by PCR analysis and Sanger sequencing 2 weeks post-infection (Figure 1B). Furthermore, immunostaining showed co-localization of cTnT and GFP in treated cells, indicating precise and complete knock-in of the P2A-GFP construct by HITI in modified cardiomyocytes (Figure 1C). Finally, we compared the gene targeting efficiencies of self-complementary and single-stranded AAV-cTnT-HITI vectors using flow cytometry analysis on treated cells 7 days post-infection. When co-delivered with AAV-SpCas9, single-stranded HITI donors outperformed their self-complementary counterpart and yielded 20% more GFP+ cells than the donor only control (Figure 1D, S1C). *Bona fide* genomic modification was detected by 5' junction PCR only in the presence of the AAV-SpCas9 vector (Figure S1D), suggesting that GFP expression in cells treated with donor vector alone was driven by low promoter activity from the AAV ITR[27, 28]. We thus provide proof-of-concept demonstration that HITI efficiently integrates transgenes at target loci in human cardiomyocytes.

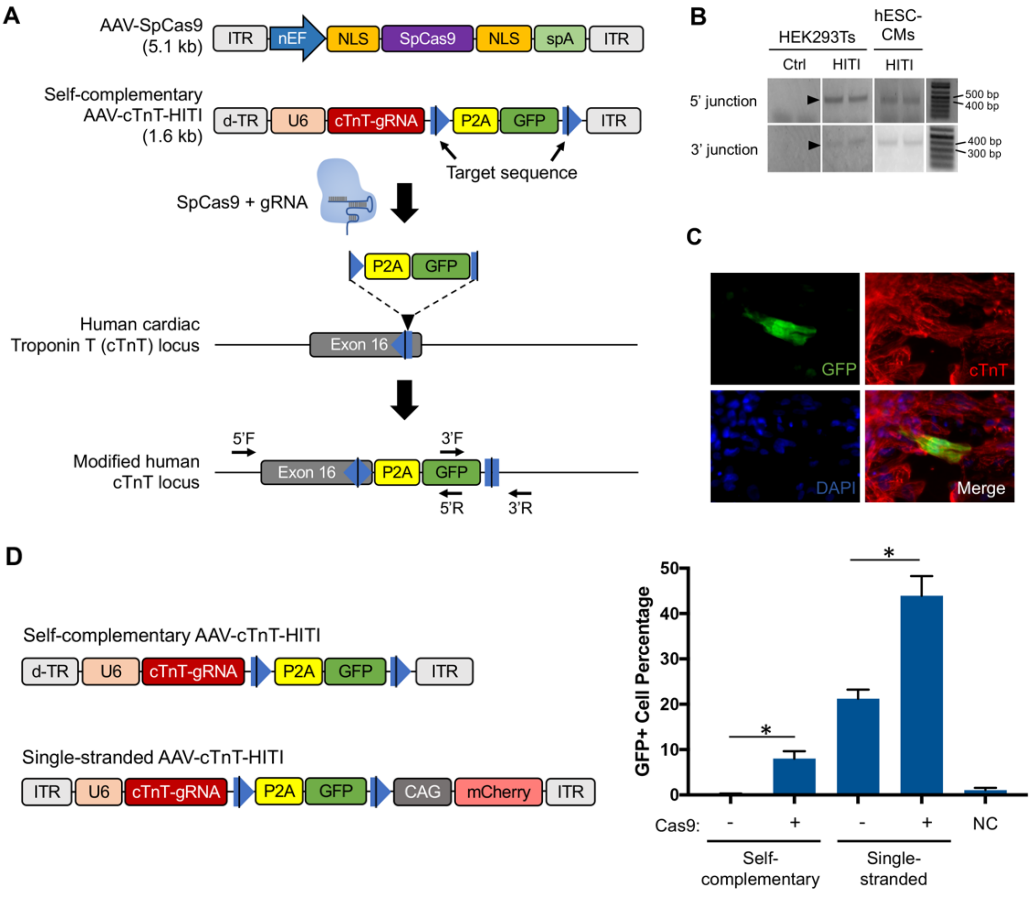


Figure 5.1. AAV-HITI mediates efficient gene targeting in hESC-CMs.

A) Schematic of human *TNNT2* targeting strategy. Co-delivery of AAV-SpCas9 and AAV-cTnT-HITI enables NHEJ-based targeted integration of the P2A-GFP donor construct into the last exon of the endogenous *TNNT2* locus, yielding GFP expression in correctly modified cardiomyocytes. ITR: inverted terminal repeat, d-TR: mutated terminal repeat for self-complementary vectors, nEF: hybrid EF1α/HTLV promoter, NLS: nuclear localization signal, spA: synthetic polyA signal, 5' F/R: forward/reverse primers for 5' junction genotyping, 3' F/R: forward/reverse primers for 3' junction genotyping

B) Targeted knock-in by HITI in HEK293Ts and hESC-CMs. PCR genotyping of 5' and 3' junctions confirmed targeted donor integration by plasmid co-transfection and AAV6-mediated delivery of AAV-SpCas9 and AAV-cTnT-HITI vectors using cTnT-g4 in HEK293T cells and hESC-CMs, respectively. 5' junction expected band size: 446 bp. 3' junction expected band size: 342 bp. Ctrl: untreated control.

C) Co-localization of GFP and cTnT in HITI-modified cardiomyocytes. Treated cells were co-stained for GFP, cTnT, and DAPI by immunocytochemistry. Representative images are shown. Overlay of GFP and cTnT indicates correct targeting of the P2A-GFP donor construct to the *TNNT2* locus.

D) Comparison of gene targeting efficiencies using self-complementary and single-stranded AAV-cTnT-HITI vectors. Schematic of donor vector design is shown (left). Percentage of GFP positive cell was quantified by flow cytometry to determine frequency of targeted knock-in (right). NC: untreated negative control.

Correction of *DMD* in ΔE51 DMD-CMs by dual-cut HITI

To harness HITI for full-length dystrophin restoration, we aimed to develop an approach that can be broadly applied to various DMD-causing mutations. Deletion of one or more exons and small mutations, including indels and point mutations, in the *DMD* gene constitute 68% and 20% of known DMD genotypes, respectively[29]. While the previously reported HITI approach, hereafter termed single-cut HITI, can be leveraged for targeted insertion of missing exons to treat deletion genotypes, it cannot readily correct small mutations that require gene replacement to restore *DMD* expression. Thus, we created an excise-and-replace strategy, herein termed dual-cut HITI, that employs dual gRNA directed excision and repair of a target exon to correct exon deletions and small mutations in the *DMD* gene (Figure 2A). Using target sites in the upstream and downstream intronic regions, a mutant exon can be deleted and replaced with a HITI donor template that contains the wildtype exon and target sequences oriented to drive donor integration in the forward direction (Figure 2A, left). To adapt dual-cut HITI for deletion genotypes, an exon immediately adjacent to the missing exon(s) is targeted for excision and substituted with the complementary DNA (cDNA) of both the target and missing exons, yielding a complete *DMD* transcript (Figure 2A, right). Importantly, while our strategy calls for exon-specific customization, it is impartial to the sequence identity of the mutant or missing genomic region, enabling a single guide RNA pair and HITI donor to potentially correct a range of genotypes for a given target exon.

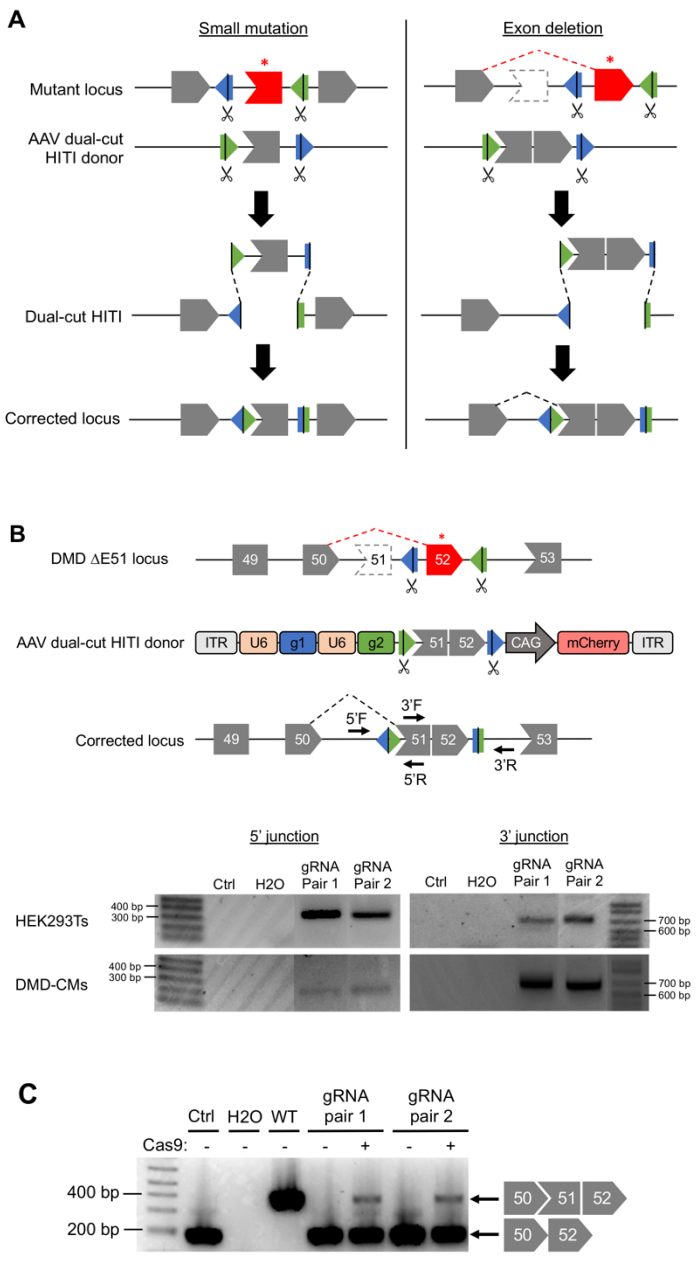


Figure 5.2. AAV dual-cut HITI corrects *DMD* locus and generates wildtype dystrophin transcripts in ΔE51 DMD-CMs.

A) Schematic of AAV dual-cut HITI mediated repair of mutant *DMD* loci. Target sites upstream and downstream of a target exon direct excision of the intervening genomic sequence. Cleavage of both the genome and the AAV dual-cut HITI donor vector enable repair of the chromosomal DSB by insertion of the donor fragment. This strategy can be applied to small *DMD* mutations (left) or *DMD* exon deletions (right).

B) Targeted donor integration in HEK293Ts and ΔE51 DMD-CMs by dual-cut HITI. PCR genotyping of 5' and 3' junctions confirmed targeted donor integration by plasmid co-transfection and AAV6-mediated delivery of AAV-SpCas9 and AAV-dual-HITI vectors in HEK293T cells and ΔE51 DMD-CMs, respectively. 5' junction expected band size: 329 bp. 3' junction expected band size: 696 bp. gRNA pair 1: DMD-g1 + DMD-g3. gRNA pair 2: DMD-g1 + DMD-g4. Ctrl: untreated control. H2O: PCR control with no DNA template.

C) Corrected *DMD* transcripts containing missing exon in treated ΔE51 DMD-CMs. RT-PCR using primers specific to *DMD* exons 50 and 52 resulted in amplicons corresponding to wildtype dystrophin transcripts in ΔE51 DMD-CMs treated with AAV dual-cut HITI. Sequencing confirmed accurate splicing of corrected transcripts. Only mutant transcripts were detected in donor only controls. *DMD* ΔE51 expected band size: 182 bp. WT *DMD* expected band size: 415 bp. WT: wildtype hESC-CM control.

In a proof-of-principle experiment, we applied our excise-and-replace strategy to correct human cardiomyocytes derived from DMD induced pluripotent stem cells (iPSCs) that harbor a deletion of *DMD* exon 51 (ΔE51 DMD-CMs). We designed candidate guide RNAs with target sites located ~100-200 base pairs (bp) upstream (DMD-g1, DMD-g2) or downstream (DMD-g3, DMD-g4) of *DMD* exon 52. Upstream and downstream guide RNAs were then paired to induce excision of the target exon and cloned into plasmids that also express SpCas9 for co-transfection into HEK293T cells. Genomic DNA was assayed for genomic deletion by PCR genotyping after 72 hours (Figure S2A). Guide RNA pairs 1 (DMD-g1 + DMD-g3) and 2 (DMD-g1 + DMD-g4)

were superior in generating modified *DMD* alleles harboring the desired deletion and thus used for HITI-based correction of our DMD cardiomyocyte model. Next, we constructed a dual-cut HITI donor template that contains the cDNA of *DMD* exons 51-52, flanked by ~100-200 bp fragments from the introns immediately upstream of exon 51 and downstream of exon 52 to retain the endogenous splice acceptor and splice donor sites. As targeted knock-in may lead to indel generation at the 5' and 3' genome-donor junctions, we incorporated these intronic "arms" to create non-coding spacer sequences that reduce the likelihood of splice site mutagenesis. Target sites of the *DMD* guide RNA pairs were then appended to each end of the donor template and positioned to favor forward donor insertion. Finally, AAV donor vectors (AAV-dual-HITI) encoding the promoterless dual-cut HITI donor template and multiplex expression of the *DMD* guide RNA pairs were assembled. Targeted donor integration was detected in genomic DNA 72 hours after plasmid co-transfection of AAV-SpCas9 and AAV-dual-HITI into HEK293T cells, as determined by 3' junction PCR analysis and verified by Sanger sequencing (Figure 2B). AAV6-mediated delivery of AAV-SpCas9 and AAV-dual-HITI also resulted in the desired modification in ΔE51 DMD-CMs 12 days post-infection (Figure 2B), demonstrating that AAV-based dual-cut HITI can facilitate site-specific gene replacement in human cardiomyocytes.

To characterize *DMD* expression in cells modified by dual-cut HITI, we performed an end-point reverse transcription PCR (RT-PCR) that assays for corrected transcript splicing in the genomic target region. Total RNA was collected from ΔE51 DMD-CMs treated with dual-cut HITI 12 days post-infection and reverse transcribed to cDNA that served as a template for RT-PCR using primers specific to *DMD* exons 50 and 52 (Figure 2C). Of note, desired gene replacement versus insertion into either of the genomic target sites can be differentiated by PCR amplicon size in our analysis. For both *DMD* guide RNA pairs, dual-cut HITI yielded PCR products corresponding in size to the wildtype *DMD* mRNA, which were sequence verified. In contrast, donor only controls produced only uncorrected transcripts. Our findings show that the splice sites of our dual-cut HITI donor templates remain intact upon targeted knock-in, hence enabling accurate splicing of the restored *DMD* transcript. We also analyzed mRNA expression in treated cells using primers specific to *DMD* exons 49 and 53, which revealed considerable deletion of target exon 52 (Figure S2B). Our observation is consistent with prior reports of guide RNA multiplexing[15, 16, 19], indicating that genomic deletion is not specific to dual-cut HITI. Therefore, our *DMD* expression studies confirm that dual-cut HITI yields correctly spliced *DMD* transcripts in treated ΔE51 DMD-CMs and suggest that genomic deletion formed by dual guide RNA directed cleavage is the dominant on-target editing event.

Triple-cut HITI strategy improves knock-in efficiency in HEK293Ts

To improve the gene targeting efficiency of our excise-and-replace strategy, we developed a novel triple-cut HITI-based method to direct targeted donor integration in the event of precise genomic deletion mediated by target sites on the same DNA strand. Our design introduces the expression of a third guide RNA, termed gRNA-A, that is complementary to the new target sequence generated upon error-free deletion, triggering further editing and potentially increasing the frequency of desired donor insertion (Figure 3A). We also explore two different architectures for our triple-cut HITI vectors, one in which the donor is flanked by the original genomic target sites (3A, left) and another in which it is flanked by target sites for gRNA-A (Figure 3A, right).

For both constructs, target sequences are oriented to drive forward donor integration, as reverse knock-in restores target sites that are subject to re-cleavage by Cas9.

To compare the knock-in frequencies of targeting approaches discussed thus far, we devised an editing scheme that tags the endogenous human *GAPDH* locus with an IRES-mCherry construct in HEK293T cells (Figure 3B). We first assessed the on-target activity of two candidate guide RNA pairs, each with protospacer adjacent motif (PAM) sequences on the same genomic strand within the *GAPDH* 3'UTR, by plasmid co-transfection into HEK293T cells and PCR genotyping (Figure S3A). Due to its superior editing efficiency, guide RNA pair 2 (GAPDH-g3, GAPDH-g4) was selected for subsequent *GAPDH* targeting experiments. Next, we designed a third gRNA, GAPDH-gA, comprised of the first 17 bases of GAPDH-g3 and the last 3 bases of GAPDH-g4. Upon perfect genomic deletion mediated by GAPDH-g3 and GAPDH-g4, the target sequence for GAPDH-gA emerges and renews editing of *GAPDH* by Cas9. Single-, dual-, and triple-cut HITI donor plasmids harboring the IRES-mCherry repair template and expressing the corresponding *GAPDH* guide RNAs were co-transfected with an SpCas9-T2A-GFP expression plasmid into HEK293Ts. Four days post-transfection, double positive cells (GFP+ mCherry+) and GFP positive cells were quantified by flow cytometry to determine the percentage of correctly modified cells in the transfected population (Figure 3B). Relative to dual-cut HITI, triple-cut HITI indeed improved efficiency of gene targeting to the *GAPDH* locus in HEK293Ts.

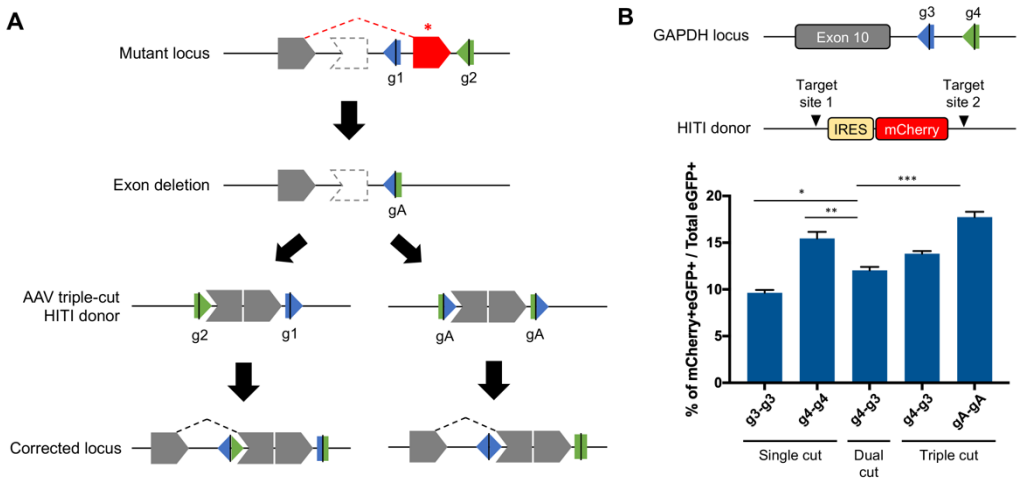


Figure 5.3. Triple-cut HITI enhances gene targeting efficiency in HEK293Ts.

A) Schematic of proposed triple-cut HITI strategy for *DMD* gene correction. Genomic target sites (g1 and g2) located on the same DNA strand can recombine to generate a new target sequence (gA) upon perfect deletion of the intervening exon. A third gRNA is supplied to direct cutting of the gA target site. Co-delivery of an AAV triple-cut HITI donor vector, containing the donor fragment flanked by either g1 and g2 (left) or by gA (right), induces targeted knock-in upon cleavage of the genome and the AAV vector. Forward donor insertion into the chromosomal DSB repairs the mutant *DMD* locus.

B) Gene targeting efficiency of single-, dual-, and triple-cut HITI at the *GAPDH* locus in HEK293Ts. HITI donor plasmids harboring the IRES-mCherry repair template and expressing the corresponding *GAPDH* guide RNAs were co-transfected with an SpCas9-T2A-GFP expression plasmid into HEK293Ts. The percentage of corrected modified cells (GFP+mCherry+) over the transfected population (GFP+) was quantified by flow cytometry.

Guide RNA sequences	
cTnT-g1	TTTAGCCTTCCCGCGGGTCT
cTnT-g2	GGTGACTTTAGCCTTCCCGC
cTnT-g3	GCGGGAAGGCTAAAGTCACC
cTnT-g4	AGGCTAAAGTCACCGGGCGC
DMD-g1	GGAGACATTCCGGAGTACCT
DMD-g2	TTTATAGAAAACAATTTAAC
DMD-g3	AGTTGCAAAGCATGCATTGA
DMD-g4	TTTATTAAAAGGTAACATTA
GAPDH-g1	GAGAGAGACCCTCACTGCTG
GAPDH-g2	GCCATGTAGACCCCTTGAAG
GAPDH-g3	CTTCCTCTTGTGCTCTTGCT
GAPDH-g4	ACAAGTAACTGGTTGAGCAC
Primers for PCR genotyping	
cTnT indel F	GTCATTCTCGCCTAGCCCTC
cTnT indel R	AGGCCAGCTCCCCATTTC
cTnT HITI 5'F	TTTGGCACCCCAGTCCTA
cTnT HDR 5'F	GGAAGCTTCTCCTTGGCTTC
cTnT HITI 5'R	AAGTCGTGCTGCTTCATGTG
cTnT HITI 3'F	ACTACAACAGCCACAACGTCTATATCA
cTnT HITI 3'R	GAGCGAGGAGCAGATCTTTG
DMD geno F	TCGGTTCTGTGCTTTTTCTG
DMD geno R	GGCATCAAAGTAGCGGAGAG
DMD HITI 5'F	TCAAGGCATTCAGACAGTGG
DMD HITI 5'R	TTGGAGATGGCAGTTTCCTT
DMD HITI 3'F	CGGGCTTGGACAGAACTTAC
DMD HITI 3'R	GGCATCAAAGTAGCGGAGAG
DMD E50 RT-F	CTGAGTGGAAGGCGGTAAAC
DMD E52 RT-R	GCCTCTTGATTGCTGGTCTT
DMD E49 RT-F	AACAACCGGATGTGGAAGAG
DMD E53 RT-R	CCTCCTTCCATGACTCAAGC
GAPDH geno F	CAACGAATTTGGCTACAGCA
GAPDH geno R	CTCCCCAGCAAGAATGTCTC

Table 5.1 Guide RNA sequences and genotyping primers.

5.3 Discussion

CRISPR/Cas9-based genome editing strategies that reframe or skip exons in the *DMD* gene have led the revolution of treatments for DMD, due to their potential for affording long-term therapeutic benefit. These approaches employ NHEJ-mediated repair to recreate in-frame dystrophin proteins through frameshifting, splice site mutagenesis, and excision of mutant or out-of-frame exons, resulting in significant improvements in skeletal muscle function in mouse[5, 10, 14, 18, 19] and canine[16] models of DMD. Partial phenotypic rescue has also been achieved in human cardiomyocytes derived from edited DMD iPSC lines[13, 15, 17]. Nonetheless, the restoration of full-length dystrophin protein bears untapped potential for overcoming functional abnormalities in cardiac muscle with enhanced potency. Furthermore, there is a need for therapeutic interventions

that directly target human cardiomyocytes, as these strategies hold greater promise for clinical translation in human DMD patients.

In this study, we establish that AAV-mediated HITI enables efficient targeted transgene integration in human cardiomyocytes *in vitro*, demonstrating its utility for gene correction in post-mitotic cardiac muscle in which DNA lesions are primarily repaired by error-prone NHEJ[30]. The low efficiency of HDR-mediated repair in non-dividing cells has driven the development of gene targeting approaches that harness the dominant NHEJ pathway to guide homology-independent targeted integration (HITI)[22, 23, 31, 32]. Specifically, the inclusion of a target site in the donor template and simultaneous cleavage of the donor and the chromosome result in ligation at the genomic DSB in human cells *in vitro*[23, 32]and in mice *in vivo*[22]. In this work, we show that HITI can efficiently insert a promoterless P2A-GFP cassette into the *TNNT2* locus at the intended target site in hESC-CMs. Of note, the frequency of site-specific forward donor insertion is likely underreported in our analysis, as detection of GFP signal by flow cytometry requires seamless, in-frame fusion while indel formation at the 5' genome-donor junction would ablate expression of the downstream reporter gene. Overall, our findings confirm that HITI can be readily applied to human cardiomyocytes to mediate targeted gene knock-in using a non-homologous repair substrate.

HITI has been previously shown to produce higher gene targeting efficiencies than HDR in mouse heart *in vivo*[22]. Interestingly, HDR donor vectors improved knock-in frequency in hESC-CMs compared to their HITI counterpart (Figure S1D, S1E), suggesting that repair occurs during the S and G2 phases of cell cycle. Nonetheless, this discrepancy may be explained by the retained, though limited, proliferative capacity of human cardiomyocytes that are generated from small molecule based differentiation of pluripotent stem cell populations[33, 34]. Evaluation in large animal models or more mature human cardiomyocytes *in vitro* may better predict knock-in rates achieved by HITI in human patients. Integration methods based on microhomology-mediated end joining (MMEJ) and homology-mediated end joining (HMEJ) have also been engineered, with demonstrated efficacy in various cell types *in vitro* and *in vivo*[35-37]. Similar to HITI repair templates, MMEJ and HMEJ donors are targeted for cleavage; however, they additionally contain microhomology (5-25 bp) or long homology (800 bp) arms that facilitate forward knock-in. By granting directionality, these strategies could potentially improve the ratio of desired insertion over other genomic rearrangements in human cardiomyocytes and thus warrant further investigation.

In seeking to adapt HITI for the treatment of DMD, we created a new excise-and-replace approach that utilizes dual gRNA directed excision of a target exon and insertion of an exogenous donor fragment to correct DMD-causing mutations that are amenable to targeted gene knock-in or replacement. In addition to exon deletions, our strategy can be applied to small *DMD* mutations, whereas classical HITI is limited to targeted knock-in and thus cannot remove a mutant exon. Due to the absence of homology arms to guide transgene integration in a specific direction, we placed target sites in the AAV dual-cut HITI donor on the strand opposite of that of the genomic target sites. This design enables restoration of intact target sites in the genome upon error-free reverse knock-in of the donor, thereby providing opportunity for further cleavage by SpCas9. Here, we report the repair of the mutant *DMD* locus using an AAV dual-cut HITI donor vector in a human cardiomyocyte model of DMD. Furthermore, targeted integration results in expression of corrected *DMD* transcripts, indicating that the integrity of the splice sites encoded in the donor was preserved during the repair process.

With optimization of gene targeting efficiency, AAV dual-cut HITI holds significant promise for gene replacement therapy. The low knock-in rates observed in our study may be attributed to competing repair mechanisms. For example, we hypothesize that paired gRNAs must exhibit similar on-target activity in order to drive deletion of the intervening genomic sequence. Repair of a single DSB prior to cleavage of the other target site can result from differences in cutting efficiency between the two gRNAs, leading to indel formation or donor knock-in at one of the sites that renders the locus refractory to correction (data not shown). Thus, screening of gRNA pairs for their ability to induce deletion is critical for dual-cut HITI. While excision of the target exon is an essential step of our repair strategy, our data suggest that permanent genomic deletion outcompetes donor insertion upon cleavage of both target sites. Importantly, due to pre-existing early termination codons that disrupt the *DMD* ORF, we do not expect deletion of the targeted exon to exacerbate phenotype in treated cells. Interestingly, several studies have shown that the majority of deletion events are perfect deletions mediated by error-free joining of the two broken DNA ends[19, 38].

To increase the frequency of donor insertion, we proposed a triple-cut HITI strategy that enables editing to resume when deletion is mediated by genomic target sites on the same DNA strand. As error-free excision creates a new target site, we provided a third gRNA that directs cleavage of this sequence to re-stimulate DNA repair. We demonstrate that triple-cut HITI is indeed superior to dual-cut HITI at inserting a transgene cassette into the human *GAPDH* locus. We also reason that the activity of the third gRNA for the chimeric target site determines the added benefit conferred by triple-cut HITI. As this sequence is recombined from the original genomic target sites and thus context-dependent, screening of gRNA combinations may be necessary to optimize the gene targeting efficiency of triple-cut HITI for a given target exon.

HITI-based excise-and-replace repair possesses several attractive properties for correction of DMD and possibly other recessive disorders. First, the vast diversity of DMD-causing mutations[29, 39] calls for simplified therapeutic strategies that can be used to treat large patient cohorts harboring a range of genotypes. Our approach only requires the design of exon-specific gRNA pairs, which can be coupled with different HITI donors to correct upstream or downstream deletions, deletions of varying length, indels, or point mutations for a given target exon. In contrast, CRISPR base editors, which have been devised to precisely edit single nucleotides independently of endogenous DNA repair pathways, can only target applicable point mutations within a limited editing window and necessitate optimization for each unique mutation within a target exon[40, 41]. Second, our approach utilizes gRNA pairs that target intronic regions and can tolerate the introduction of indels at genome-donor junctions. As precise in-frame knock-in is not necessary for correction, target sites and AAV dual-cut or triple-cut HITI donors can be designed with great flexibility. Finally, while donor insertion is the preferred editing event, genomic deletion can, in some cases, fortuitously convert DMD to the milder BMD by skipping the target exon and restoring the *DMD* ORF. For compatible DMD genotypes, it is therefore possible to design editing schemes that rescue full-length dystrophin expression upon targeted knock-in and generate partially functional dystrophin upon genomic deletion.

In summary, we present novel excise-and-replace strategies for gene targeting in post-mitotic cells and corroborate their ability to mediate site-specific gene integration and gene replacement in human cardiomyocytes. Our results establish the suitability of AAV dual-cut and triple-cut HITI for correction of the mutant *DMD* locus, as demonstrated by successful restoration

of full-length dystrophin expression in an *in vitro* human disease model of DMD. With improvements to knock-in efficiency, these strategies potentially offer greater clinical benefit, particularly in cardiac muscle, than reframing and exon skipping for DMD patients by rescuing complete dystrophin function in corrected cells. Lastly, HITI-based excise-and-replace repair may be a promising treatment for other recessive disorders, particularly those that call for early intervention in dividing cells and those for which AAV's carrying capacity limits the efficacy of therapeutic transgene delivery.

5.4 Future Directions

Because triple-cut HITI increased gene targeting to the *GAPDH* locus in HEK293Ts as previously discussed, we will apply this approach to ΔE51 DMD-CMs using AAV vectors. To confirm that AAV dual-cut and triple-cut HITI can produce full-length dystrophin, we will probe for restored expression of the target protein by Western blot analysis on treated cells. Finally, we will characterize on-target editing using a recently reported transposon-based unidirectional sequencing method[19, 42]. Adapted from Illumina's Nextera library preparation platform, this approach entails TN5 transposon-mediated "tagmentation" of genomic DNA[43] and PCR-based target enrichment using an "anchor" primer specific to the genomic target region and a primer specific to the exogenous tag sequence introduced by the transposon. A second PCR is then used to add experimental barcodes and Illumina adaptors for Illumina Miseq sequencing. We expect to detect a range of genomic modification events including indels, deletions, inversions, and AAV genome integrations (partial and full length), as diverse editing outcomes have been observed in loci targeted with Cas9 gRNA pairs[19, 42]. In fact, gRNA target sites incorporated into AAV HITI donors may lead to more heterogeneity in genomic rearrangements through insertion of cleaved vector fragments. For example, we predict the knock-in of undesired AAV genome fragments, such as those containing the U6-gRNA expression cassettes or the CAG-mCherry reporter cassette, due to the recombinogenic nature of the AAV ITRs[19, 44]. Moreover, this analysis will enable quantification of correction efficiency and can further be applied to cDNA from treated cells to measure resulting transcript changes.

5.4 Materials and Methods

Plasmid construction

For SpCas9 and guide RNA co-expression, complementary oligonucleotides encoding guide RNAs were synthesized, annealed, and individually cloned into pX458 (Addgene #48138) using the BbsI restriction sites. For *TNNT2* targeting experiments, the pAAV-nEFCas9 (Addgene #87115) plasmid was used for AAV-mediated expression of SpCas9. The 5' LTR of the nEF promoter in pAAV-nEFCas9 was removed to generate the EF1α shortened (EFS) promoter, and the resulting pAAV-EFSCas9 plasmid was applied in *DMD* targeting experiments.

To construct the AAV-cTnT-HITI donor vectors, the P2A-GFP donor fragment was first amplified from the Oct4-2A-eGFP-PGK-Puro plasmid (Addgene #31938) by PCR using primers with overhangs that encode the gRNA target site. This amplicon and the corresponding U6-gRNA cassette amplified from pX458 were then spliced by overlap extension PCR. The resulting

construct was ligated into a self-complementary AAV backbone[45] by restriction cloning. To generate the single-stranded equivalent, the same construct and a CAG-mCherry expression cassette were cloned into a single-stranded AAV backbone derived from pX601 (Addgene #61591) using unique restriction sites for single-tube assembly.

To build the AAV-DMD-HITI donor vectors, *DMD* intron 50 with exon 51 and *DMD* exon 52 with intron 52 were first amplified by PCR from genomic DNA extracted from HEK293Ts. The two fragments were then spliced by overlap extension PCR using primers with overhangs that encode the gRNA target sites. The resulting construct and the corresponding U6-gRNA expression cassettes were then ligated into a single-stranded AAV backbone containing a CAG-mCherry expression cassette by restriction cloning in a single-tube reaction.

For the construction of the *GAPDH* targeting vectors, the IRES-mCherry donor fragment was first amplified by PCR using primers with overhangs that encode the gRNA target sites. The resulting amplicon and the corresponding U6-gRNA expression cassettes were then ligated into the pX601 backbone digested with XbaI and NotI by restriction cloning.

AAV vector production and purification

All recombinant AAV vectors were packaged in HEK293T cells, as previously described[46, 47]. In summary, recombinant AAV vectors were produced by triple transient transfection of a helper plasmid encoding adenoviral helper genes, an AAV helper plasmid encoding AAV *rep* and *cap* genes, and a transfer plasmid containing donor constructs flanked by AAV ITRs, into HEK293T cells using polyethylenimine (PEI) (Polysciences, 23966-1). Culture media was changed 14-16 hours post-transfection to reduce PEI-associated toxicity. At 72 hours post-transfection, cells were dissociated from culture dishes using a cell scraper and pelleted by centrifugation at 1500 *g* for 2.5 min. Cells were then resuspended in AAV lysis buffer (50 mM Tris base, 150 mM NaCl, pH 8.2-8.5) and lysed with 3 freeze/thaw cycles using a dry ice / ethanol bath and a 37°C water bath. Supernatant was collected following centrifugation at 10,000 rpm for 10 min in a tabletop centrifuge. For additional purification of vector preparations (optional), clarified cell lysate and culture supernatant can be combined with a solution of 40% polyethylene glycol (PEG) 8000 (Sigma-Aldrich, P2139) and 2.5 M NaCl at a 4:1 ratio, incubated at 4°C overnight, centrifuged at 2500 *g* for 30 min to harvest precipitated AAV particles, and resuspended in AAV lysis buffer[48]. Crude lysates were treated with 10 U benzonase (Sigma-Aldrich, E8263) per mL of lysate for 30 min at 37°C.

For AAV purification by iodixanol density centrifugation, discontinuous gradients comprised of 15%, 25%, 40%, and 54% iodixanol layers were set up in Optiseal tubes (Beckman Coulter, 362185) using OptiPrep (Axis-Shield, AVS-1114542). Benzonase-treated lysates were then loaded onto iodixanol gradients and centrifuged at 174,000 *g* for 2 hr at 18°C. Tubes were punctured using 21-gauge needles attached to a 1 mL syringe at the 40% / 54% interface, and the 40% iodixanol fraction containing AAV particles was collected. Vector preps were then buffer exchanged and concentrated into PBS with 0.001% or 0.00025% Tween 20 (Sigma-Aldrich, 9416) using Amicon filtration (EMD Millipore, UFC910024) at 3000 *g*. DNase-resistant (Sigma-Aldrich, 04536282001) viral genomic titers were measured by real-time qPCR using SYBR Green I (Invitrogen, S7567) and the CFX96 Real-Time PCR cycler (Bio-Rad, 1855195).

Cell culture

H9 embryonic stem cells (NIH Stem Cell Registry #0062) and DMD ΔE51 induced pluripotent stem cells (Stanford University Cardiovascular Institute Biobank, SCVI630) were cultured in 6-well plates coated with Matrigel (Corning, 354277) and maintained in mTeSR1 media (STEMCELL Technologies, 85850) with 1% penicillin-streptomycin (Gibco, 15140122). At confluency, cells were dissociated with Accutase (STEMCELL Technologies, 07920) and seeded in mTeSR1 supplemented with 5 μM ROCK inhibitor, Y-27632 (Selleckchem, S1049). Human embryonic kidney 293T (HEK293T) cells from the American Type Culture Collection (Manassas, CRL-3216) were cultured in DMEM (GIBCO, 12800082) with 10% fetal bovine serum (Invitrogen, 10437028) and 1% Antibiotic-Antimycotic (GIBCO, 15240062) at 37°C and 5% CO_2.

Cardiomyocyte differentiation

H9 ESCs were differentiated into cardiomyocytes (hESC-CMs) using small molecule inhibitors of Wnt signaling[34]. Briefly, H9 ESCs were dissociated into single cells with Accutase at 37°C for 5 min and seeded at 100,000 cells/cm^2 in mTeSR1 supplemented with 5 μM Y-27632 (day -3) for 24 hours. Cells were then cultured in mTeSR1 with daily media changes. On day 0, cells were treated with 5.5 μM CHIR99021 (Selleckchem, S2924) in RPMI 1640 (Gibco, 11875093) for 24 hours, after which culture media was changed to RPMI supplemented with B27-insulin (Gibco, A1895601) . On day 3, conditioned media prepared from culture media and fresh RPMI/B27-insulin (1:1 ratio) was added with 5 μM IWP2 (Tocris, 3533) to cells. Finally, media was changed to RPMI/B27-insulin on day 5 and then to RPMI/B27 (Gibco, 17504044) on day 7. Cardiac progenitor cells were maintained in RPMI/B27 with media changes every 3 days to complete differentiation into cardiomyocytes.

DMD ΔE51 iPSCs were differentiated into cardiomyocytes (ΔE51 DMD-CMs) using the STEMdiff cardiomyocyte differentiation kit (STEMCELL Technologies, 05010). At confluence, DMD ΔE51 iPSCs were dissociated into single cells with Accutase at 37°C for 5 min and passaged at a 1:1.5 ratio into a 12-well plate coated with Matrigel. Cells were treated according to the manufacturer's protocol.

For AAV infection experiments, day 6 cardiac progenitors were dissociated with Accutase for 5-10 min, resuspended in RPMI supplemented with 20% FBS and 5 μM Y-27632, and seeded into plates coated with gelatin (Sigma, G1393). Cultures were changed to RPMI/B27 or STEMdiff cardiomyocyte maintenance basal media the next day and maintained as previously described to complete differentiation. Recombinant AAV vectors were added to cultures on day 14 or later for transduction of fully differentiated cardiomyocytes. Alternatively, day 6 cardiac progenitors can be resuspended in RPMI supplemented with 30% FBS, 10% DMSO, and 5 μM Y-27632 for long-term cryostorage. To thaw, cells were transferred to RPMI/B27 and centrifuged at 200 *g* for 5 min. Cells were then resuspended in RPMI with 20% FBS and 5 μM Y-27632 and seeded into plates coated with gelatin (Sigma, G1393).

Guide RNA design and validation

Candidate guide RNAs (gRNAs) were identified using the guide design tool provided by F. Zhang's laboratory (http://crispr.mit.edu/). Plasmids co-expressing SpCas9 and individual

gRNAs were transfected into HEK293Ts using PEI, and genomic DNA was extracted from treated cells 72 hours post-transfection to analyze on-target editing efficiency. For editing of the *TNNT2* locus, the target region was first amplified by PCR using GoTaq Green Master Mix (Promega, M7122). Heteroduplexes were formed from resulting PCR amplicons and incubated with Surveyor nuclease, according to the instructions provided with the Surveyor mutation detection kit (IDT, 706020). Treated products were run on a 3% agarose gel to detect Surveyor-mediated cleavage of mismatches. For dual gRNA directed deletion at the *DMD* and *GAPDH* loci, the target region was amplified by PCR using GoTaq and PCR amplicons were directly visualized on a 3% agarose gel.

Genomic DNA extraction and genotyping analysis

Genomic DNA was isolated from treated cells using either QuickExtract DNA extraction solution (Lucigen, QE09050) according to the manufacturer's instruction or the Quick-DNA Miniprep Plus Kit (Zymo Research, D4068). Endpoint PCR was performed on extracted genomic DNA using GoTaq Green Master Mix (Promega, M7122) to amplify 5' and 3' genome-donor junctions. Genotyping primer sequences are listed in Table 1. PCR products were gel purified using the PureLink quick gel extraction and PCR purification combo kit (Invitrogen, K220001) and directly submitted for Sanger sequencing.

RT-PCR analysis

Total RNA was isolated from treated cells using the Direct-Zol RNA microprep/miniprep kits (Zymo Research, R2061/R2051) and reverse transcribed using the iScript cDNA synthesis kit (Biorad, 1708891) according to the manufacturer's instructions. Endpoint RT-PCR was performed on cDNA using GoTaq Green Master Mix to amplify the *DMD* target region. RT-PCR primer sequences are listed in Table 1. PCR products were gel purified using the PureLink kit and directly submitted for Sanger sequencing.

Flow cytometry

For the *TNNT2* targeting experiment, treated hESC-CMs were washed with PBS, trypsinized for 5-10 min to dissociate into single cells, and neutralized with RPMI/B27 media. Resuspended cells were transferred to a round bottom ultra-low attachment 96-well microplate (Corning, 7007), and the percentage of GFP positive cells was quantified using the Guava easyCyte flow cytometer (Millipore).

For the *GAPDH* targeting experiment, HEK293Ts were seeded at a density of 175,000 cells per well in a 24-well plate and transfected the following day using PEI. Culture media was changed the next morning, and cells were passaged at a ratio of 1:2 on day 2 post-transfection. On day 4, treated HEK293Ts were washed with PBS, trypsinized for 2-3 min, and neutralized with DMEM containing 10% FBS. Resuspended cells were transferred to Eppendorf tubes and analyzed by flow cytometry using the Attune NxT Acoustic Focusing Cytometer (Thermo Fisher).

Immunocytochemistry

Treated hESC-CMs were fixed with 4% paraformaldehyde for 15 min and blocked with PBS plus 0.4% Triton X-100 and 5% non-fat dry milk (Bio-Rad) for 1 hr at room temperature.

Fixed cells were then co-stained for cardiac Troponin T (Abcam, ab8295), GFP (Invitrogen, A-11122), and DAPI (Thermo Fisher, D1306).

5.5 Acknowledgements

The authors thank Christopher Nelson and Adrian Oliver for their generous counsel on characterization of on-target editing. Chengzu Long, Rhonda Bassel-Duby, and Efrain Sanchez-Ortiz kindly provided guidance on dystrophin Western blotting. We also thank Thomas Gaj for engaging in insightful discussions and the SCVI for supplying the DMD ΔE51 iPSC line.

5.6 Funding

This work was supported by the National Institutes of Health grant (NIH R01 EY022975). SS was supported by the American Heart Association predoctoral fellowship and NIH R21 EB021572-01.

Conflict of interest statement. DVS is an inventor on patents involving AAV directed evolution and a co-founder of a company developing AAV vectors for clinical gene therapy.

5.7 References

1. Mendell, J.R., et al., *Evidence-based path to newborn screening for Duchenne muscular dystrophy.* Ann Neurol, 2012. **71**(3): p. 304-13.
2. Muntoni, F., S. Torelli, and A. Ferlini, *Dystrophin and mutations: one gene, several proteins, multiple phenotypes.* Lancet Neurol, 2003. **2**(12): p. 731-40.
3. Duan, D., *Systemic AAV Micro-dystrophin Gene Therapy for Duchenne Muscular Dystrophy.* Mol Ther, 2018. **26**(10): p. 2337-2356.
4. Lim, K.R.Q., C. Yoon, and T. Yokota, *Applications of CRISPR/Cas9 for the Treatment of Duchenne Muscular Dystrophy.* J Pers Med, 2018. **8**(4).
5. Bengtsson, N.E., et al., *Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy.* Nat Commun, 2017. **8**: p. 14454.
6. Ishizu, T., et al., *Targeted Genome Replacement via Homology-directed Repair in Non-dividing Cardiomyocytes.* Sci Rep, 2017. **7**(1): p. 9363.
7. Pawelczak, K.S., et al., *Modulating DNA Repair Pathways to Improve Precision Genome Engineering.* ACS Chem Biol, 2018. **13**(2): p. 389-396.

8. Zetsche, B., et al., *Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system.* Cell, 2015. **163**(3): p. 759-71.

9. Bin Moon, S., et al., *Highly efficient genome editing by CRISPR-Cpf1 using CRISPR RNA with a uridinylate-rich 3'-overhang.* Nat Commun, 2018. **9**(1): p. 3651.

10. Nelson, C.E., et al., *In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy.* Science, 2016. **351**(6271): p. 403-7.

11. Long, C., et al., *Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy.* Science, 2016. **351**(6271): p. 400-3.

12. Tabebordbar, M., et al., *In vivo gene editing in dystrophic mouse muscle and muscle stem cells.* Science, 2016. **351**(6271): p. 407-411.

13. Zhang, Y., et al., *CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice.* Sci Adv, 2017. **3**(4): p. e1602814.

14. Amoasii, L., et al., *Single-cut genome editing restores dystrophin expression in a new mouse model of muscular dystrophy.* Sci Transl Med, 2017. **9**(418).

15. Young, C.S., et al., *A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle Cells.* Cell Stem Cell, 2016. **18**(4): p. 533-40.

16. Amoasii, L., et al., *Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy.* Science, 2018. **362**(6410): p. 86-91.

17. Long, C., et al., *Correction of diverse muscular dystrophy mutations in human engineered heart muscle by single-site genome editing.* Sci Adv, 2018. **4**(1): p. eaap9004.

18. Min, Y.L., et al., *CRISPR-Cas9 corrects Duchenne muscular dystrophy exon 44 deletion mutations in mice and human cells.* Sci Adv, 2019. **5**(3): p. eaav4324.

19. Nelson, C.E., et al., *Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy.* Nat Med, 2019. **25**(3): p. 427-432.

20. Nicolas, A., et al., *Becker muscular dystrophy severity is linked to the structure of dystrophin.* Hum Mol Genet, 2015. **24**(5): p. 1267-79.

21. Ho, R., M.L. Nguyen, and P. Mather, *Cardiomyopathy in becker muscular dystrophy: Overview.* World J Cardiol, 2016. **8**(6): p. 356-61.

22. Suzuki, K., et al., *In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration.* Nature, 2016. **540**(7631): p. 144-149.

23. He, X., et al., *Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair.* Nucleic Acids Res, 2016. **44**(9): p. e85.

24. Zhang, C., et al., *Homology-independent multiallelic disruption via CRISPR/Cas9-based knock-in yields distinct functional outcomes in human cells.* BMC Biol, 2018. **16**(1): p. 151.

25. Rapti, K., et al., *Effectiveness of gene delivery systems for pluripotent and differentiated cells.* Mol Ther Methods Clin Dev, 2015. **2**: p. 14067.

26. Guan, X., et al., *Use of Adeno-Associated Virus to Enrich Cardiomyocytes Derived from Human Stem Cells.* Hum Gene Ther Clin Dev, 2015. **26**(3): p. 194-201.

27. Flotte, T.R., et al., *Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter.* J Biol Chem, 1993. **268**(5): p. 3781-90.

28. Wang, L., et al., *Productive life cycle of adeno-associated virus serotype 2 in the complete absence of a conventional polyadenylation signal.* J Gen Virol, 2015. **96**(9): p. 2780-7.

29. Aartsma-Rus, A., I.B. Ginjaar, and K. Bushby, *The importance of genetic diagnosis for Duchenne muscular dystrophy.* J Med Genet, 2016. **53**(3): p. 145-51.

30. Branzei, D. and M. Foiani, *Regulation of DNA repair throughout the cell cycle.* Nat Rev Mol Cell Biol, 2008. **9**(4): p. 297-308.

31. Cristea, S., et al., *In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration.* Biotechnol Bioeng, 2013. **110**(3): p. 871-80.

32. Maresca, M., et al., *Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining.* Genome Res, 2013. **23**(3): p. 539-46.

33. Lian, X., et al., *Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling.* Proc Natl Acad Sci U S A, 2012. **109**(27): p. E1848-57.

34. Lian, X., et al., *Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/beta-catenin signaling under fully defined conditions.* Nat Protoc, 2013. **8**(1): p. 162-75.

35. Nakade, S., et al., *Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9.* Nat Commun, 2014. **5**: p. 5560.

36. Yao, X., et al., *Homology-mediated end joining-based targeted integration using CRISPR/Cas9.* Cell Res, 2017. **27**(6): p. 801-814.

37. Zhang, J.P., et al., *Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage.* Genome Biol, 2017. **18**(1): p. 35.

38. Guo, T., et al., *Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing.* Genome Biol, 2018. **19**(1): p. 170.

39. Flanigan, K.M., et al., *Mutational spectrum of DMD mutations in dystrophinopathy patients: application of modern diagnostic techniques to a large cohort.* Hum Mutat, 2009. **30**(12): p. 1657-66.

40. Gaudelli, N.M., et al., *Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage.* Nature, 2017. **551**(7681): p. 464-471.

41. Komor, A.C., et al., *Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage.* Nature, 2016. **533**(7603): p. 420-4.

42. Giannoukos, G., et al., *UDiTaS, a genome editing detection method for indels and genome rearrangements.* BMC Genomics, 2018. **19**(1): p. 212.

43. Bruinsma, S., et al., *Bead-linked transposomes enable a normalization-free workflow for NGS library preparation.* BMC Genomics, 2018. **19**(1): p. 722.

44. Miller, D.G., L.M. Petek, and D.W. Russell, *Adeno-associated virus vectors integrate at chromosome breakage sites.* Nat Genet, 2004. **36**(7): p. 767-73.

45. Fu, H., et al., *Self-complementary adeno-associated virus serotype 2 vector: global distribution and broad dispersion of AAV-mediated transgene expression in mouse brain.* Mol Ther, 2003. **8**(6): p. 911-7.

46. Grieger, J.C., V.W. Choi, and R.J. Samulski, *Production and characterization of adeno-associated viral vectors.* Nat Protoc, 2006. **1**(3): p. 1412-28.

47. Gaj, T. and D.V. Schaffer, *Adeno-Associated Virus-Mediated Delivery of CRISPR-Cas Systems for Genome Engineering in Mammalian Cells.* Cold Spring Harb Protoc, 2016. **2016**(11).

48. Ayuso, E., et al., *High AAV vector purity results in serotype- and tissue-independent enhancement of transduction efficiency.* Gene Ther, 2010. **17**(4): p. 503-10.

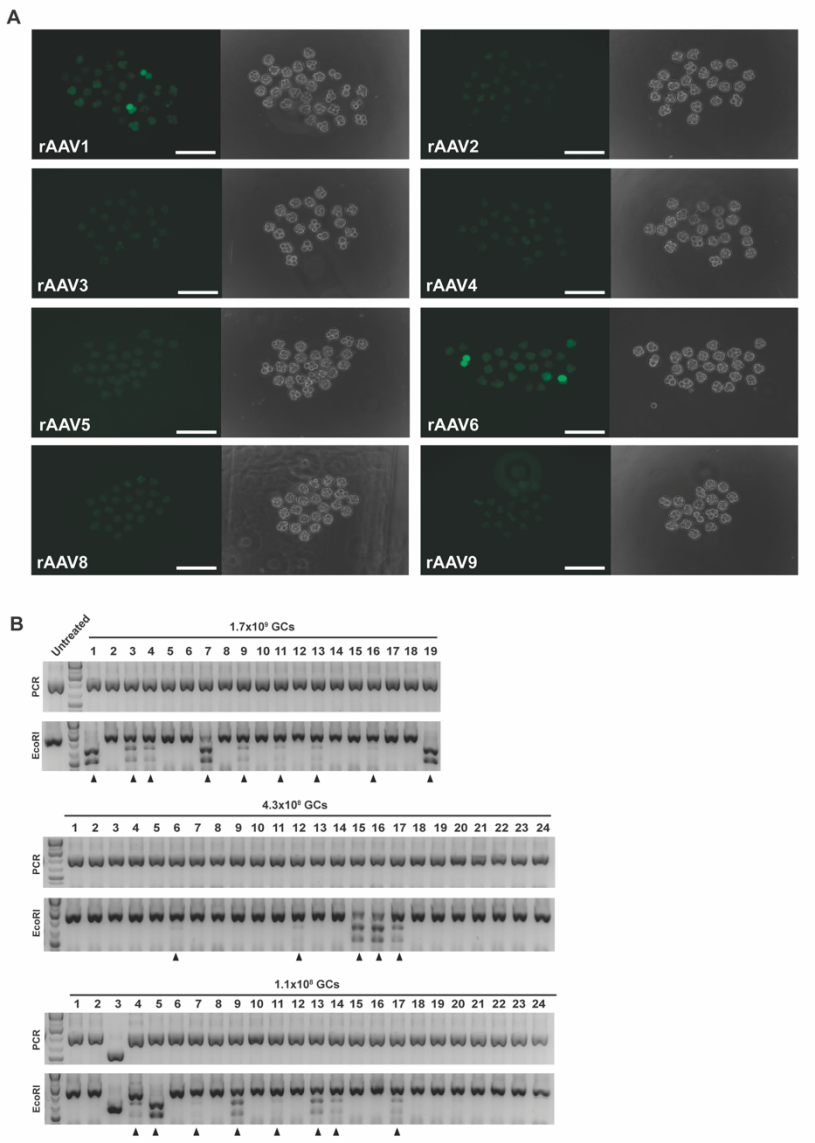
Appendix A: Supplementary Material for Chapter 4

This appendix is adapted from a manuscript published as

S. Sun*, S. Chen*, D. Moonen, C. Lee, A.Y. Lee, D.V. Schaffer†, L. He†. CRISPR-READI: Efficient generation of knock-in mice by CRISPR RNP electroporation and AAV donor infection. *Cell Reports*, 2019. **27**(13): p.3780-3789.

*Indicates co-first authors
†Indicates co-corresponding authors

A.1 Supplementary Figures



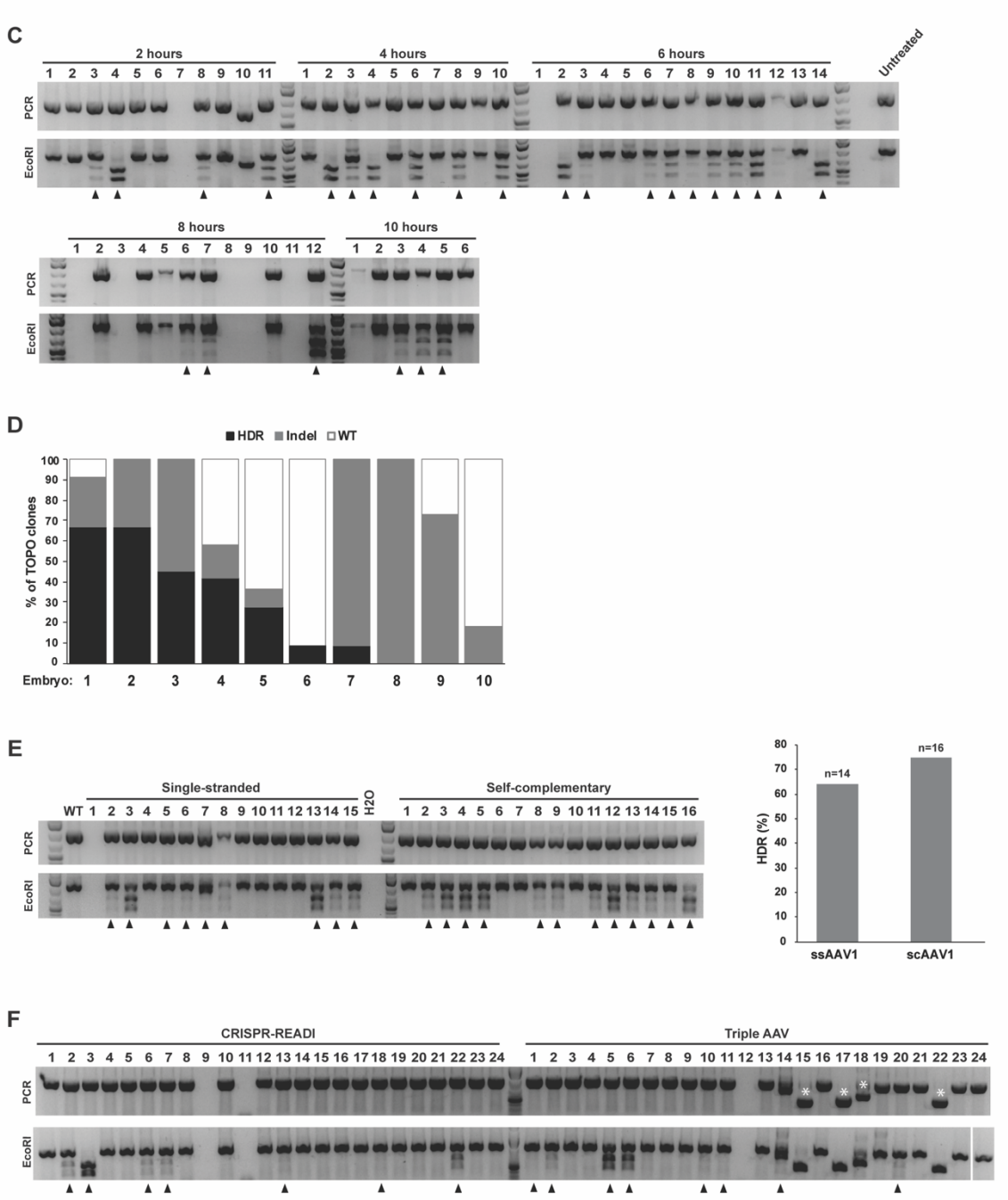


Figure A.1. Optimization of CRISPR-READI for HDR-mediated editing in mouse embryos. Related to Figure 4.1.

A) Mouse zygote transduction efficiency of eight AAV serotypes. Zygotes were transduced with rAAV vectors packaged with a CMV-eGFP reporter and imaged by fluorescent microscopy 48 hours post-transduction. Representative fluorescent and brightfield images of 4-cell stage embryos are shown for each tested serotype. Scale bars = 30 μm.

B) RFLP analysis for rAAV1 dosage optimization. Zygotes were transduced with scAAV1-Tyr at a dose of 1.1×108, 4.2×108, or 1.7×109 GCs, electroporated with RNPs 5 hours post-transduction, and then returned to rAAV1 incubation for another 19 hours. Treated embryos were cultured to the morula stage and then lysed. PCR products amplified from the Tyr locus were digested with EcoRI. Edited embryos yield 650 bp and 420 bp bands upon EcoRI digestion of the PCR amplicon (black arrows).

C) RFLP analysis for optimization of RNP electroporation timing relative to rAAV transduction. Zygotes were transduced with scAAV1-Tyr at a dose of 1.7×109 GCs and electroporated at various time points post-transduction (2, 4, 6, 8, or 10 hours), and returned to rAAV incubation for a total of 24 hours. Treated embryos were cultured to the morula stage, lysed, and assessed by RFLP analysis.

D) Analysis of indel and HDR frequency in individual treated embryos. Ten embryos treated with CRISPR-READI using scAAV1-Tyr (1.7x109 GCs) were cultured to morula. After lysis, the targeted region of the Tyr locus was amplified by PCR and cloned into a TOPO-TA subcloning vector. Ten clones from each embryo were sequenced to characterize mosaicism of editing at the target site.

E) Comparison of HDR editing efficiency using self-complementary and single-stranded rAAV1 Tyr donor vectors. Treated embryos were cultured to the morula stage, lysed, and assessed by RFLP analysis (left). HDR rate was quantified for each condition (right).

F) Comparison of HDR editing efficiency using CRISPR-READI and a triple AAV strategy. Zygotes were either treated with CRISPR-READI using the scAAV1-Tyr donor vector or transduced with ssAAV1-SpCas9, scAAV1-Tyr-gRNA, and scAAV1-Tyr simultaneously. All AAVs were delivered at 1×108 GCs. After 24 hours, all embryos were removed from AAV incubation and transferred to fresh media. Treated embryos were cultured to the morula stage, lysed, and assessed by RFLP analysis. Large deletions (*) were observed with the triple AAV strategy.

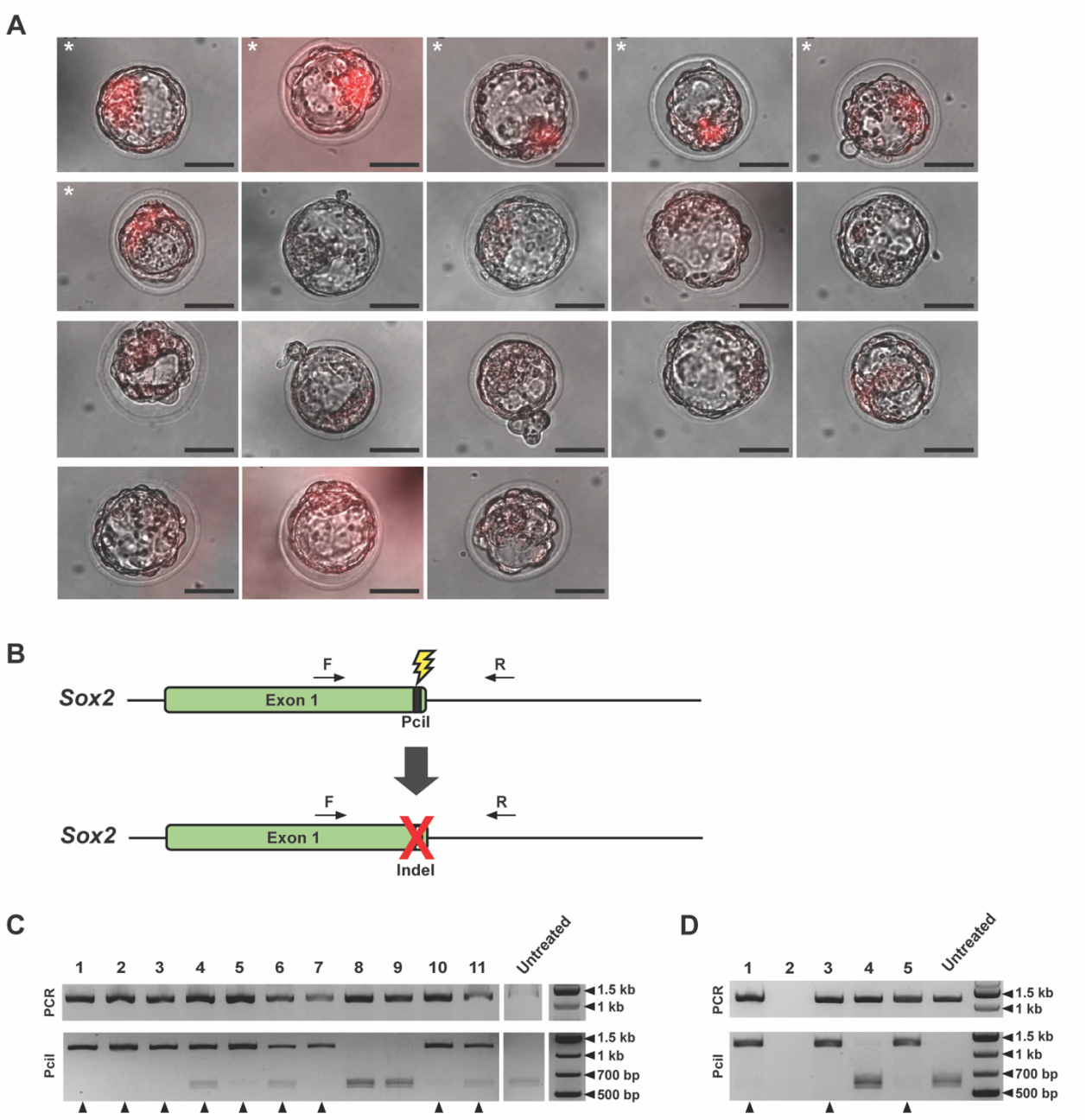


Figure A.2. Characterization of editing at the Sox2 locus in embryos and mice treated with CRISPR-READI. Related to Figure 4.2 and Figure 4.3.

A) Merged brightfield and fluorescent images of blastocysts treated with CRISPR-READI to express a fluorescent reporter from the endogenous Sox2 locus. Embryos were treated with scAAV1-Sox2-mStr at a dose of 1.3×108 GCs, cultured to the late blastocyst stage, and imaged by fluorescent microscopy. Successfully edited blastocysts (*) show localization of mStrawberry fluorescence to the inner cell mass, recapitulating the endogenous Sox2 expression pattern.

B) Schematic of RLFP strategy used to identify indels on the Sox2 target locus. An endogenous PciI restriction site lies directly upstream of the PAM; indel formation at the target site destroys the PciI recognition sequence, retaining the 1.1 kb band of the PCR amplicon. F/R: Forward and Reverse primers used for RFLP analysis.

C-D) RFLP analysis of live pups generated with CRISPR-READI using scAAV1-Sox2-mStr (C) and ssAAV1-Sox2-CreERT2 (D). The targeted region of the Sox2 locus was amplified by PCR and digested with PciI to quantify indel frequency. Correctly edited alleles were excluded from this analysis by limiting extension time of PCR amplification. Unedited alleles yield 579 bp and 506 bp bands upon PciI digestion of the PCR amplicon, while indel-containing alleles result in retention of the 1.1 kb PCR product (black arrows).

Appendix B: Supplementary Material for Chapter 5

B.1 Supplementary Figures

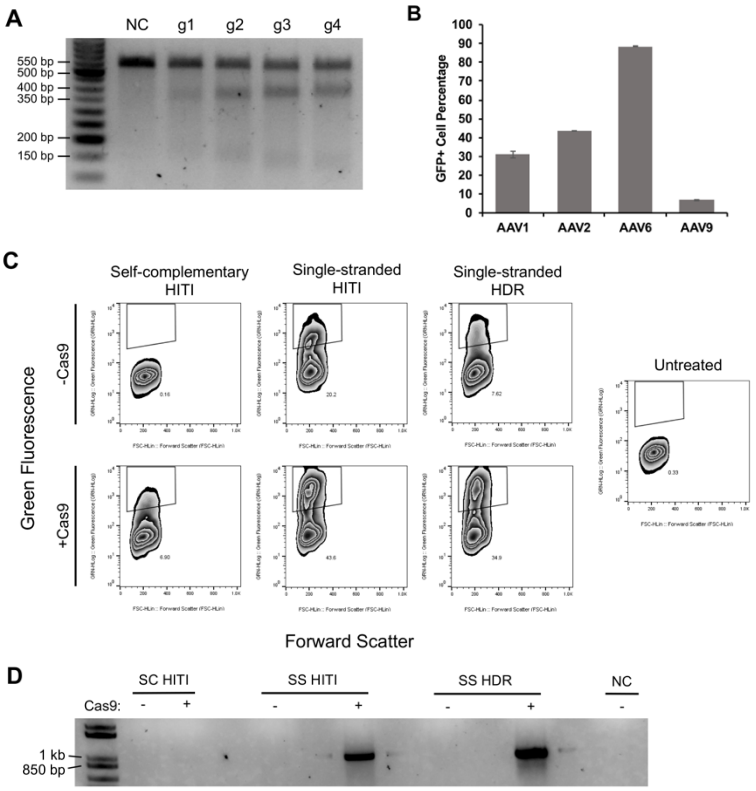


Figure B.1. Optimization of gene targeting efficiency in hESC-CMs by AAV-HITI. Related to Figure 5.1.

A) Screen of candidate cTnT-gRNAs for on-target cleavage activity by Surveyor nuclease assay. Mismatches were detected in the *TNNT2* locus of HEK293Ts following transfection of plasmids encoding SpCas9 and cTnT-gRNAs. Treatment with Surveyor nuclease yields expected fragment sizes of 362-388 bp and 136-162 for all cTnT-gRNAs. Unmodified alleles produce a 524 bp band. NC: untreated control.

B) AAV infectivity of hESC-CMs. Cardiomyocytes were transduced with self-complementary AAV vectors encoding a CAG-eGFP reporter cassette at an MOI of 10,000. GFP-positive cells were quantified by flow cytometry 6 days post-infection.

C) Density plots from flow cytometry analysis on hESC-CMs treated with AAV-cTnT-HITI and AAV-cTnT-HDR vectors. Addition of the AAV-SpCas9 vector (bottom row) increases the percentage of GFP-positive cells relative to their respective donor only controls (top row).

D) PCR genotyping of the 5' genome-donor junction in hESC-CMs treated with AAV-cTnT-HITI and AAV-cTnT-HDR vectors. Only cells infected with AAV-SpCas9 and either the single-stranded HITI or HDR donor vector resulted in site-specific transgene integration. Corresponding donor only controls were negative for targeted knock-in, despite the detection of GFP expression in these conditions by flow cytometry.

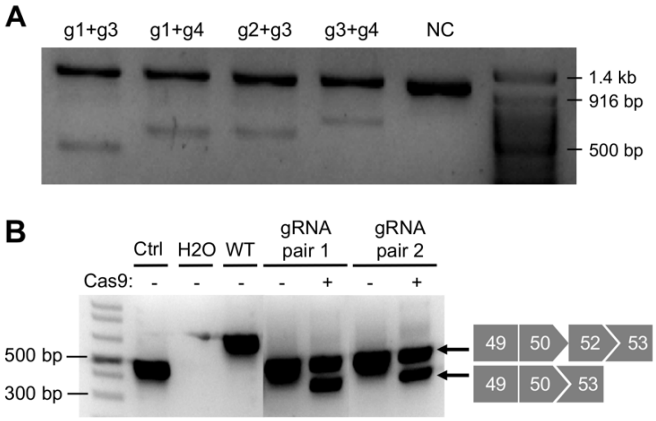


Figure B.2. Characterization of on-target editing at the human *DMD* locus by AAV dual-cut HITI. Related to Figure 5.2.

A) Screen of candidate *DMD* guide RNA pairs for on-target genomic deletion. Plasmids encoding SpCas9 and guide RNAs with target sites upstream or downstream of human *DMD* exon 52 were paired and co-transfected into HEK293Ts. Induced genomic deletion at the target region was detected by PCR genotyping for all guide RNA pairs. Dual gRNA directed deletion of *DMD* exon 52 yields an expected fragment size of 378-574 bp. Unmodified alleles produce a 993 bp band. NC: untreated control.

B) Deletion of target exon 52 in treated ΔE51 DMD-CMs. *DMD* transcripts without exons 51 and 52 were detected by reverse transcription PCR using primers specific to *DMD* exons 49 and 53 in cells treated with AAV dual-cut HITI. gRNA pair 1: DMD-g1 + DMD-g3. gRNA pair 2: DMD-g1 + DMD-g4. Ctrl: untreated control. H2O: PCR control with no cDNA template. WT: wildtype hESC-CM control. *DMD* ΔE51 expected band size: 469 bp. *DMD* ΔE51 expected band size: 351 bp.

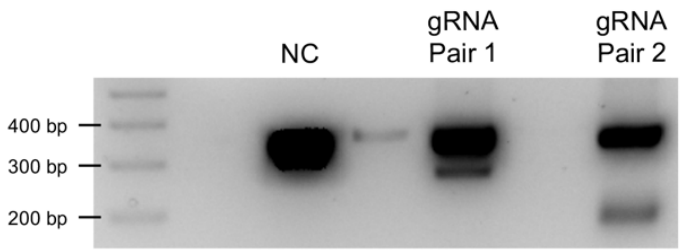


Figure B.3. Screening for dual gRNA directed deletion at the human *GAPDH* locus. Related to Figure 5.3.

Plasmids encoding SpCas9 and guide RNAs targeting the 3'UTR of the *GAPDH* locus were paired and co-transfected into HEK293Ts. Induced genomic deletion at the target region was detected by PCR genotyping for both guide RNA pairs. Dual gRNA directed deletion of the *GAPDH* 3'UTR yields an expected fragment size of 275 bp for gRNA pair 1 and 182 bp for gRNA pair 2. Unmodified alleles produce a 351 bp band. gRNA pair 1: GAPDH-g1 + GAPDH-g2. gRNA pair 2: GAPDH-g3 + GAPDH-g4. NC: untreated control.

Appendix C: Display of Engineered High-Affinity Affibodies on the AAV Capsid for Vector Re-Targeting

C.1 Background

Cardiac gene therapy using adeno-associated viral (AAV) vectors with targeted delivery to diseased cardiomyocytes may enable correction of molecular signaling abnormalities in these cells. AAV's minimal risk of insertional mutagenesis, low immunogenicity, and ability to transduce and persist in non-dividing cells contribute to its strong therapeutic potential for single-administration treatments for cardiac diseases. However, its broad tropism, or affinity for other host tissues besides the heart, necessitates high vector doses to overcome vector sequestration in peripheral tissues, elevates the risk of immune presentation and reaction, and raises the possibility of side effects due to off-target transgene expression.

Re-targeting AAV vectors has been explored through the conjugation of receptor-specific affibodies and designed ankyrin repeat proteins (DARPins) to viral protein 2 (VP2) of the AAV capsid[1, 2]. These ligands were identified through phage display and exhibit high binding affinity for receptors that are highly and uniquely expressed in rare cell types[3], such as tumor cells and CD4+ lymphocytes. Grafting these small proteins onto the AAV capsid effectively re-targeted the vector to cell types overexpressing the given receptor. This precedent motivated my hypothesis that directed evolution can be harnessed to engineer AAV variants with high-affinity ligands that target vectors to a unique surface receptor on cardiomyocytes.

C.2 Results

Functional display of known high affinity affibodies on the AAV capsid

The N-terminus of VP2 has recently been used for incorporation of small proteins that classical insertion sites in VP3 fail to tolerate[4]. Fusion of large ligands, including GFP and luciferase, to VP2 has shown that insertion at this position does not interfere with capsid assembly[5]. To validate these previous findings, I built and optimized an expression system that enables assembly of AAV vectors with ligands conjugated to the N-terminus of VP2 (Figure 1A).

As a first proof-of-concept experiment, I fused the known immunoglobulin (IgG) binding C domain from protein A of *Staphylococcus aureus* to the AAV capsid. This small protein, known as an affibody, rapidly folds into a three-helical structure, exposing a large 800-Å surface that can achieve up to picomolar affinity for its target binding partner. To verify proper expression and incorporation of the C domain-fused VP2 into the AAV capsid, I performed Western blots on purified vector that confirmed the presence of the three VPs at the expected stoichiometric ratio (Figure 1B), demonstrating that fusion of this small protein does not disrupt viral particle assembly. However, tagging different types of single-chain antibodies to the same site produced variable levels of VP2 (Figure 1B), suggesting that incorporation is ligand-dependent.

After validating that an affibody-fused VP2 can be properly assembled with wildtype AAV VP1 and VP3 into viral particles, I tested functional display of the ligand on the AAV capsid. To verify that the C domain retains its strong affinity for IgG while presented on the capsid surface, I designed a sandwich enzyme-linked immunosorbent assay (ELISA) that captures C domain-fused AAV viral particles (Cdom-AAV) on a surface coated with rabbit IgG and detects intact capsids with an A20 antibody, which binds to a conformational epitope on assembled capsids that is absent in denatured and native capsid proteins. By specifically probing for intact AAV capsids with C domain affibodies that can bind to rabbit IgG, the assay showed that rabbit IgG engages with Cdom-AAVs while its interaction with AAV vectors lacking the C domain-fused VP2 (VP2KO) is undetectable (Figure 1C).

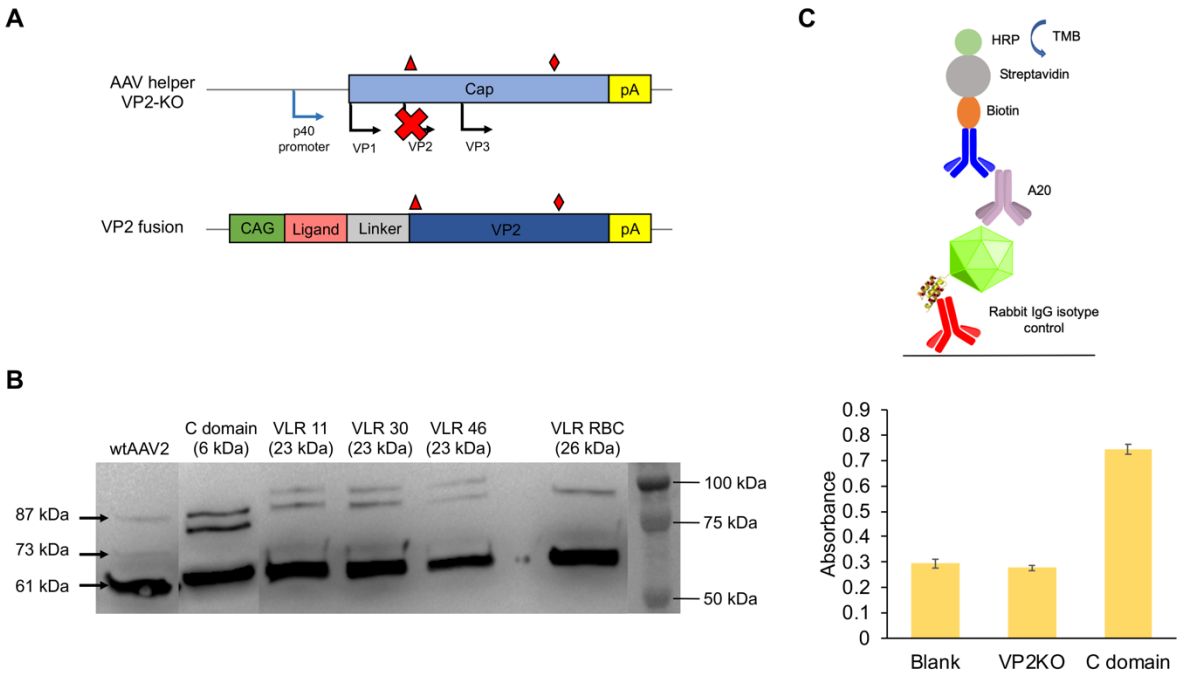


Figure C.1. Validation of AAV VP2 fusion system for display of ligands on the AAV capsid.

A) Plasmid constructs for AAV VP2 fusion system. In the AAV helper plasmid, the VP2 start codon of the AAV *cap* gene (red diamond) and residues conferring recognition of cell surface glycans (red diamond) were mutagenized to knock out VP2 expression and ablate native binding, respectively. These mutations were applied to base serotypes, AAV2 and AAV9, such that resulting capsids cannot engage with heparan sulfate and galactose for cell entry. The VP2 fusion construct was assembled to encode a CAG promoter driving expression of a ligand fused to the N-terminus of the mutant VP2 fragment via a glycine-alanine linker. p40 promoter: native promoter driving *cap* expression. VP1-3: AAV viral proteins 1-3. pA: polyadenylation signal sequence.

B) Western blot analysis of AAV capsids displaying various ligands on VP2. The IgG-binding C domain derived from protein A of *S. aureus* and different engineered variants of variable lymphocyte receptors (VLRs) derived from lampreys were conjugated to AAV2 VP2. VP2 fusion constructs were co-expressed with AAV2 VP1 and VP3 to produce viral particles, which were then denatured for characterization by Western blot. Modified VP2 proteins were detected for all ligands, except for VLR RBC, and showed a band shift, relative to the wildtype VP2 control, that corresponded to the size of the attached ligand.

C) Activity of Cdom-AAV on its native substrate. A custom sandwich ELISA assay was designed using a rabbit IgG isotype control to capture AAV capsids displaying functional C domain ligands. Bound, intact capsids were then probed for using the A20 antibody. Signal was amplified with a biotinylated secondary antibody and HRP-conjugated streptavidin, generated with TMB substrate, and measured using a spectrophotometer. Binding activity was detected for Cdom-AAV, compared to the blank and VP2KO controls.

Engineering affibody-fused AAV variants with target specificity for the mCrhr2 receptor

To develop an affibody variant that re-targets AAV vectors to cardiomyocytes, I am applying directed evolution to affibody libraries displayed on AAV capsids and selecting for mutants that utilize the mouse corticotropin-releasing factor receptor (mCrhr2) to enter cells. Crhr2 is a G protein-coupled receptor implicated in mediating cardiovascular homeostasis and is highly and selectively expressed in mouse and human cardiomyocytes[6]. Crhr2 expression is also upregulated in failing mouse cardiomyocytes, making it a potential therapeutically relevant target.

First, two affibody libraries were generated through mutagenesis of 13 surface-exposed residues in helices 1 and 2, which was produced in-house, or helices 2 and 3, which was purchased from Creative Biolabs (Figure 2A). The affibody variants were then cloned into AAV2 and AAV9 VP2 plasmid backbones for N-terminal presentation and packaged into AAV vectors. Interestingly, Sanger sequencing analysis revealed that vector production enriches for affibody variants encoding premature stop codons (Figure 2B). Early termination of the affibody leads to assembly of viral particles that only contain VP1 and VP3, as determined by Western blot analysis (Figure 2C). These data suggest that VP2 knockout (VP2KO) mutants undergo efficient viral particle assembly; in contrast, incorporation of an affibody likely impedes packaging. Nonetheless, through site-directed mutagenesis of residues in AAV2 and AAV9 that confer native cell surface glycan binding to enable subsequent cellular entry, we show that the infectivity of VP2KO mutants is ablated. Therefore, the presence of these undesired mutants is not expected to confound downstream selections.

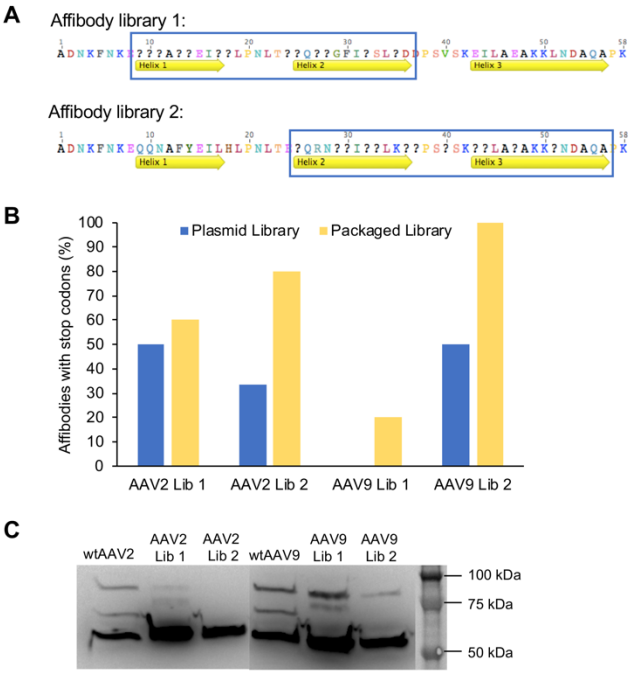


Figure C.2. Construction and characterization of affibody libraries conjugated to AAV VP2.

A) Amino acid sequences of affibody libraries based on the C domain. Helices 1 and 2 were mutagenized for affibody library 1, and 2 and 3 for affibody library 2. ?: degenerate codon.

B) Quantification of affibody variants encoding stop codons before and after AAV vector production. 10 clones each from plasmid and packaged AAV libraries were analyzed by Sanger sequencing. Affibody mutants encoding early termination are enriched during viral particle assembly across all libraries.

C) Western blot analysis of AAV vectors displaying affibody libraries. Capsids comprised of only AAV VP1 and VP3 are detected upon assembly of affibody libraries into viral particles. Presence of VP2KO mutants appears to be correlated with the abundance of variants encoding stop codons within the affibody library.

To select for mCrhr2-mediated entry, a HEK293T cell line overexpressing mCrhr2 and Cre recombinase (mCrhr2-Cre OE) was produced using a lentiviral donor system and AAV vectors displaying the affibody libraries were added to these cells. A Cre-dependent approach was applied to selectively recover variants that not only bind and enter the cells, but also traffic to the nucleus and complete all steps necessary for gene expression. After stripping the cell surface of non-specific binders with a low pH buffer, genomic DNA of low molecular weight was extracted from treated cells, and Cre-recombined AAV genomes were amplified by PCR. Affibody variants containing stop codons were detected, though at lower frequencies than in the pre-selection libraries, by sequencing analysis. We hypothesize that these false positives may arise from VP2KO mutants that bind non-specifically to the cell surface or cross-packaging of non-functional genomes by functional capsid variants. Then, to identify clones that specifically engage with mCrhr2 rather than another surface receptor for cellular entry, libraries recovered from the positive selection round were incubated with wildtype HEK293T cells for 1 hour and variants that cannot bind and enter cells without mCrhr2 were collected from the culture supernatant. These experiments are in progress, and Sanger sequencing will be used to evaluate the relative abundances of individual clones that have emerged from positive and negative selection for mCrhr2 target specificity.

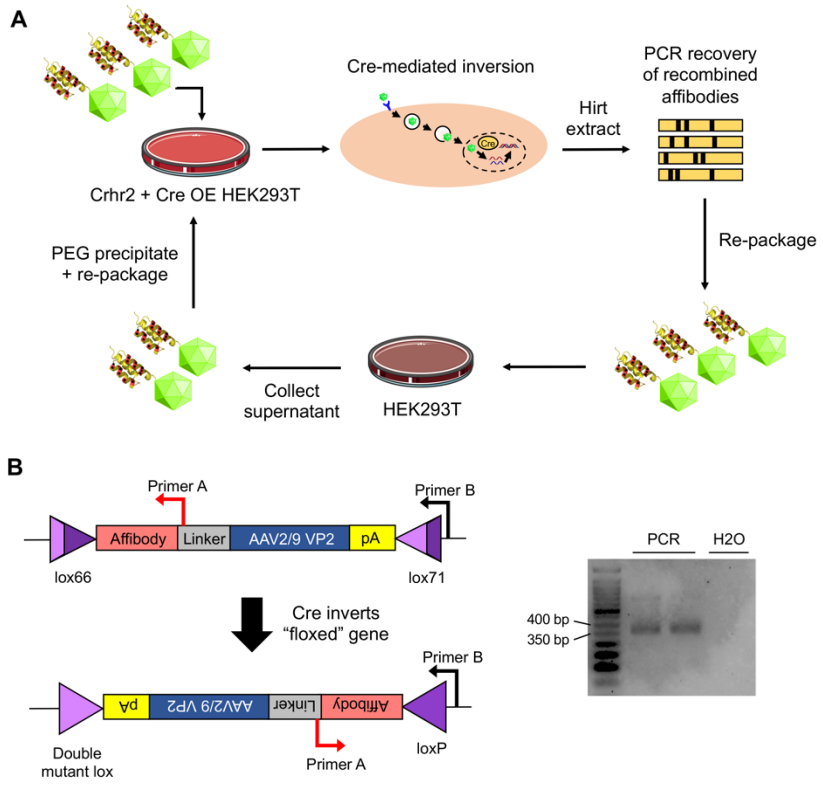


Figure C.3. Selection of affibody-fused AAV libraries for mCrhr2 target specificity.

A) Schematic of selection strategy. HEK293T cells stably overexpressing the target mCrhr2 receptor and Cre recombinase are transduced with affibody-fused AAV libraries. Cellular internalization, intracellular trafficking, nuclear import, and capsid uncoating are necessary for Cre-mediated inversion of AAV genomes. DNA of low molecular weight is extracted from treated cells, and recombined affibodies are recovered by PCR using primers specific to inversion. The recovered affibody pool is then re-packaged into viral particles using the AAV VP2 fusion system and applied to wildtype HEK293Ts. Culture supernatant is collected to select against clones that utilize non-mCrhr2 receptors for cellular entry and OE: overexpression. PEG: polyethylene glycol.

B) Recovery of affibody clones following selection on mCrhr2 OE HEK293Ts. Cre recombined affibody variants were amplified by PCR using primers specific to inversion. Expected band size: 370 bp.

C.3 Future Directions

Characterization of evolved affibody-fused AAV variants for mCrhr2-expressing cells *in vitro* and *in vivo*

To determine *in vitro* specificity for mCrhr2-expressing cells, candidate vectors will be packaged with a GFP reporter and flow cytometry will be used to quantify transduction efficiency of Crhr2-Cre OE and wildtype HEK293T cells. To assess whether Crhr2-mediated entry enables cardiomyocyte targeting, the same GFP vectors will be intravenously administered to adult C57Bl/6 mice, and immunohistochemistry and confocal microscopy will be used to measure vector localization to cardiomyocytes relative to other cell types of the heart. Finally, to evaluate *in vivo*

biodistribution, the variants will be packaged with a firefly luciferase reporter and again systemically administered into adult C57Bl/6 mice. Luciferase expression from major tissues will be quantified using a luminometer. Data from these experiments will be collectively analyzed to isolate the top variant that transduces cardiomyocytes with the greatest efficiency and specificity.

C.4 References

1. Munch, R.C., et al., *Displaying high-affinity ligands on adeno-associated viral vectors enables tumor cell-specific and safe gene transfer.* Mol Ther, 2013. **21**(1): p. 109-18.
2. Hagen, S., et al., *Modular adeno-associated virus (rAAV) vectors used for cellular virus-directed enzyme prodrug therapy.* Sci Rep, 2014. **4**: p. 3759.
3. Munch, R.C., et al., *Off-target-free gene delivery by affinity-purified receptor-targeted viral vectors.* Nat Commun, 2015. **6**: p. 6246.
4. Warrington, K.H., Jr., et al., *Adeno-associated virus type 2 VP2 capsid protein is nonessential and can tolerate large peptide insertions at its N terminus.* J Virol, 2004. **78**(12): p. 6595-609.
5. Lux, K., et al., *Green fluorescent protein-tagged adeno-associated virus particles allow the study of cytosolic and nuclear trafficking.* J Virol, 2005. **79**(18): p. 11776-87.
6. Tsuda, T., et al., *Corticotropin releasing hormone receptor 2 exacerbates chronic cardiac dysfunction.* J Exp Med, 2017. **214**(7): p. 1877-1888.