Title
Expanding the Potential of Adeno-Associated Virus for the Treatment of Intractable Inherited Retinal Degenerations

Author
Day, Timothy Paul

Publication Date
2017

Peer reviewed|Thesis/dissertation

Expanding the Potential of Adeno-Associated Virus for the Treatment of Intractable Inherited Retinal Degenerations

By

Timothy Paul Day

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John G. Flannery, Co-Chair
Professor David V. Schaffer, Co-Chair
Professor Marla Feller
Professor Ehud Isacoff
Professor Laurent Coscoy

Fall 2017

Abstract

Expanding the Potential of Adeno-Associated Virus for the Treatment of

Intractable Inherited Retinal Degenerations

by

Timothy Paul Day

Doctor of Philosophy in Neuroscience

University of California, Berkeley

Professor John G. Flannery, Co-Chair

Professor David V. Schaffer, Co-Chair

Inherited retinal degenerations, effecting about 1:3000 people, have historically been challenging to treat. A large number of mutations have been identified in these patients, which share a similar phenotype of a gradual loss of photoreceptor cells leading to blindness. In recent decades, a promising new form of treatment called gene therapy has been developed to address patient needs. Encouragingly, there have been several successful clinical trials using adeno-associated virus (AAV) to deliver genetic cargo to the necessary retinal cells. Viral vectors based on AAV have seen success because of the ability of the virus to achieve high transfection efficiency and a strong safety profile. The seminal gene therapy clinical trial utilizing AAV was the delivery of *RPE65* to the retinas of patients with Leber congenital amaurosis (LCA), which proved to have long-term safety and showed significant efficacy. But, challenges still remain for broad application of AAV. It is therefore necessary to develop AAV variants with novel properties in order to address the wide range of retinal degenerations.

There are a number of natural AAV serotypes that exhibit different tropisms for different retinal cells allowing for the potential of treatments targeted to different cell types. Importantly, the route of delivery strongly affects the cells targeted and depends on whether the virus is administered subretinally or intravitreally. Natural AAV serotypes perform poorly when delivered intravitreally due to structural barriers, such as the inner limiting membrane. Therefore, initial clinical trials chose to deliver AAV subretinally, which has the advantage of local delivery to the photoreceptors or retinal pigment epithelium and limited immune response. But, expression is restricted to only a portion of the retina and there is a risk of damage from retinal detachment in a tissue that is already degenerating. Intravitreal injection would be preferred and would allow for pan-retinal expression and would not cause retinal detachment. Importantly, recent evidence has shown species specificity for AAV variants necessitating future AAV engineering approaches to take into account the specific challenges of the human retina. One method of AAV engineering, directed evolution, has yielded many novel AAV variants with advantageous properties. Directed evolution is a biomimetic process that emulates how viruses

naturally evolve. Large libraries of highly diverse AAV variants are iteratively selected utilizing pressures for a desired attribute.

This dissertation focuses on the directed evolution of novel variants that are able to intravitreally deliver cargo in large animal models. Canines serve as a strong preclinical model for retinal degenerations and contain a visual streak for high acuity vision, which is lacking in rodents. Directed evolution was performed in a canine model and variants were selected utilizing deep sequencing of each round of evolution resulting in AAVs that had high outer nuclear layer or bipolar cell tropism when delivered intravitreally. Directed evolution was also performed in non-human primates (NHPs) because the retina shares many similarities to the human retina, including a cone-rich macula and fovea, and has the highest chance of yielding the most clinically relevant AAVs.

To generate a NHP variant, several new approaches and tools were developed. First, novel AAV libraries were generated including AAV4-, AAV5-, and Ancestral-based peptide insertion libraries. Multiple plasmid backbones with unique restriction and primer binding sites were created to allow for tuning of different pools of libraries throughout the directed evolution process. A packaging strategy that mitigated the possibility of replication and immune response during NHP directed evolution was developed leading to a 10-fold increase in safety. Serum screening for animals without pre-existing antibodies to AAV also increased safety. Implementing these tools and strategies for directed evolution in a NHP model led to an AAV2-based peptide insertion variant that showed increase cone targeting and decreased ganglion cell targeting when intravitreally delivered. For both the canine and NHP directed evolution studies, deep sequencing of each round of selection revealed “hidden” variants and elucidated the biases in the selection process due to initial overrepresentation of particular variants. An additional study to increase the packaging capacity of AAV was performed and demonstrated the size limits of AAV cargo. The research performed in this dissertation led to AAV variants with increased therapeutic properties and insights into the mutability of AAV and performance of individual variants throughout the directed evolution process.

Table of Contents

Acknowledgements

First, I would like to thank my parents Ken and Linda Day for providing a wonderful childhood and for always lovingly fostering and supporting my interests. I would also like to thank my undergraduate mentor Dr. Sunil David for taking me into his lab, and encouraging me to overcome my limiting beliefs and to pursue challenging goals. Without them having a career in science would not be possible. I have been very fortunate to be mentored by two wonderful professors, Dr. John Flannery and Dr. David Schaffer, during my time in graduate school. They have both provided excellent guidance and have allowed me to explore, learn, and grow. I'd like to thank John for taking me as a summer undergraduate researcher and helping me to realize my potential over the years. I will always think fondly on our conversations about science, jazz, and world travel. John fosters a very positive work environment and served as a great example on the power of networking and forming long lasting relationships. I'd like to thank Dave for his insightful guidance, always being available to help, and for serving as an excellent example of having a strong work ethic. Dave gives a lot of intellectual freedom to his students allowing them to grow into competent independent scientists and I very much appreciated that opportunity. I'd like to thank the other members of my thesis committee Marla Feller, Ehud Isacoff, and Laurent Coscoy for their feedback and thoughtful suggestions over the years.

I have been fortunate to work with a number of fantastic colleagues throughout my time at Berkeley. First, I'd like to thank Deniz Dalkara and Meike Visel for mentoring me in molecular biology, setting me up with a strong foundation of fundamental skills, and always being available to help. I'd like to thank the individuals who helped me as undergraduate researchers Emilia Zin, Nicoletta Commins, and Yebon Oh. It was a pleasure to work with all of them and it has been even more enjoyable to watch them go off and accomplish great things on their own. I'd especially like to thank Leah Byrne who I have worked closely with and whose collaboration has been an integral part to my graduate work and dissertation. It has been great to watch someone with such strong drive flourish and I have greatly benefited from our many scientific conversations and brainstorming sessions. I'd like to thank David Ojala for being a springboard for scientific ideas and for the many plasmid aliquots. I'd like to thank Kanye West and Elon Musk for serving as examples of self-actualization. Special thanks goes to Benjamin Gaub, Trevor Lee, Fakhra Khalid, Jas Mann, Melissa Kotterman, Jorge Santiago-Ortiz, Sabrina Sun, Thomas Gaj, Benjamin Epstein, Chris Habrian, and Joshua Levitz. They have all helped to contribute to my work and to my growth as a scientist.

Finally, I'd like to thank Cécile Fortuny who has been the most important source of support. She has been there for the tough times and for the celebrations. Her feedback, help, and moral support were invaluable and I will always be grateful.

Chapter 1: Introduction

In the last two decades gene therapy has moved from the realm of science fiction to an actual reality that is directly improving the lives of patients. Leading the charge are therapies designed to treat genetic diseases in the retina. This tissue has many advantages for gene therapy and serves as a beachhead for treating diseases throughout the rest of the body. Recent advances in genome editing strategies with CRISPR/Cas9 have further increased the scope of diseases to which gene therapy can be applied. Major challenges still exist in delivery and in the fundamental understanding of underlying mutations, which will be needed to maximize the potential of gene therapy. This dissertation aims to contribute to a field that is working hard to address these challenges. As this area of science continues to progress, we will soon enter an age of genetic medicine and bring treatments for the first time to patients that currently have very few therapeutic options.

The retina

Humans rely heavily on vision for perceiving the external world and this process begins in the retina. The eyes of vertebrates are based on a similar structural plan, in which light passes through a lens and creates an inverted image on the retina found at the back of the eye. The retina is the neural tissue responsible for vision and is the first step in visual processing via intricate circuitry. It is a highly organized structure with cell bodies in three different layers, which are separated by two synaptic layers. There is estimated to be over 55 distinct cell types with different cell functions. These are classified as photoreceptors, horizontal cells, bipolar cells, ganglion cells, Müller glia, and amacrine cells. The outer nuclear layer (ONL) is located furthest from the vitreoretinal interface and contains the soma of both rods and cones. Found in the inner nuclear layer (INL) are the bipolar cells, horizontal cells, amacrine cells, and Müller glia (the primary glial cell type in the retina). Ganglion cells as well as additional amacrine cells are located in the ganglion cell layer (GCL). The inner and outer plexiform layers located between the cell layers contain the synapses linking neurons together and allows for both vertical and horizontal information transfer (**Fig. 1**). Photoreceptor cells synapse onto bipolar cells, which relay information from photoreceptors to ganglion cells. The ganglion cells send information to the lateral geniculate nucleus of the thalamus and on to the visual cortex in the occipital lobe for processing, leading to the percept of sight.

Counterintuitively, the light-sensing photoreceptors are the furthest cell layer from incoming light. They are in physical contact with the retinal pigment epithelium (RPE), which supports the photoreceptors in a number of ways. The RPE forms part of the blood-retina barrier, engages in phagocytosis of the continually shed outer segments, recycles photopigments, provides nutrients, absorbs excess photons to prevent glare, and protects photoreceptors from high intensity light[1]. There are two main types of photoreceptor cells: rods, which are involved in scotopic, low light vision and cones, which are involved in photopic, high-acuity color vision. Cones can be categorized by the wavelength of light with which they have the greatest sensitivity. In humans, there are short-wavelength, medium-wavelength, and long-wavelength cones corresponding to sensitivity to blue, green, and red, respectively. Rods outnumber cones in mammals by approximately 20-fold, and are able to detect even a single photon of light[2].
Interestingly, there is evidence that cones evolved before rods, which makes intuitive

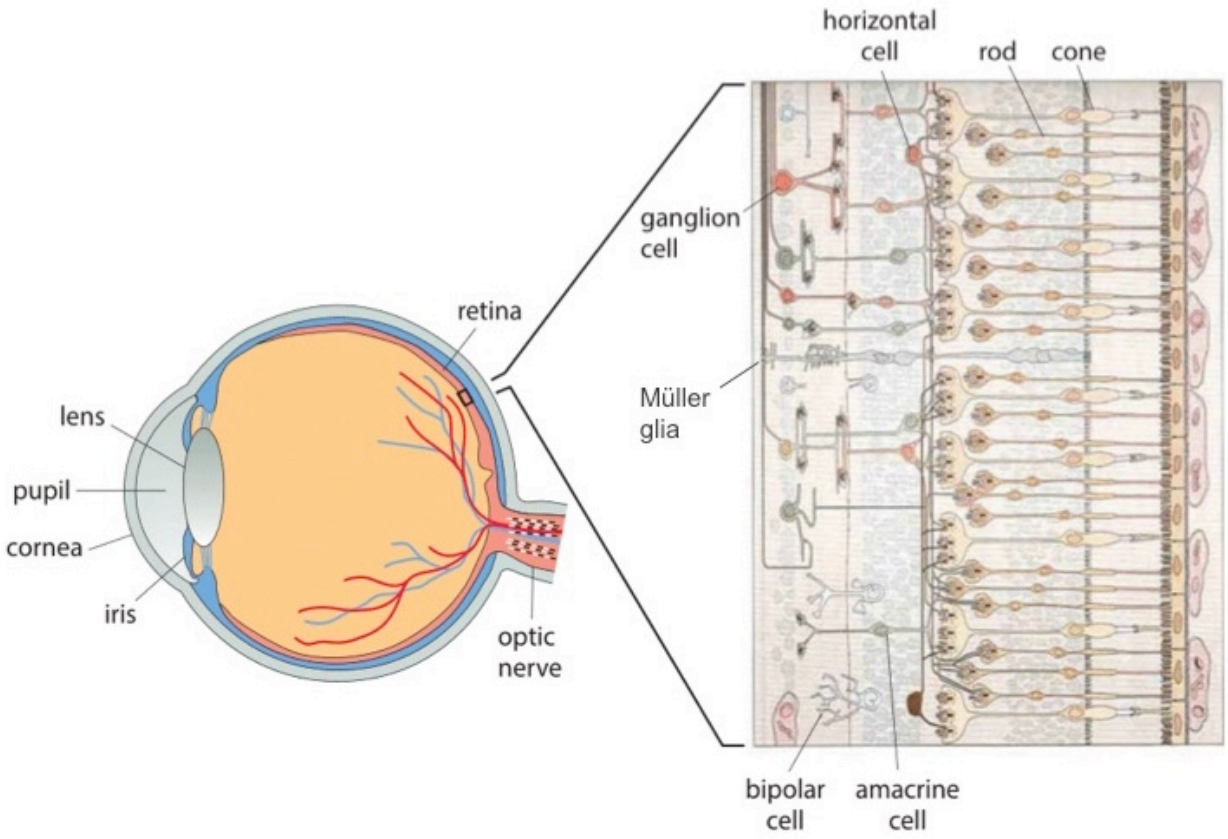


Figure 1. Organization of the retina. The eyes of vertebrates are based on a similar structural plan, in which light passes through a lens and creates an inverted image on the retina, the light sensitive neural tissue found at the back of the eye. The retina is a highly organized with cell bodies in three different layers, which are separated by two synaptic layers. The light-sensing cells, photoreceptors (rods and cones), are furthest in the cell layers from incoming light. They synapse onto bipolar cells and horizontal cells laterally interconnect and integrate across multiple photoreceptor cells. Müller glia span the entire retina and provide support for the neurons. The signal from bipolar cells is passed onto the ganglion cells either directly or indirectly via amacrine cells. The signal from ganglion cells is then carried to the brain. Adapted from Rodieck[3].

sense because the early light-sensing cell would initially evolve for periods of maximum light. Cones are associated with a complex network of postsynaptic cells while there is minimal circuitry strictly associated with rods. And although rods outnumber cones, there are 8-10 times greater number of cone driven neurons than rod driven neurons in the retina[2].

The photosensitive molecule that allows for light-sensing in the retina is 11-*cis*-retinal, found in discs in the outer segments of photoreceptor cells, that undergoes a conformational change once a photon is absorbed. This change is able to activate the protein opsin, which then signals transducin initiating the activity of cyclic-GMP (cGMP) phosphodiesterase. Interestingly, in the dark cGMP-gated ion channels remain open, leading to maintenance of the depolarized membrane potential. Upon photo-activation, the biochemical cascade results in the closing of cGMP-gated ion channels and hyperpolarization of the membrane altering the glutamate release and the ONL electrical microenvironment. In this way, photoreceptors work “backwards” by hyperpolarizing when being exposed to light[4].

Bipolar cells receive signaling from photoreceptor cells. When there is a change in glutamate release from photoreceptor cells, bipolar cells will respond depending on the receptor expressed in that particular bipolar cell. For example, ON-bipolar cells express metabotropic glutamate receptors, mainly mGluR6. The activated G-protein-coupled receptor utilizes a second messenger pathway allowing for sign reversal. ON-bipolar cells depolarize in response to decreases in glutamate, after photoreceptors are hyperpolarized by light. On the other hand, OFF-bipolar cells express an iontropic glutamate receptor where glutamate directly opens the cation channel. This subset of bipolar cells then depolarize in a response to increases in the amount of glutamate released into the synaptic cleft. ON- and OFF-bipolar cells are relatively equally represented in numbers and there are several distinct subclasses of both which allow for separate channels of information between transient and sustained signals. Bipolar cells transfer information via the release of glutamate to both amacrine and ganglion cells[5].

Horizontal cells are also able to influence bipolar cells. Horizontal cells comprise less than 5% of cells in the INL. They have a wide receptive field allowing them to gather information from a large number of photoreceptor cells as well as couple to each other and photoreceptors via gap junctions. All photoreceptors receive feedback from horizontal cells and they have traditionally been thought to enhance contrast between light and dark regions. Light that activates the center of a receptive field leads to a larger depolarization response while stimuli in the surround leads to hyperpolarization causing a center-surround organizational feature and greater contrast[6].

Amacrine cells are a diverse class of retinal neurons representing a range of functions. Most amacrine cells and all ganglion cells receive their main bipolar cell synapse from cone bipolar cells. Therefore, signals from rods reach ganglion cells via an indirect route using as its final path the cone bipolar cell axon terminals. Amacrine cells are thought to correlate firing among ganglion cells. Dopaminergic amacrine cells globally adjust the responsiveness of the retina under bright or dim light and starburst amacrine cells are important for direction selectivity. Amacrine cells make inhibitory synapses on the axon terminals of bipolar cells and in this way are able to control their output to ganglion cells[7].

Retinal ganglion cells receive their input from cone bipolar cells but most synaptic input comes from amacrine cells. Similar to bipolar cells they are also characterized by ON or OFF center-surround receptive fields. The two main types of ganglion cells in primate are the parvocelluar P cells with smaller receptive fields that are sensitive to color and the magnocellular M cells that have large receptive fields, are sensitive to color changes, and transmit signals with quick conduction velocity. The axons of the majority of retinal ganglion cells carry the totality of the information in the retina and form the optic nerve. The main efferent pathway projects to the lateral geniculate nucleus, which then projects to the visual cortex for further processing leading to conscious vision. Ganglion cells also innervate the suprachiasmatic nucleus, superior colliculus, and preoptic area, which operate below the level of consciousness. In addition, there is a subset of ganglion cells that are intrinsically photosensitive by expressing the photopigment melanopsin. This cell projects to the suprachiasmatic nucleus of the hypothalamus and the olivary pretectal nucleus to mediate circadian rhythm and pupillary light reflex, respectively[8]. The retina is often underappreciated for the amount of computations that it performs before information is transmitted to the brain. This includes many important computations such as

signal-to-noise filtering of light, motion detection, motion anticipation, saccadic suppression, and latency coding. We are only beginning to uncover the mysteries of the retina[9].

Inherited retinal degenerations

The retina is a very complex and sensitive system, especially in regards to photoreceptors, in which small defects are capable of causing cell death. This leads to loss of sight, which is devastating since humans rely heavily on vision. There are over 250 different genes associated with inherited retinal degenerations[10]. Degenerations can come in a variety of forms including light damage, oxidative stress, and mutations in genes both directly related to retina-specific pathways and throughout the body converging on the phenotype of photoreceptor cell death and loss of vision. The majority of mutations, however, are found in photoreceptors and the RPE[11]. Many affect rhodopsin, transducin, and cGMP phosphodiesterase, which are all players in the phototransduction cascade. Leber congenital amaurosis (LCA) is the leading type of congenital blindness and has been shown to affect 1/30000 people worldwide with 70% of the cases having the underlying genetic mutation identified[12]. Until recently there were no treatments for any of the 14 different LCAs[13]. Retinitis pigmentosa (RP), affecting about 1 in 4000 people worldwide, is another example of an inherited retinal degeneration characterized by the loss of rods and subsequent loss of night and peripheral vision in adolescence followed by the loss of central vision due to the death of cones[14-17]. After photoreceptor cell death, the RPE cells migrate inward towards the retinal blood vessels resulting in a spiculated, pigmented appearance on fundus images[18-20]. It is the largest and most genetically heterogeneous of the inherited retinal diseases[21]. There are three main categories of RP with 50-60% of cases able to be genotyped arising from autosomal recessive mutations, 30-40% from autosomal dominant mutations, and the remaining arising from X-linked mutations[12]. In approximately 50% of RP cases, the underlying mutation cannot be identified (**Fig. 2**).

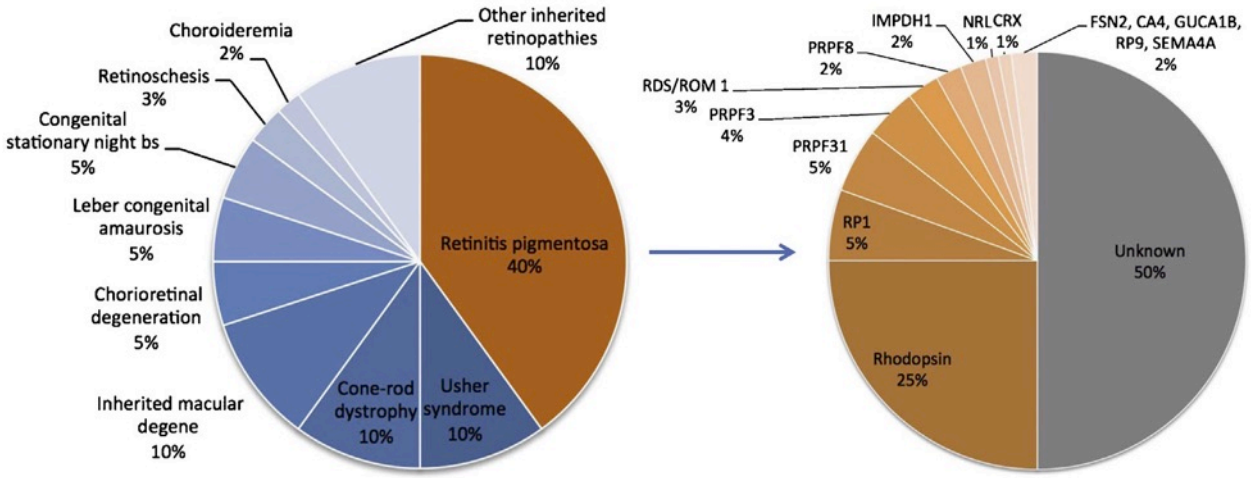


Figure 2. Percentage of broad categories of inherited retinal degenerations. There are nine broad categories of inherited retinal degenerations with retinitis pigmentosa being the most prominent. Within retinitis pigmentosa, the majority of known mutations are found in rhodopsin and nearly half still are caused by unknown idiopathic mutations. Adapted from Dalkara *et al.*[22]

Gene therapy in the retina

Gene therapy is the introduction of nucleic acids into a cell in order to prevent, halt, or reverse a pathological process by altering the expression of genes[23,24]. In the case of many recessive diseases, gene therapy can provide a normal copy of the mutated gene allowing for effective treatment. Autosomal dominant diseases require gene therapies that provide the knockdown of the mutant protein and potentially followed by replacement with a normal copy of the gene. Knockdown can be accomplished using either a specific nuclease or RNA interference (RNAi)[25]. The retina has many advantages as a tissue for gene therapy. Two eyes allow for a built-in internal control for each patient and a large number of animal models exist to serve as a proof-of-principle for different gene therapy strategies. The retina is considered immunoprivileged because it is separated from the general vasculature by the blood-retina barrier which increases safety and limits systemic exposure of the gene delivery vehicle[26]. The retina is a very accessible tissue for delivery and it can be non-invasively monitored. Also, the amount of vector required is relatively small compared to other tissues, the cell types are well characterized and understood, and the underlying gene mutations for many diseases have been elucidated.

There are multiple approaches that can be taken as gene delivery strategies. First gene replacement which is the introduction of wild-type DNA to replace the mutant endogenous activity, especially in the case of loss-of-function mutations. This has found success with diseases like X-linked retinoschisis which showed marked improvement when *RS1* was delivered and expressed by photoreceptor cells[27,28]. Dominant mutations have been more challenging in the retina but RNA-mediated gene knockdown strategies could be implemented[22]. Genome editing is becoming a more viable approach in which the genome can be permanently edited for a beneficial effect. The toolkit to accomplish this task is ever expanding with Zinc-fingers, TALENs, and now CRISPR-Cas9[29]. These can be used both to knockdown gain-of-function mutations through nonhomologous end joining or can replace specific areas of the genome through homology directed repair and providing donor DNA[30]. Specifically for the retina, this has been applied to the CEP290 mutation (LCA10) through the elimination of a cryptic splice site preventing an intronic sequence from being expressed[31]. Sometimes the underlying mutation is not well understood and a broader approach must be taken. Delivery of trophic factors has been shown to slow down the progression of degeneration and photoreceptor cell death. BDNF, CNTF, and GDNF are factors that have been shown to promote survival and/or delay cell death in dominant mutations of rhodopsin[32-36]. Expression of exogenous photosensitive ion channels in the field of optogenetics is another gene therapy approach, especially for retinas that have entirely lost photoreceptors. There are examples in animal models expressing light-gated ionotropic glutamate receptors and injected with a maleimide-azobenzene-glutamate photoswitch imbuing bipolar cells and ganglion cells with light sensitivity[37]. There is also the ectopic expression of rhodopsin in bipolar cells that led to a light response and successful visually-guided behavior studies in blind *rd1* mice[38].

AAV gene therapy

In order to deliver a therapeutic transgene, a vector, such as a virus, must be used. Adeno-associated virus (AAV) has been the vector of choice for recent clinical trials in the retina[39]. AAV is a nonpathogenic member of the *Parvoviridae* family that requires the co-infection of a

helper virus, such as adenovirus or herpes simplex virus, in order to replicate. The non-enveloped capsid is approximately 25 nm in diameter and contains a single-stranded DNA genome[40]. AAV is able to pack an incredibly dense amount of information in its genome and encodes for a number of proteins through a combination of alternative splicing, alternative start codons, and different open reading frames (ORFs). The genome contains three ORFs flanked by inverted terminal repeats (ITRs), which form hairpin structures required for viral packaging and second strand synthesis. The first open reading frame, *rep*, encodes the four proteins necessary for viral replication and integration into human chromosome 19[41]. The second ORF termed *cap* is necessary for the expression of the three viral proteins VP1, VP2, and VP3. These three structural proteins are assembled into a 60-mer to form the icosahedral viral capsid. A third ORF encoding an assembly-activating protein (AAP) needed for localizing the capsid proteins into the nucleolus and for capsid assembly has been discovered as an alternate reading frame inside of *cap* (**Fig. 3**)[42,43].

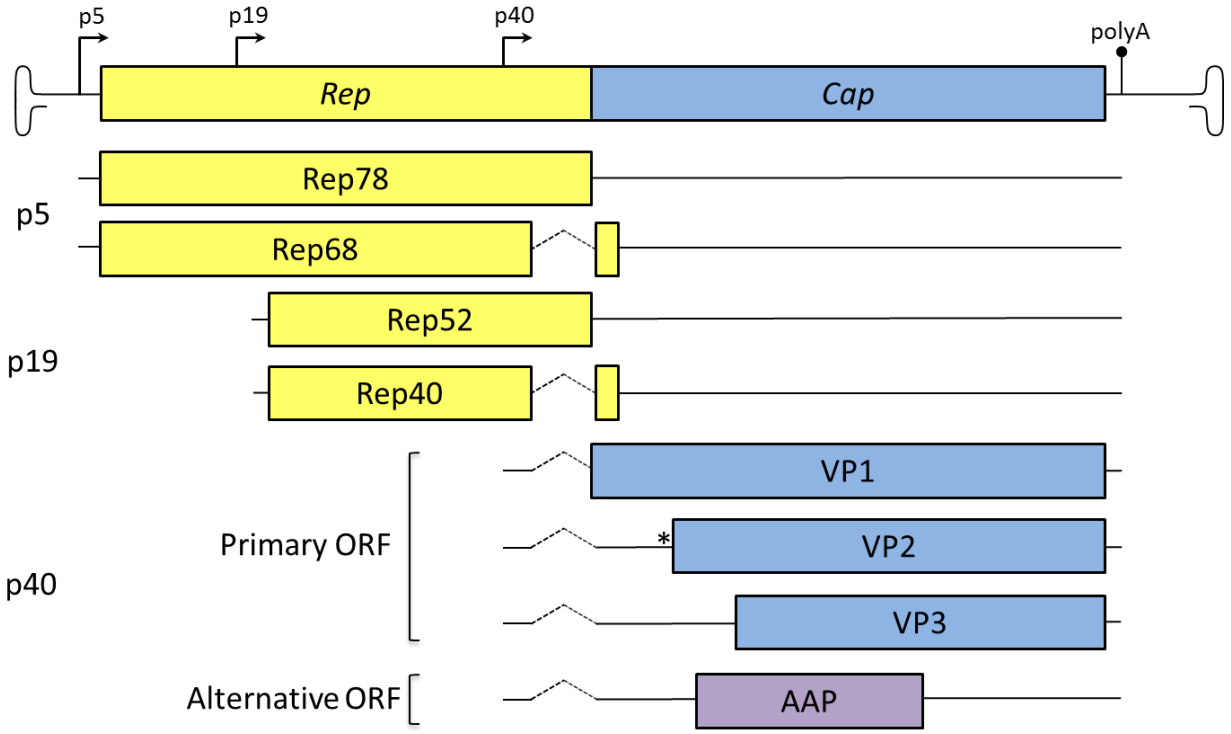


Figure 3. AAV Genome. Through the combination of two different start sites and alternative splicing, *rep* encodes for the four proteins Rep78, Rep68, Rep52, and Rep40. The p5 promoter drives Rep78 and Rep68 while the p19 promoter drives Rep52 and Rep40. These proteins are necessary for AAV genome replication. The *cap* gene encodes three different structural capsid proteins VP1, VP2, and VP3 that then form a 60-mer icosahedral viral capsid shell. All three proteins are driven by the p40 promoter and translated from a single mRNA with two splice forms. VP1 is the largest protein and comes from the longer splice form and VP2 and VP3 come from the shorter splice form with VP2 driven by the alternative start codon ACG* surrounded by an optimal Kozak context. AAP is encoded by an alternative ORF in *cap* that initiates from an unconventional CUG start site. AAP is important for viral capsid assembly promoting activity. Adapted from Ojala *et al*[44].

AAV cell specificity is governed by the ability of the capsid to transduce cells via binding to a cell surface receptor that leads to eventual entry into the nucleus. Recently, a type I transmembrane protein called AAV receptor (AAVR) was identified which serves as a universal

AAV receptor necessary for rapid endocytosis from the plasmid membrane and trafficking to the *trans*-Golgi network for AAV serotypes in general[45]. There is evidence that supports a role for both the clathrin-coated pit pathway and clathrin-independent pathways[46,47,48]. Following escape from the early endosome, AAV traffics to the perinuclear area. There is evidence that supports both uncoating prior to trafficking into the nucleus or AAV uncoating after trafficking into the nucleus to facilitate viral DNA entering into the nucleus[49-51]. Once the single-stranded DNA has entered into the nucleus it undergoes second-strand synthesis converting the single-stranded DNA into double-stranded transcriptionally available DNA[52,53]. Absence of a helper virus causes AAV to enter into a latent life cycle. This leads to the integration of the viral genome selectively into the AAV1S locus on human chromosome 19, which matches a sequence found in the AAV ITRs[54-56].

AAV has many advantages as a viral vector for inherited diseases affecting the retina including the ability to infect quiescent cells, long-term episomal transgene expression, various serotypes that exhibit tropism for different subsets of retinal cells, low immunogenicity, and low insertional mutagenesis risk. AAV2, the best characterized AAV serotype, has been used for the rescue of multiple animal models and in the seminal clinical trials for Leber congenital amaurosis type 2 (LCA2), an autosomal recessive retinal dystrophy caused by a mutation in *RPE65*, by three independent groups in 2008[57-60]. In the LCA2 trials, subretinal administration of the AAV2 vector carrying *RPE65* was well tolerated and led to marked improvement in vision, especially in younger patients. A follow-up five years later reported no adverse events for the patients[22]. In the initial clinical trials only one eye was treated but for some patients the second eye was treated months or years later. The second injection resulted in minimal inflammation and improved visual function proving that subretinal readministration is safe and effective[61]. It is important to note that in both human patients and in older canine models that received this gene therapy treatment that degeneration of photoreceptors continued at the same rate even though there was observed long-term improvements in vision[62].

To render a recombinant AAV amenable for gene therapy, the viral genome is substituted with the gene of interest driven by a promoter and flanked by ITRs. The ITRs must be *cis* elements because they are recognized by the Rep proteins for packaging in the capsid but the *cap* and *rep* genes can be supplied in *trans*. A helper plasmid that supplies the necessary elements from adenovirus is also required, so a triple transfection is used to produce recombinant AAV[63]. AAV infection of non-dividing cells results in persistent, long-term episomal expression that is stabilized by the concatemerization of the transgenes. Over 100 different AAVs have been isolated from a variety of different species leading to diverse tropisms in different animals, tissues, and cell types. A number of receptors for AAV, including heparan sulfate proteoglycan, have been identified and work continues to fully understand the underlying mechanisms of cell entry and trafficking[64,65].

Specifically for the retina, different capsids have different natural tropisms for different subsets of cells. The tropism also depends on the route of administration of the vector during surgery as well. AAV2 leads to strong expression in ganglion cells when injected into the vitreous[66]. AAV4 is able to efficiently transduce RPE cells when injected subretinally, while AAV5 has strong tropism for photoreceptors when subretinally injected[67]. AAV8 is capable of infecting Müller glia when injected subretinally, as well as RPE and photoreceptors[68]. AAV9 has

the unique property of crossing the blood-retina barrier when delivered intravenously and has been observed to have a greater affinity for cones over rods when injected subretinally[69].

Engineering AAV for gene therapy

Fundamental understanding of AAV capsid structure can lead to a rational design approach to increase the efficacy of AAV for retinal gene therapy. Tyrosine residues on the surface of the viral capsid are susceptible to phosphorylation by tyrosine kinases labeling the capsid for ubiquitination and proteasome-mediated degradation. By mutating the tyrosine residues to phenylalanines, phosphorylation is avoided and transgene expression can be increased 10- to 30-fold[70]. This has also been shown to reduce the likelihood of an immune response to the capsid as well[71]. Dalkara *et al.* demonstrated in neonatal mice that mutating two tyrosines on the capsid of AAV9 resulted in increased transgene expression[69]. The tyrosine mutants do not significantly alter tropism but could reduce the possibility of an immune response by lowering the effective dose required.

Despite the success of natural and rationally designed AAV vectors, additional properties that are unlikely to offer a selective advantage to natural AAV serotypes or are counter to natural selection are still required to meet the needs of human therapeutics. In addition, the incomplete mechanistic understanding and complexity of the structure-function relationships of AAV particles limit rational design strategies. Another approach that has found success is directed evolution, which is a strategy that mimics natural evolution through genetic diversification and selection pressure to improve function. But in the case of directed evolution, the selection pressure is controlled in the laboratory to result in AAV variants with improved clinical properties. Directed evolution has been used to generate multiple different AAV variants with improved properties such as enhanced neutralizing antibody evasion and cell specificity[72].

Directed evolution begins with the generation of large libraries of AAV variants. This can be accomplished through error-prone polymerase chain reaction (PCR), shuffling of DNA, or the insertion of peptide sequences into *cap*[73,74]. More advanced libraries have been generated through the use of SCHEMA which is a computational method that predicts the optimal crossover points for DNA shuffling and leads to a larger number of viable variants[75]. Because evolution traverses a fitness landscape, certain serotypes of AAV may require too large of an evolutionary jump to reach the desired trait[76], which would be difficult to achieve with natural AAV serotypes as the starting point. To address this issue in AAV, a library based on an ancestral reconstruction was also developed which is more thermally stable, promiscuous, and allows for a greater exploration of evolutionary space[77]. In the next step in the directed evolution process, highly diverse AAV libraries are then packaged in such a way as to have each library genome contain the *cap* sequence that encodes for the capsid shell it is packaged in. In this way, the genome can serve as a barcode and link the genotype and phenotype of the virus. The virus then undergoes a selective pressure, such as specific cell tropism or neutralizing antibody evasion. The *cap* sequences are recovered, PCR amplified, and packaged again. Iterative rounds of selection are used to drive convergence on the top performing variants. An additional round of mutations can be applied in the middle of the selection process to increase diversity. The top variants are then individually screened for the desired trait and insights on novel structure function relationships can be gleaned (**Fig. 4**).

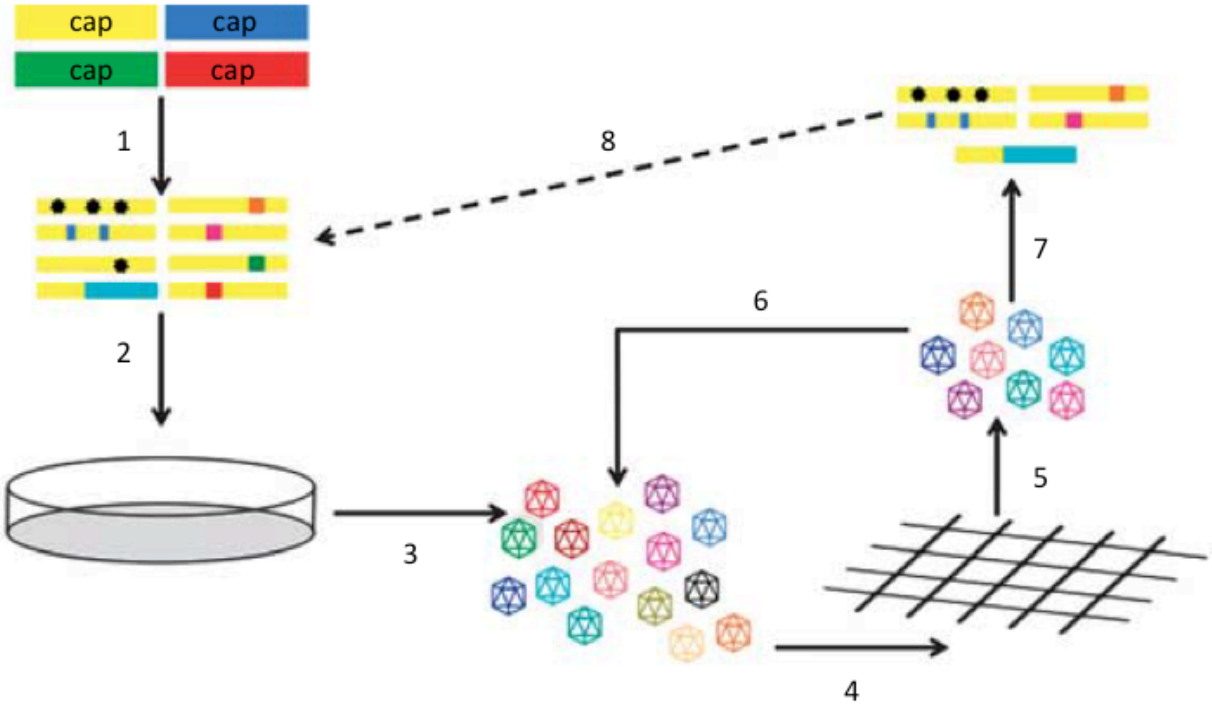


Figure 4. Directed evolution schematic. 1) The AAV *cap* libraries are generated through the introduction of mutations, inserting of peptides, or shuffling serotypes. **2)** Triple transfection using HEK293T cells is used to produce viruses, such that each genome represents the capsid shell it is packaged in. **3)** The viruses are harvested and purified. **4)** Selection pressure for a desired trait is applied to the library. **5)** Successful variants are recovered via PCR amplification of *cap* sequences. **6)** Iterative rounds of selection are performed to enrich for successful variants. **7)** Isolated *cap* genes are recovered and mutagenized to act as a new starting point for evolution. **8)** Additional rounds of selection are performed on mutagenized variants until sequence convergence is reached. Adapted from Bartel *et al.*[78]

Directed evolution in the retina was used to overcome significant challenges in AAV gene delivery including engineering AAV variants that are able to be administered in a safer, more effective manner and increasing cell specificity. There are multiple surgical routes to administer AAV to the retina including subretinal and intravitreal injections. Positioning a needle between the photoreceptor layer and the RPE layer, and then introducing a "bleb" of injected liquid is a subretinal injection. The advantage of subretinal injections is that the virus is delivered locally to the necessary cells. The disadvantages is that only the cells in the bleb are transduced and the risk of retinal detachment in a retina that is already damaged from degeneration. An intravitreal injection is the direct delivery into the vitreous humor. The advantages of intravitreal injections are that the surgical procedure is less invasive, there is the potential for pan-retinal expression, and there is no risk of inducing retinal detachment. The challenge has been that is it difficult for natural AAV serotypes to cross the inner limiting membrane (ILM). A studied showed that the ILM is the main barrier by intravitreally injecting fluorescently labeled AAVs and observing an accumulation at the ILM. But, mild digestion of the ILM using a protease significantly enhanced transduction in the retina[79]. In order to overcome the ILM, directed evolution was utilized to engineer an AAV variant that is capable of transducing the retina via an intravitreal injection. Six rounds of selection were performed in transgenic mice with GFP

expression in photoreceptors so that variants capable of reaching photoreceptors could be recovered from GFP-positive cells via fluorescence-activated cell sorting. This led to 7m8, which is an AAV2-based variant that has a seven amino acid sequence inserted on loop four. This variant showed improved success in two different inherited retinal degeneration animal models: Leber congenital amaurosis and X-linked retinoschisis. The vector showed improved transduction in the primate retina as well[80]. Another key example of directed evolution in the retina was the development of a Müller glia-specific variant called ShH10. Directed evolution was performed on primary human astrocytes and then, due to shared properties, screened for Müller cell tropism *in vivo*[74,81,82]. Müller glia targeting is important because Müller cells span the entire retina and would therefore be the ideal target for gene delivery of secretable neuroprotective factors to slow or halt retinal degeneration. ShH10 is closely related to AAV6 and showed, via an intravitreal injection in the rat retina, a transduction profile of 94% Müller cells, 2% interneurons, and 4% retinal ganglion cells.

This dissertation focuses on the development of novel AAV vectors for the treatment of retinal degenerations and better understanding of the fundamental biology of proteins involved in retinal degeneration in order to identify and develop effective gene therapy strategies. The experiments describe both the adaptability of AAV and the limitations of the underlying properties of AAV. This work emphasizes the importance of having a deep understanding of the mechanism of directed evolution as an engineering tool leading to AAV variants that have the best chance of maximizing clinical outcomes for patients suffering from retinal degenerations. They also lay a foundation of understanding that can lead to additional gene therapy treatments. We are entering a phase of exponential growth in the gene therapy field and clinical trials in the retina are paving the way for a better future not only for patients who suffer from genetic eye diseases but also for patients who suffer from a range of genetic diseases.

References

1. Sparrow, J. R., Hicks, D. & Hamel, C. P. The retinal pigment epithelium in health and disease. *Curr. Mol. Med.* **10,** 802–823 (2010).
2. Masland, R. H. The fundamental plan of the retina. *Nature Neuroscience* **4,** 877–886 (2001).
3. Rodieck, R. W. The First Steps in Seeing. *Sinauer* 1–8 (1998).
4. Koch, K.-W. & Dell'Orco, D. Protein and Signaling Networks in Vertebrate Photoreceptor Cells. *Front Mol Neurosci* **8,** 67 (2015).
5. Euler, T., Haverkamp, S., Schubert, T. & Baden, T. Retinal bipolar cells: elementary building blocks of vision. *Nat. Rev. Neurosci.* **15,** 507–519 (2014).
6. Poché, R. A. & Reese, B. E. Retinal horizontal cells: challenging paradigms of neural development and cancer biology. *Development* **136,** 2141–2151 (2009).
7. Masland, R. H. The tasks of amacrine cells. *Vis. Neurosci.* **29,** 3–9 (2012).
8. Sanes, J. R. & Masland, R. H. The types of retinal ganglion cells: current status and implications for neuronal classification. *Annu. Rev. Neurosci.* **38,** 221–246 (2015).
9. Gollisch, T. & Meister, M. Eye smarter than scientists believed: neural computations in circuits of the retina. *Neuron* **65,** 150–164 (2010).
10. Hafler, B. P. CLINICAL PROGRESS IN INHERITED RETINAL DEGENERATIONS: GENE THERAPY CLINICAL TRIALS AND ADVANCES IN GENETIC SEQUENCING. *Retina (Philadelphia, Pa.)* **37,** 417–423 (2017).
11. Gupta, P. R. & Huckfeldt, R. M. Gene therapy for inherited retinal degenerations: initial successes and future challenges. *J Neural Eng* **14,** 051002 (2017).
12. Sahel, J.-A., Marazova, K. & Audo, I. Clinical characteristics and current therapies for inherited retinal degenerations. *Cold Spring Harbor Perspectives in Medicine* **5,** a017111 (2014).
13. Hollander, den, A. I., Roepman, R., Koenekoop, R. K. & Cremers, F. P. M. Leber congenital amaurosis: Genes, proteins and disease mechanisms. *Progress in Retinal and Eye Research* **27,** 391–419 (2008).
14. Berson, E. L. Retinitis pigmentosa. The Friedenwald Lecture. *Investigative Ophthalmology & Visual Science* **34,** 1659–1676 (1993).
15. Pagon, R. A. Retinitis pigmentosa. *Survey of Ophthalmology* **33,** 137–177 (1988).
16. van Soest, S., Westerveld, A., de Jong, P. T., Bleeker-Wagemakers, E. M. & Bergen, A. A. Retinitis pigmentosa: defined from a molecular point of view. *Survey of Ophthalmology* **43,** 321–334 (1999).
17. Milam, A. H., Li, Z. Y. & Fariss, R. N. Histopathology of the human retina in retinitis pigmentosa. *Progress in Retinal and Eye Research* **17,** 175–205 (1998).
18. Stone, J. L., Barlow, W. E., Humayun, M. S., de Juan, E. & Milam, A. H. Morphometric analysis of macular photoreceptors and ganglion cells in retinas with retinitis pigmentosa. *Arch. Ophthalmol.* **110,** 1634–1639 (1992).
19. Li, Z. Y., Possin, D. E. & Milam, A. H. Histopathology of bone spicule pigmentation in retinitis pigmentosa. *Ophthalmology* **102,** 805–816 (1995).
20. Rodrigues, M. M. *et al.* Dominantly inherited retinitis pigmentosa. Ultrastructure and biochemical analysis. *Ophthalmology* **92,** 1165–1172 (1985).
21. Sullivan, L. S. & Daiger, S. P. Inherited retinal degeneration: exceptional genetic and clinical heterogeneity. *Mol Med Today* **2,** 380–386 (1996).

22. Dalkara, D. & Sahel, J.-A. Comptes Rendus Biologies. *Comptes rendus - Biologies* **337,** 185–192 (2014).
23. Naldini, L. Gene therapy returns to centre stage. *Nature* **526,** 351–360 (2015).
24. Somia, N. & Verma, I. M. Gene therapy: trials and tribulations. *Nat Rev Genet* **1,** 91–99 (2000).
25. Kay, M. A. State-of-the-art gene-based therapies: the road ahead. *Nat Rev Genet* **12,** 316–328 (2011).
26. Petit, L. & Punzo, C. Gene therapy approaches for the treatment of retinal disorders. *Discov Med* **22,** 221–229 (2016).
27. Byrne, L. C. *et al.* Retinoschisin gene therapy in photoreceptors, Müller gliaor all retinal cells in the. *Gene Ther* 1–8 (2014). doi:10.1038/gt.2014.31
28. Zeng, Y. *et al.* Retinal Structure and Gene Therapy Outcome in Retinoschisin-Deficient Mice Assessed by Spectral-Domain Optical Coherence Tomography. *Investigative Ophthalmology & Visual Science* **57,** OCT277–87 (2016).
29. Ovando-Roche, P., Georgiadis, A., Smith, A. J., Pearson, R. A. & Ali, R. R. Harnessing the Potential of Human Pluripotent Stem Cells and Gene Editing for the Treatment of Retinal Degeneration. *Curr Stem Cell Rep* **3,** 112–123 (2017).
30. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* (2013). doi:10.1126/science.1231143
31. Ruan, G.-X. *et al.* CRISPR/Cas9-Mediated Genome Editing as a Therapeutic Approach for Leber Congenital Amaurosis 10. *Molecular Therapy* **25,** 331–341 (2017).
32. Liang, F. Q. *et al.* Long-term protection of retinal structure but not function using RAAV.CNTF in animal models of retinitis pigmentosa. *Mol. Ther.* **4,** 461–472 (2001).
33. Okoye, G. *et al.* Increased expression of brain-derived neurotrophic factor preserves retinal function and slows cell death from rhodopsin mutation or oxidative damage. *J. Neurosci.* **23,** 4164–4172 (2003).
34. Buch, P. K. *et al.* In contrast to AAV-mediated Cntf expression, AAV-mediated Gdnf expression enhances gene replacement therapy in rodent models of retinal degeneration. *Mol. Ther.* **14,** 700–709 (2006).
35. Dalkara, D. *et al.* AAV Mediated GDNF Secretion From Retinal Glia Slows Down Retinal Degeneration in a Rat Model of Retinitis Pigmentosa. *Molecular Therapy* **19,** 1602–1608 (2011).
36. McGee Sanftner, L. H., Abel, H., Hauswirth, W. W. & Flannery, J. G. Glial cell line derived neurotrophic factor delays photoreceptor degeneration in a transgenic rat model of retinitis pigmentosa. *Mol. Ther.* **4,** 622–629 (2001).
37. Gaub, B. M. *et al.* Restoration of visual function by expression of a light-gated mammalian ion channel in retinal ganglion cells or ON-bipolar cells. *Proceedings of the National Academy of Sciences* **111,** E5574–83 (2014).
38. Gaub, B. M., Berry, M. H., Holt, A. E., Isacoff, E. Y. & Flannery, J. G. Optogenetic Vision Restoration Using Rhodopsin for Enhanced Sensitivity. *Molecular Therapy* **23,** 1562–1571 (2015).
39. Grob, S. R., Finn, A., Papakostas, T. D. & Eliott, D. Clinical Trials in Retinal Dystrophies. *Middle East Afr J Ophthalmol* **23,** 49–59 (2016).
40. Schaffer, D. V., Koerber, J. T. & Lim, K.-I. Molecular Engineering of Viral Gene Delivery Vehicles. *Annu. Rev. Biomed. Eng.* **10,** 169–194 (2008).
41. Pereira, D. J., McCarty, D. M. & Muzyczka, N. The adeno-associated virus (AAV) Rep

protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. *J. Virol.* **71,** 1079–1088 (1997).
42. Naumer, M. *et al.* Properties of the adeno-associated virus assembly-activating protein. *J. Virol.* **86,** 13038–13048 (2012).
43. Sonntag, F. *et al.* The Assembly-Activating Protein Promotes Capsid Assembly of Different Adeno-Associated Virus Serotypes. *J. Virol.* **85,** 12686–12697 (2011).
44. Ojala, D. S., Amara, D. P. & Schaffer, D. V. Adeno-associated virus vectors and neurological gene therapy. *Neuroscientist* **21,** 84–98 (2015).
45. Pillay, S. *et al.* An essential receptor for adeno-associated virus infection. *Nature* **530,** 108–112 (2016).
46. Bartlett, J. S., Wilcher, R. & Samulski, R. J. Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J. Virol.* **74,** 2777–2785 (2000).
47. Duan, D. *et al.* Dynamin is required for recombinant adeno-associated virus type 2 infection. *J. Virol.* **73,** 10371–10376 (1999).
48. Nonnenmacher, M. & Weber, T. Adeno-associated virus 2 infection requires endocytosis through the CLIC/GEEC pathway. *Cell Host Microbe* **10,** 563–576 (2011).
49. Sonntag, F., Bleker, S., Leuchs, B., Fischer, R. & Kleinschmidt, J. A. Adeno-associated virus type 2 capsids with externalized VP1/VP2 trafficking domains are generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus. *J. Virol.* **80,** 11040–11054 (2006).
50. Johnson, J. S. & Samulski, R. J. Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus. *J. Virol.* **83,** 2632–2644 (2009).
51. Johnson, J. S. *et al.* Mutagenesis of adeno-associated virus type 2 capsid protein VP1 uncovers new roles for basic amino acids in trafficking and cell-specific transduction. *J. Virol.* **84,** 8888–8902 (2010).
52. Lux, K. *et al.* Green fluorescent protein-tagged adeno-associated virus particles allow the study of cytosolic and nuclear trafficking. *J. Virol.* **79,** 11776–11787 (2005).
53. 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,** 3227–3234 (1996).
54. Kotin, R. M. *et al.* Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. U.S.A.* **87,** 2211–2215 (1990).
55. Duan, D. *et al.* Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. *J. Virol.* **72,** 8568–8577 (1998).
56. Nakai, H., Storm, T. A. & Kay, M. A. Recruitment of single-stranded recombinant adeno-associated virus vector genomes and intermolecular recombination are responsible for stable transduction of liver in vivo. *J. Virol.* **74,** 9451–9463 (2000).
57. Bainbridge, J. W. B. *et al.* Effect of Gene Therapy on Visual Function in Leber's Congenital Amaurosis. *N Engl J Med* **358,** 2231–2239 (2008).
58. Cideciyan, A. V. *et al.* Human gene therapy for RPE65 isomerase deficiency activates the retinoid cycle of vision but with slow rod kinetics. *Proceedings of the National Academy of Sciences* **105,** 15112–15117 (2008).
59. Hauswirth, W. W. *et al.* Treatment of Leber Congenital Amaurosis Due to RPE65Mutations by Ocular Subretinal Injection of Adeno-Associated Virus Gene Vector: Short-Term Results of a Phase I Trial. *Human Gene Therapy* **19,** 979–990 (2008).

60. Maguire, A. M. *et al.* Safety and Efficacy of Gene Transfer for Leber's Congenital Amaurosis. *N Engl J Med* **358,** 2240–2248 (2008).
61. Bennett, J. *et al.* AAV2 gene therapy readministration in three adults with congenital blindness. *Science Translational Medicine* **4,** 120ra15 (2012).
62. Cideciyan, A. V. *et al.* Human retinal gene therapy for Leber congenital amaurosis shows advancing retinal degeneration despite enduring visual improvement. *Proceedings of the National Academy of Sciences* **110,** E517–25 (2013).
63. Gonçalves, M. A. F. V. Adeno-associated virus: from defective virus to effective vector. *Virol. J.* **2,** 43 (2005).
64. Summerford, C. & Samulski, R. J. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J. Virol.* **72,** 1438–1445 (1998).
65. Mietzsch, M., Broecker, F., Reinhardt, A., Seeberger, P. H. & Heilbronn, R. Differential adeno-associated virus serotype-specific interaction patterns with synthetic heparins and other glycans. *J. Virol.* **88,** 2991–3003 (2014).
66. Flannery, J. G. *et al.* Efficient photoreceptor-targeted gene expression in vivo by recombinant adeno-associated virus. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 6916–6921 (1997).
67. Yang, G. S. *et al.* Virus-mediated transduction of murine retina with adeno-associated virus: effects of viral capsid and genome size. *J. Virol.* **76,** 7651–7660 (2002).
68. Black, A. *et al.* Adeno-associated virus 8-mediated gene therapy for choroideremia: preclinical studies in in vitro and in vivo models. *J Gene Med* **16,** 122–130 (2014).
69. Dalkara, D. *et al.* Enhanced gene delivery to the neonatal retina through systemic administration of tyrosine-mutated AAV9. **19,** 176–181 (2011).
70. Zhong, L. *et al.* Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proceedings of the National Academy of Sciences* **105,** 7827–7832 (2008).
71. Martino, A. T. *et al.* Engineered AAV vector minimizes in vivo targeting of transduced hepatocytes by capsid-specific CD8+ T cells. *Blood* **121,** 2224–2233 (2013).
72. Maheshri, N., Koerber, J. T., Kaspar, B. K. & Schaffer, D. V. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat. Biotechnol.* **24,** 198–204 (2006).
73. Koerber, J. T., Maheshri, N., Kaspar, B. K. & Schaffer, D. V. Construction of diverse adeno-associated viral libraries for directed evolution of enhanced gene delivery vehicles. *Nat Protoc* **1,** 701–706 (2006).
74. Koerber, J. T., Jang, J.-H. & Schaffer, D. V. DNA Shuffling of Adeno-associated Virus Yields Functionally Diverse Viral Progeny. *Molecular Therapy* **16,** 1703–1709 (2008).
75. 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. *Molecular Therapy* (2017). doi:10.1016/j.ymthe.2017.09.006
76. Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **10,** 866–876 (2009).
77. Santiago-Ortiz, J. *et al.* AAV ancestral reconstruction library enables selection of broadly infectious viral variants. *Gene Ther* **22,** 934–946 (2015).
78. Bartel, M. A., Weinstein, J. R. & Schaffer, D. V. Directed evolution of novel adeno-associated viruses for therapeutic gene delivery. **19,** 694–700 (2012).
79. Dalkara, D. *et al.* Inner Limiting Membrane Barriersto AAV-mediated Retinal TransductionFrom the Vitreous. *Molecular Therapy* **17,** 2096–2102 (2009).

80. Dalkara, D. *et al.* In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous. *Science Translational Medicine* **5,** 189ra76 (2013).
81. Koerber, J. T. *et al.* Molecular Evolution of Adeno-associated Virus for Enhanced Glial Gene Delivery. *Molecular Therapy* **17,** 2088–2095 (2009).
82. Klimczak, R. R., Koerber, J. T., Dalkara, D., Flannery, J. G. & Schaffer, D. V. A novel adeno-associated viral variant for efficient and selective intravitreal transduction of rat Müller cells. *PLoS ONE* **4,** e7467 (2009).

Chapter 2: Efficient AAV gene delivery to canine and primate outer retina from intravitreal injection

Manuscript in preparation as a letter

Abstract

Efficient delivery to the outer retina remains a significant barrier to effective and permanent gene therapies for retinal disease. Here, we apply the process of directed evolution, guided by deep sequencing, and followed by secondary selection of high-performing vectors with a GFP-barcoded library, to create AAV viral capsids with new capabilities of overcoming anatomic barriers and delivering genes to the outer retina. Direct comparison of gene delivery in canines, primates and mice revealed important species differences in vector performance. Intravitreal injection of the vectors resulted in efficient targeting of photoreceptors, bipolar cells and RPE cells in dog retina and substantially increased gene delivery to primate foveal cones. These new viral vectors will enable permanent and pan-retinal gene therapies targeting outer retina cell types. Further, these data support that directed evolution is a powerful approach for development of AAV vectors specialized for a multitude of physical barriers and cellular targets.

Introduction

Inherited retinal degeneration (RD) is caused by mutations in >50 genes[1], the majority of which are expressed in photoreceptors or RPE, leading to vision loss and dramatically decreasing quality of life. Gene therapy is a promising approach to treating RD, although efficient gene delivery to the outer retina remains a significant hurdle. Here, we used directed evolution, guided by deep sequencing, to create AAVs with the ability to efficiently deliver genes to outer retina in canines and primates via intravitreal injection. Deep sequencing revealed the dynamics of directed evolution. Then, secondary screening of subsets of AAVs, barcoded, pooled and co-injected into mouse, canine or primate retinas, highlighted species specificities of the evolved AAV vectors, and revealed the most efficient AAVs for photoreceptors, RPE, and bipolar cell gene delivery in dog or primate retina. Efficient gene delivery to the outer retina may enable more effective, permanent gene therapies for RD.

Canine directed evolution

Directed evolution has created efficient viral vectors for outer retinal gene delivery in mice[2-4]. However, the anatomical features and structural barriers that exist in large animal retinas are substantially different from mouse[5], and vectors that were selected for in the context of mouse retina are not optimally efficient in large animal models[4]. Large animals have a specialized area for high acuity vision (a retinal streak in dogs or fovea in primates) and a thicker inner limiting membrane[6]. In order to create AAV vectors with the enhanced ability to target the outer retina in large animals with eyes more like humans, we used directed evolution, performed in canines, a well-established preclinical model for RD (**Fig. 1**)[7,8].

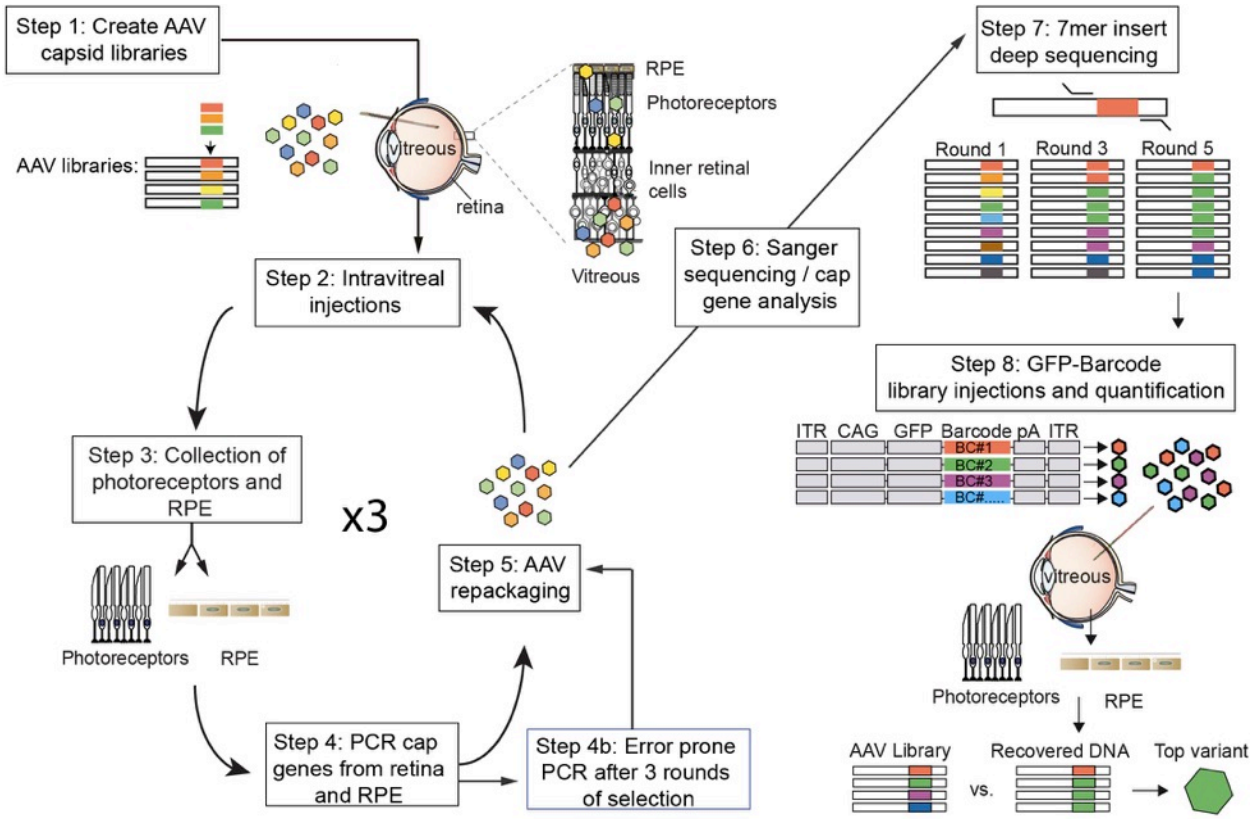


Figure 1. Summary of the experimental approach. Highly diverse (~E+7) libraries of AAV variants were packaged such that each virus contained a genome encoding its own capsid. Libraries were pooled and injected intravitreally in canines or primates. After AAV infection had occurred, retinal tissue and RPE cells were collected, and cap gene variants were PCR amplified, then re-cloned and repackaged for the following round of injection. Five rounds of selection were performed, with error prone PCR performed after the third round to introduce additional diversity into the library. Following rounds of selection, libraries were subjected to deep sequencing, in order to analyze convergence of the libraries and behavior of every individual variant in the library. Variants were then chosen for secondary screening, based on their increase in representation relative to the original library. A barcoded library containing top candidate vectors was then constructed, in which the variants were packaged with a construct encoding GFP fused to a unique DNA barcode sequence. Variants were pooled in equal amounts, and injected intravitreally. Retinal cells (bipolar cells, photoreceptors or RPE cells) were harvested, and GFP barcodes were PCR amplified from the collected tissue. Deep sequencing was used to quantify the relative abundance of barcodes. The top-performing variants were indicated by the greatest fold increase of barcodes recovered from collected tissue relative to the injected library.

AAV libraries were packaged so that each AAV particle contained the genome encoding its own capsid. AAV libraries included in the screen were an AAV2 error prone library[9], AAV2-7mer[10], Loop Swap[11] and SCHEMA libraries[12]. Libraries were pooled and injected intravitreally into both eyes of a WT dog. All large animals used in these studies were previously screened for neutralizing antibodies to AAV2, with titers <1:25. After 6 weeks, retinas were collected and AAV cap genes were recovered using PCR from total neuroretina and RPE separately. AAV genomes were repackaged and re-injected, and an additional 2 rounds of selection were conducted. Following the third round of selection, a round of error prone PCR was performed in order to introduce additional diversity into the library. Two additional rounds of selection were performed, for a total of 5 rounds of selection (**Supplementary Table 1**), after which Sanger sequencing revealed convergence of the library onto a single consensus sequence, originating from the AAV2-7mer library (AAV2-LAHQDTTKNA). Injection of this variant into the canine

eye revealed partial targeting of the outer retina. Interestingly, intravitreal injection of the variant in mouse retina did not result in appreciable GFP expression (**Supplementary Fig. 1**), suggesting that variants selected for in the context of canine retina may not be optimized for other species.

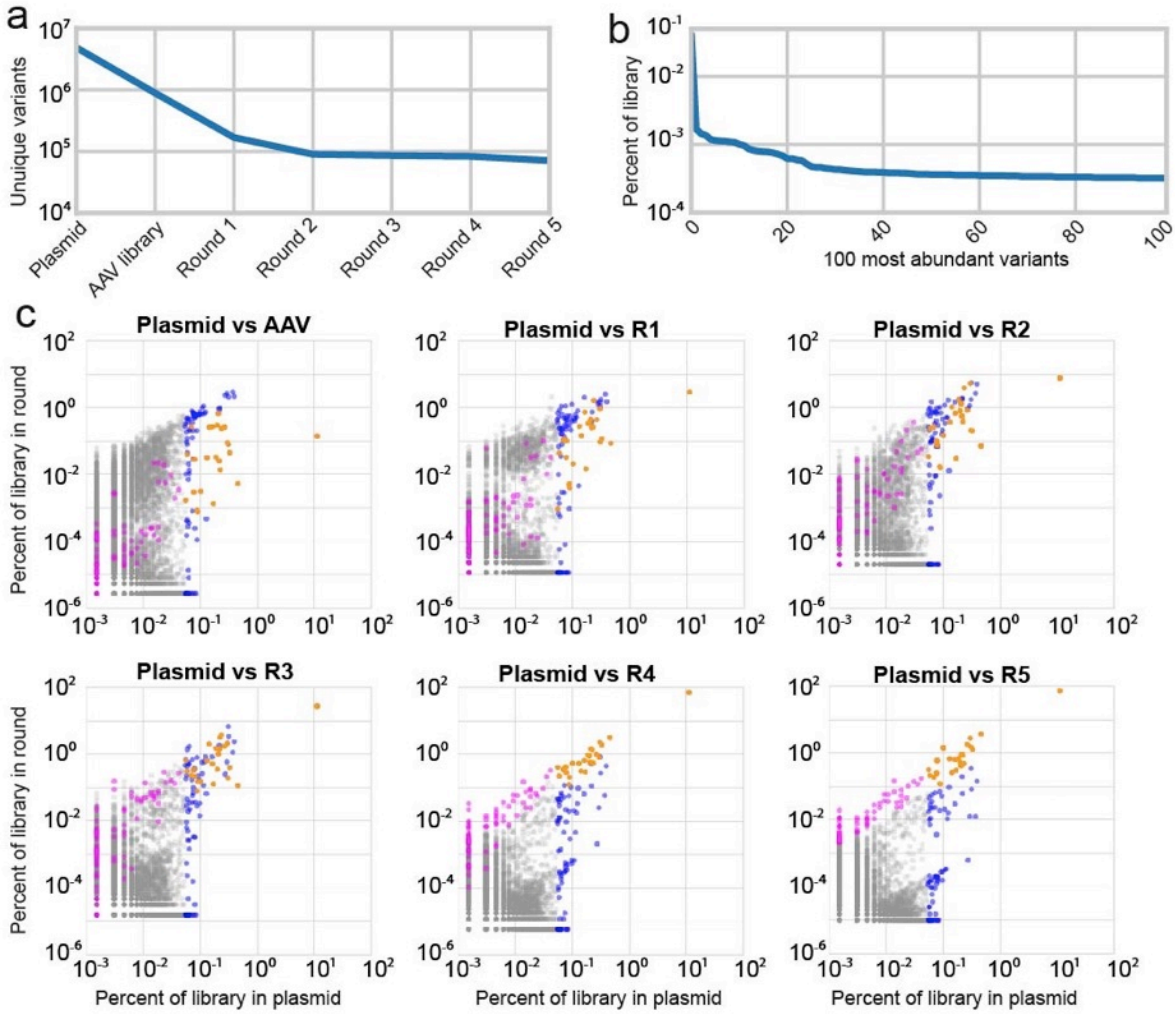


Figure 2. Directed evolution of AAV in canine retina. a) Deep sequencing revealed that the number of unique variants in the AAV2-7mer library was ~4.8E+6 variants, which converged to ~7.1E+4 variants in the final round of selection. **b)** Analysis of the copy number of variants in the plasmid library revealed that variants were not equally represented, with some variants represented by >100X copies than others. **c)** Scatterplots illustrate the behavior of individual variants through each round of selection. Each dot represents an individual variant. Variants overrepresented in the original library are colored blue. Variants that had the greatest fold increase in representation in the final round of selection are colored magenta. Variants that were both overrepresented in the original library and increased significantly in representation over rounds of selection (an overlap of blue and magenta) are colored orange.

Next, in order to understand the dynamics of directed evolution selections, and to identify additional vectors with enhanced ability to target the outer retina in canines, we used high throughput Illumina sequencing to track convergence of variants in the AAV2-7mer library, which made up 100% of clones following the last round of selection. Deep sequencing showed that the libraries had an initial diversity of 4.8E+6 variants, which decreased over five rounds of selection to 7.1E+4 variants. Thus, a large subset of variants was carried through the rounds of

selection, although only a single sequence was visible through Sanger sequencing (**Fig 2a**). Additionally, variants were not equally represented in initial libraries. A minority of variants in the library were significantly overrepresented, suggesting that these variants may be better positioned to expand in their representation as rounds of selection continue (**Fig 2b**).

Additional promising variants were then identified from the screen, based on their fold increase relative to the starting library over five rounds of selection. Top variants were chosen based on the following calculation: % of total in final round / % of total in library (**Supplementary Table 2**). A pseudo-count of 1 was added to each variant in every round, in order to mitigate effects of small number increases and allow analysis of variants with a zero count in the first round[13]. Scatterplots illustrate the performance of all individual variants in each round of selection (**Fig. 2c**), and allow for the visualization of the behavior of overrepresented variants. A pool of successful variants emerged from the library over 5 rounds of selection, while variants that are unable to infect the retina decreased in their representation. Twenty top variants that increased most over rounds of selection were chosen from the pool of viruses for further analysis. A sequence logo, constructed from the top 100 canine-derived variants, illustrates sequence convergence of successful variants, and the presence of 'DTTK' as a significant motif (**Supplementary Fig. 3**).

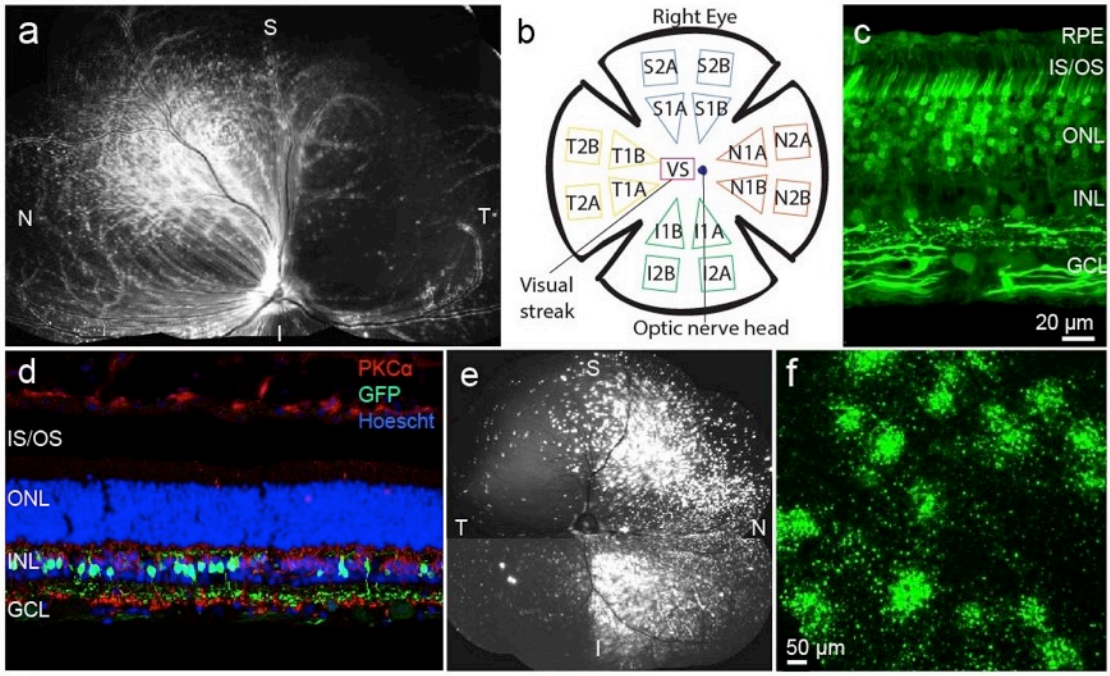


Figure 3. Barcoded and bipolar variants in canine retina. a) A scCAG-eGFP-barcoded library made up of top variants, and co-injected intravitreally into a WT dog eye resulted in robust GFP expression in the retina 3 weeks after injection. **b)** A map illustrates the location of samples collected from the GFP-barcoded library-injected retinas. **c)** Confocal imaging of samples collected from GFP-barcoded library-injected canine retinal samples showed GFP expression in all retinal layers. **d)** Immunohistochemical labeling of retinas injected with GFP-barcoded libraries driven by a bipolar-specific promoter confirmed ON bipolar cell expression in canine retina. **e)** Fluorescent fundus imaging of canine retinas injected with the bipolar specific GFP-barcoded library revealed strong GFP expression in the canine retina. **f)** Confocal imaging of native GFP expression from bipolar specific GFP-barcoded libraries confirmed large numbers of bipolar cells in dog retina expressing GFP.

Then, in order to test each of the selected variants in a quantitative manner with head to head comparison, the vectors were packaged individually with a ubiquitous CAG promoter driving expression of GFP fused to a unique DNA barcode (**Fig. 1 Step 8**). Twenty of the most promising variants were packaged, with AAV2 as a control. Vectors were titer matched, mixed, and injected intravitreally into both eyes of 3 WT dogs and 5 WT mice. Six weeks after injection of the library, GFP expression was bright in fluorescence fundus images of the canine eyes (**Fig. 3a**). Eyes were collected, and tissue samples were collected from across the retina (**Fig. 3b**). The RPE was separated from neural retina, and photoreceptors were collected using transverse sectioning on a cryostat (**Supplementary Fig. 3**). GFP expression, observed by confocal imaging of native GFP expression, was present in every layer of dog retina (**Fig. 3c**). Following DNA extraction, barcodes were PCR amplified from photoreceptors and RPE. Additionally, in order to quantify performance of variants in bipolar cells, a separate library was prepared, in which a bipolar specific promoter[14] was used to drive expression of GFP-barcode constructs (**Fig. 3d**). Library expression was strongest in superior, inferior and nasal retina (**Fig. 3e-f**). Then, mRNA was collected, and used to create cDNA. Illumina sequencing was used to quantify the representation of each of the variants in library and tissue samples. Variants were ranked on the basis of the increase in their representation in the library, compared to the injected AAV library (% of total in AAV library / % of total in recovered library) (**Fig. 4**). Selected variants largely outperformed AAV2 (magenta). Top-performing variants from the outer retina in canine (in orange) did not outperform other variants in mouse retina. In addition, variants that performed best in mouse (in blue) did not perform as well in canine, providing additional indication that variants had species specificity. The convergent variant (LAHQDTTKNA, in green) from the canine screen did not outperform other variants in canine retina, indicating that overall quantity of representation in the final round of selection may not be the best indicator of fitness.

Mouse ONL		Mouse RPE		Canine ONL		Canine RPE		Canine Bipolar (intravitreal)	
LAPDSTTRSA	(1.82)	LAPDSTTRSA	(2.88)	LAKDATKTIA	(3.53)	PAPQDTTKKA	(2.88)	LAPDSTTRSA	(5.36)
LALGETTRPA	(1.70)	LALGETTRPA	(1.62)	PAPQDTTKKA	(1.73)	LAKDATKTIA	(1.62)	LALGETTRPA	(4.72)
LAKSDQSKPA	(1.59)	LAKSDQSKPA	(1.36)	LALGETTRPA	(1.66)	LATTSQNKPA	(1.36)	LQKNARPASTESVNFQ	(2.83)
LAANQPSKPA	(1.42)	LATTSQNKPA	(1.34)	PAHQDTTKNA	(1.57)	LAKSDQSKPA	(1.34)	LAVDGAQRSA	(2.67)
LATTSQNKPA	(1.22)	LAKDATKTIA	(1.32)	LAPDSTTRSA	(1.17)	LAKGTELKPA	(1.32)	LATTSQNKPA	(2.35)
LAVSDSTKAA	(1.13)	SCHEMA_k94	(1.25)	LAKSDQSKPA	(1.13)	LAANQPSKPA	(1.25)	LAKSDQSKPA	(1.40)
LAISDQTKHA	(1.00)	LAVDGAQRSA	(0.96)	LATTSQNKPA	(1.10)	LALGETTRPA	(0.96)	PAPQDTTKKA	(0.48)
LAKGTELKPA	(0.96)	PAPQDTTKKA	(0.79)	LAVDGAQRSA	(1.07)	LPHQDTTKNA	(0.79)	LANQEHVKNA	(0.42)
LAVDGAQRSA	(0.93)	LAANQPSKPA	(0.79)	LAKGTELKPA	(1.04)	LAHQDTTKNA	(0.79)	AAV24YF+TV	(0.09)
LAHQDTTKNA	(0.90)	LAKGTELKPA	(0.75)	LAANQPSKPA	(0.99)	LAVDGAQRSA	(0.75)	TGLDATRDHGLSPVTGT	(0.06)
SCHEMA_k94	(0.88)	LAHQDTTKNA	(0.66)	LSHQDTTKNA	(0.86)	LAVSDSTKAA	(0.66)	NGAVADYTRGLSPATGT	(0.05)
LAKDATKNA	(0.88)	PAHQDTTKNA	(0.65)	LAHQDTTKNA	(0.71)	PAHQDTTKNA	(0.65)	PAHQDTTKNA	(0.04)
PAPQDTTKKA	(0.71)	LAIIDATKNA	(0.63)	LPHQDTTKNA	(0.70)	SCHEMA_k94	(0.63)	LAKDATKTIA	(0.03)
LAIIDATKNA	(0.69)	LAVSDSTKAA	(0.58)	LAVSDSTKAA	(0.68)	LAPDSTTRSA	(0.58)		
LAKDATKTIA	(0.67)	LPHQDTTKNA	(0.52)	LAIIDATKNA	(0.47)	LSHQDTTKNA	(0.52)		
LAKQQSASTA	(0.59)	LSHQDTTKNA	(0.50)	LAISDQTKHA	(0.43)	LAKDATKNA	(0.50)		
LPHQDTTKNA	(0.58)	LAISDQTKHA	(0.45)	SCHEMA_k94	(0.36)	LAISDQTKHA	(0.45)		
AAV2 control	(0.56)	LAKDATKNA	(0.40)	LAKDATKNA	(0.32)	AAV2 control	(0.40)		
PAHQDTTKNA	(0.44)	LAKQQSASTA	(0.39)	LAKQQSASTA	(0.23)	LAKQQSASTA	(0.39)		
LSHQDTTKNA	(0.38)	IARGVAPSSA	(0.35)	AAV2 control	(0.21)	LAIIDATKNA	(0.35)		
IARGVAPSSA	(0.19)	AAV2 control	(0.12)	IARGVAPSSA	(0.07)	IARGVAPSSA	(0.12)		

Figure 4. Ranking of top variants. GFP-barcode library injection results, for mouse and dog ONL and RPE and dog bipolar cells. The lists of variants are ordered from best (top) to worst (bottom) performing vectors, along with a value indicating the extent to which the variant competed with other vectors, expressed as: % of total in AAV library / % of total in recovered library. The best performing dog variants are highlighted in orange. Best performing mouse variants are highlighted in blue. The convergent dog variant is highlighted in green. AAV2 control vectors are highlighted in pink.

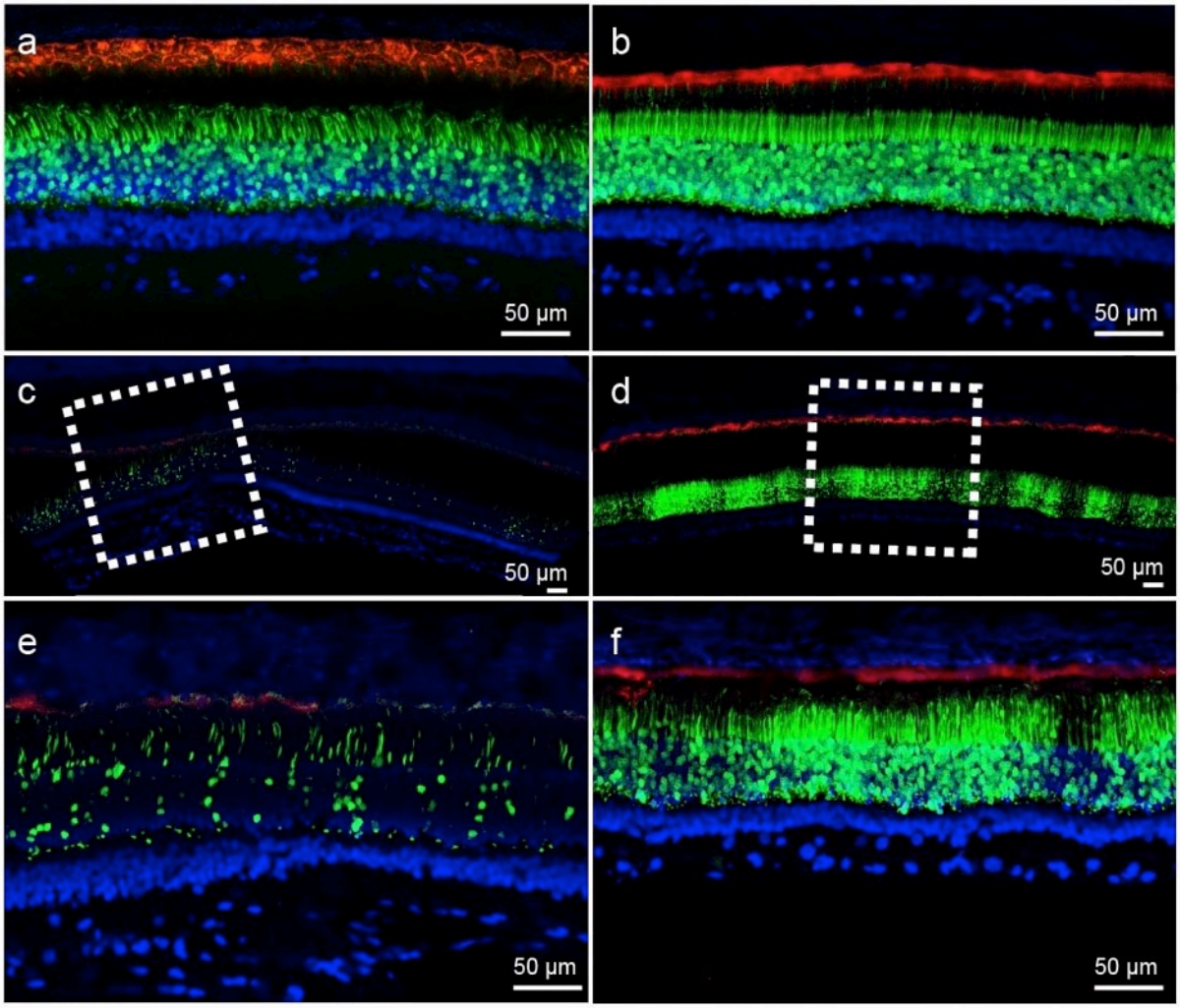


Figure 5. Validation of the best performing variant in canine retina through co-injection of vectors encoding hRho-eGFP and hVMD2-tdTomato. a,b) Subretinal injections of AAV24YF+TV **(a)** and AAV2-PAPQDTTKKA **(b)** showed that both variants infected the retina well from this injection route. **c-f)** Intravitreal injections of AAV24YF+TV **(c,e)** resulted in modest GFP and tdTomato expression in rods and RPE. In contrast, injections of AAV2-PAPQDTTKKA resulted in robust reporter gene expression in outer retina **(d,f)**. **(e,f)** are higher resolution images of the boxed areas in **(c,d)**.

Variant PAPQDTTKKA, which performed well in ONL and RPE, was then packaged with a human rhodopsin promoter driving expression of GFP and a human VMD2 promoter driving expression of tdTomato. Vector packaged with these two constructs was mixed in equal ratios and injected intravitreally into the eye of WT dogs (2.0E+11 vg or 6.25 vg injected per eye). In the contralateral eye, equal titers of the same constructs were packaged into an AAV2-4YF+TV capsid[15]. Identical mixtures were also injected subretinally into 2 eyes. Subretinal injections of both AAV24YF+TV (**Fig. 5a**) and AAV2-PAPQDTTKKA (**Fig. 5b, Supplementary Fig. 4**) resulted in strong fluorophore expression in rods and RPE cells. However, only intravitreal injection of AAV2-PAPQDTTKKA resulted in strong fluorophore expression in the outer retina in canines (**Fig. 5c-f, Supplementary Fig. 4**). A pattern of expression was visible in each of the canine eyes injected in this study, in which expression is greatest in superior, inferior and nasal quadrants of the retina, and less efficient in the visual

streak. Additional studies into the composition of physical barriers, including the vitreous and inner limiting membrane in the area centralis may inform additional studies. Alternatively, additional screening performed only in the visual streak may improve outcomes in this area, as results from the current screen represent an average across the retina.

Primate directed evolution

The species specificity observed in canine screening led us to complete an additional course of directed evolution in primate retina, which has a retinal anatomy like that of the human, and is the animal model most closely related to humans. Importantly, primates have a fovea like that in human retina, which is specialized for high acuity and color vision. Nine libraries were packaged and included in the primate screen: EP2, EP5, EP6, EP8, EP9, EP-Ancestral, AAV2-7mer, Ancestral-7mer[16] and LoopSwap[17]. To increase the safety of the experiment, and to reduce the possibility of AAV replication in primates potentially harboring a co-infection of an AAV helper virus, AAVs were packaged in replication-incompetent libraries containing stop codons in the rep open reading frame (**Supplementary Fig. 6, Appendix B**). Libraries were injected, harvested and repackaged for 5 sequential rounds of selection, with one round of error prone PCR performed after round 3 (**Supplementary Table 3**). AAV cap genes were PCR amplified from ONL, which was isolated from transverse cryosections of retina, and in parallel from overlying RPE. EP libraries were abandoned at round 3, because no variants from these libraries were recovered from retinal tissue. At round 4, additional libraries (AAV4-7mer and AAV5-7mer) were added to the selection, using a separate backbone that was isolated from other libraries by separate PCR binding sites and restriction sites (**Supplementary Fig. 7**).

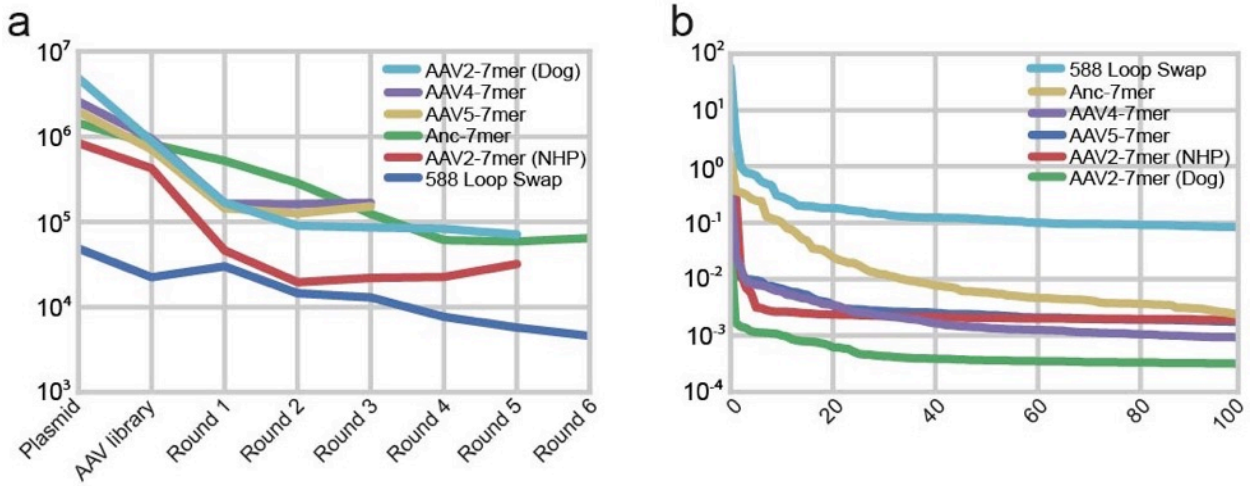


Figure 6. Convergence and overrepresentation of AAV libraries in primate retina. a) Deep sequencing of variant libraries revealed convergence of variants over rounds of selection. **b)** In each of the libraries evaluated, a small proportion of variants are overrepresented in the plasmid library.

Deep sequencing revealed that, similar to observations from the canine screen, libraries contained ~E+6 - ~E+7 individual variants, which converged to ~E+4 - ~E+5 variants over 6 rounds of selection, a diversity not possible to observe through Sanger sequencing (**Fig. 6a**). In each of the libraries analyzed, a small portion of library members were over-represented in the plasmid library (**Fig. 6b**).

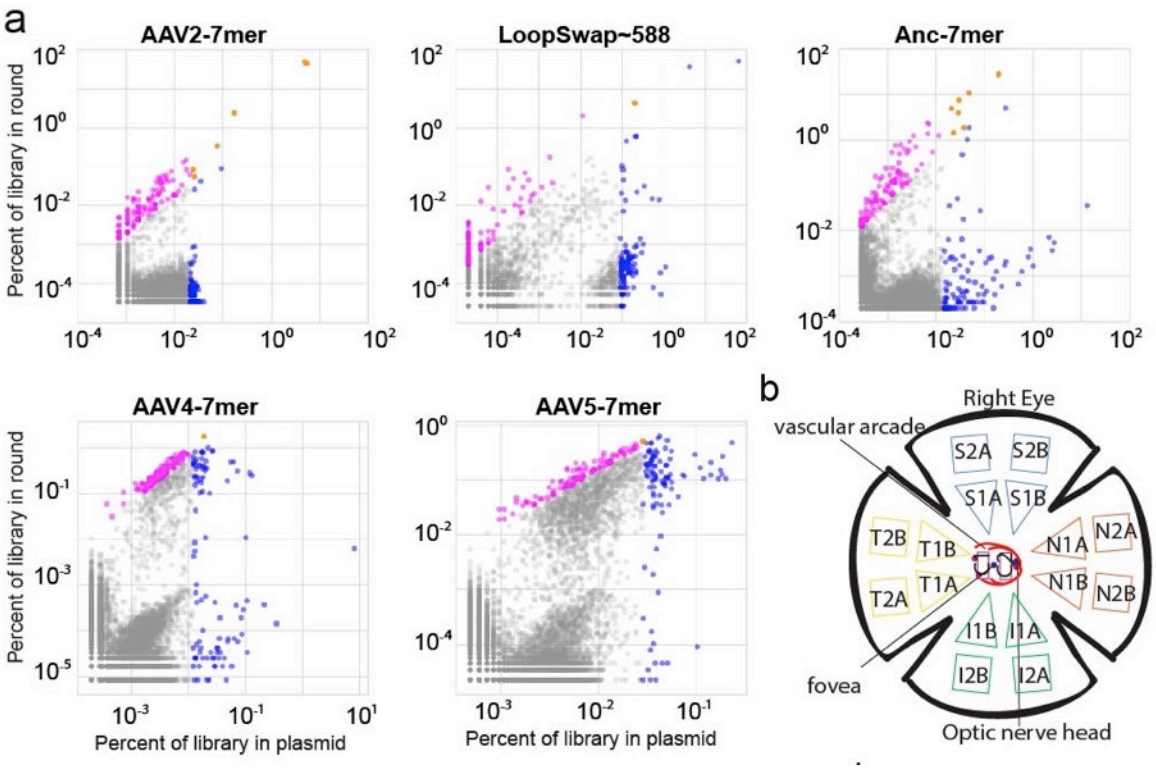


Figure 7. High throughput sequencing of AAV variants and retina biopsy map. a) Scatterplots illustrate the behavior of individual variants at the final round of selection for each of the libraries injected in primate retinas. **b)** A map of the primate retina shows the distribution of samples that were collected for rounds of selection and the GFP-barcode library. Color coding of variants is the same as in **Fig. 2**.

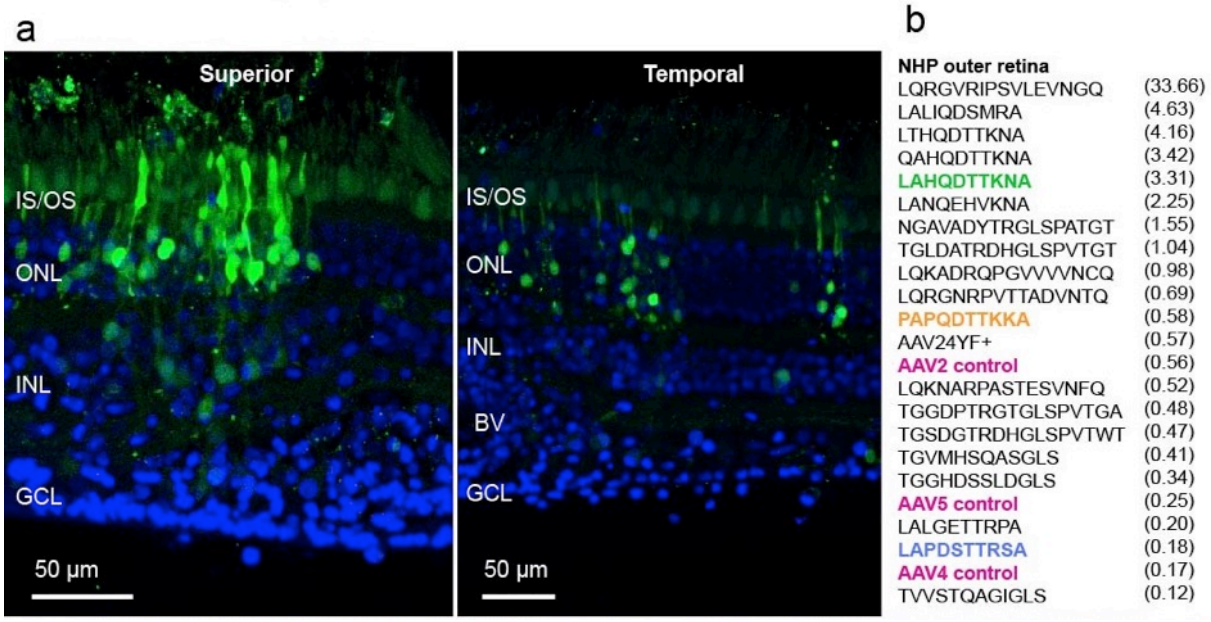


Figure 8. Performance of barcode libraries in primate retina. a) GFP expression resulting from the barcoded library revealed that expression was shifted to an outer retinal tropism in selected variants. **b)** GFP-barcode library injection results, for primate outer retina. The lists of variants are ordered from best (top) to worst (bottom) performing vectors, along with a value indicating the extent to which the variant competed with other vectors, expressed as: % of total in AAV library / % of total in recovered library. The convergent dog variant is colored green. The best performing dog variant is colored orange. AAV2, AAV4 and AAV5 control vectors are colored pink. The best performing mouse variant is colored blue.

Analysis of results from high throughput sequencing, over the rounds of selection, revealed a subset of variants that increased significantly in their representation in each of the libraries (**Fig. 7a, Supplementary Table 4**). Sixteen variants, from the 5 libraries, were selected to be included in a secondary round of selection with GFP-barcoded libraries, along with AAV2, AAV24YF+TV, AAV4 and AAV5 as controls. The library was injected in both eyes of a primate, and 3 weeks after injection, biopsies were collected from locations across the retina (**Fig. 7b**). GFP expression that resulted from injection of the GFP-barcode libraries was primarily found in photoreceptors, as well as some inner retinal cells, a tropism that is shifted from AAV2 or 7m8, which result in stronger inner retinal expression[4] (**Fig. 8a**). Quantification of vector performance in outer retina revealed that AAV2-based variants outperformed viruses based on other serotypes (**Fig. 8b**).

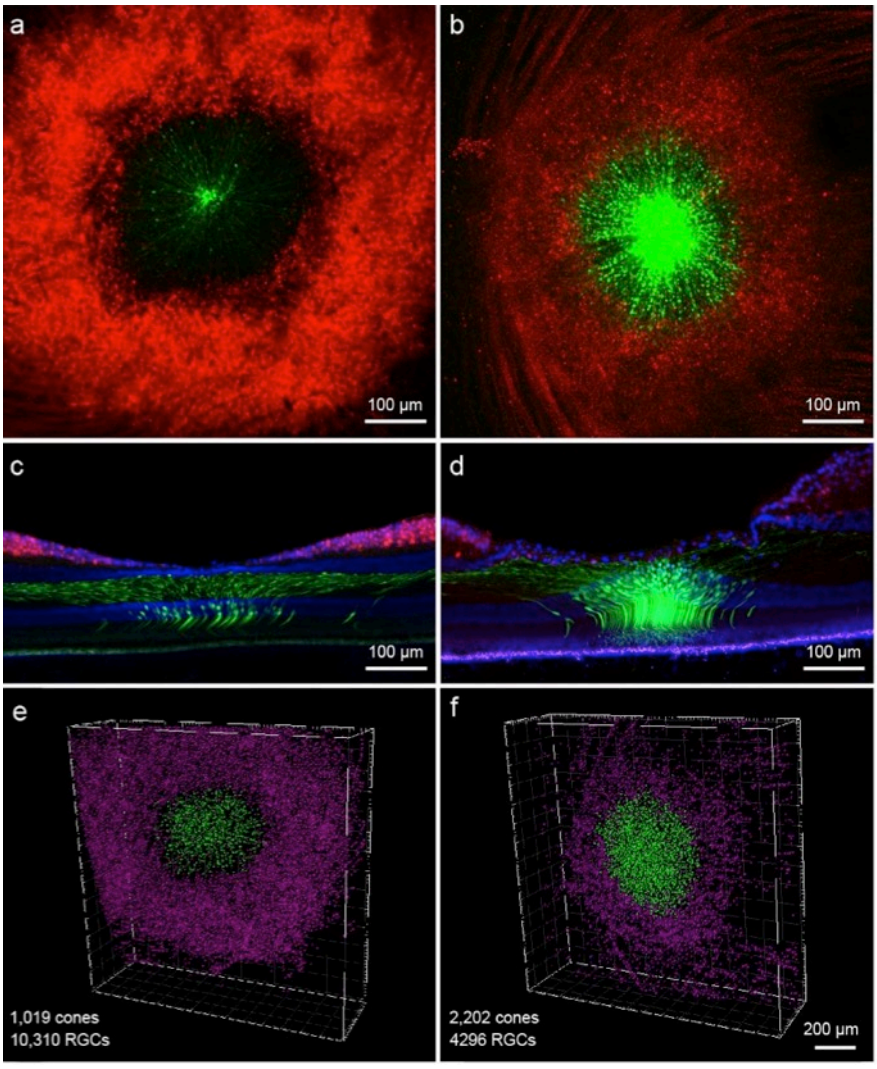


Figure 9. Validation of variant AAV2-LALIQDSMRA in primate retina by co-injection of SCNG-tdTomato and pR2.1-eGFP. Intravitreal injection of 7m8 **(a,c,e)** resulted in robust tdTomato expression in ganglion cells and expression of GFP in foveal cones. In contrast, injection of AAV2-LALIQDSMRA **(b,d,f)** resulted in reduced ganglion cell expression, and increased GFP expression in cones. **e,f)** Quantification of ganglion cells and cones, counted from 3D reconstructions of confocal imaging in foveas, performed in Imaris software, confirmed reduced numbers of ganglion cells and increased numbers of foveal cones.

One vector, Loop Swap variant AAV2 588~LQRGVRIPSVLEVNGQ, outperformed other variants, but generated low viral titers (~5E+11 vg/mL). AAV4-7mer and AAV5-7mer variants underperformed relative to other vectors, which indicates that 2 rounds of selection were not sufficient for successful variants to emerge. AAV2-LALIQDSMRA, the second ranking variant, which packaged at high titers (~5E+13 vg/mL), was therefore selected for validation studies, and was packaged with both an SNCG promoter driving tdTomato in RGCs[18] and the pR2.1 promoter driving expression of GFP in cones. Vectors encoding both these constructs were mixed in equal ratios, and injected intravitreously into the vitreous of a cynomolgous monkey. 7m8, packaged with equal titers of the same constructs was injected into the vitreous of the contralateral eye. Expression of tdTomato reporter in RGC's was reduced in AAV2-LALIQDSMRA-injected eyes compared to 7m8 (**Fig. 9a-d, Supplementary Fig. 9**), while expression in foveal cones was greatly increased. Counting of labeled cells, performed with Imaris software, confirmed a substantial increase in the number of cones targeted with AAV2-LALIQDSMRA (**Fig. 9e,f**).

Together these results indicate that directed evolution, guided by high throughput sequencing, produced AAV viruses with enhanced ability to deliver transgenes to the outer retina in canines and primates, two important preclinical models of retinal degenerative disease. GFP-barcoded libraries represent a sensitive method for evaluating many variants in parallel, in the same animal, allowing for direct, head-to-head comparison and fewer animals used. The species specificity of the vectors created in the context of mouse, canine and primate highlight differences between species and cellular targets, and indicate that it is important to develop AAV vectors with specific targets in mind. Testing of gene therapies in preclinical models may have to take into account the anatomical differences that exist between species. Further validation of the vectors developed in this study need to be conducted to confirm their ability to target additional cell types, and to mediate rescue with gene therapy approaches.

Methods

Construction of the pRepSafeStop directed evolution backbone

The pRepSafeStop plasmid with NotI was created by Quikchange site-directed mutagenesis on pSub2Cap2[14] introducing stop codons in rep at amino acids 5 and 235 using pRepSafeStop SDM primers. Unique pRepSafeStop backbones containing AscI and SpeI sites were created via Gibson Assembly in order to maintain separation of libraries through rounds of selection. Libraries were PCR amplified and digested with HindIII and NotI/AscI/SpeI and ligated into the pRepSafeStop construct.

Rep intron helper cloning

Nebulin intron 8 was amplified from genomic DNA isolated from HEK293T cells using the Neb Genomic primers and cloned into a TOPO vector. To create the pRepIntronHelper, pAAV2/rh10 digested with PmeI and BsmI, Klenow reaction was performed, and the DNA was ligated. The first AGG sequence after the Rep52/Rep40 start codon was used as the site for intron insertion based on a computational analysis of splice signal motifs. Infusion assembly was used to insert the nebulin intron using the IFA primers generating the final plasmid.

Adenovirus rescue experiments

HEK293T cells were infected with AAV at an MOI of 10^5 and then incubated at 37º C and 5% CO_2 for 48 hours. Next, 10 µL adenovirus 5 (Ad5) was infected and the plates were incubated at 37º C and 5% CO_2 for an additional 48 hours. This infection level of Ad5 led to an observable cytopathic effect in the cells that were then harvested, pelleted, resuspended in 100 µL of lysis buffer (0.15M NaCl, 50 mM Tris HCl, 0.05% Tween, pH 8.5). Freeze/thaws were used to lyse the cells and 5 µL of the crude lysate was used for titering.

AAV selection in dog neural retina and RPE

Six weeks after intravitreal injection, dogs were euthanized, and both eyes were flattened by making relief cuts in the globe. 2 mm punches of retina and RPE were collected from superior, inferior, temporal, and nasal regions of the retina, as well as from the area centralis, and flash frozen. DNA was extracted from samples using a Qiagen DNeasy blood and tissue kit, according to the manufacturer's instructions.

AAV selection in primate outer retina

Three weeks after intravitreal injection, the primate was euthanized, and both eyes were briefly submerged in 4% paraformaldehyde. Superior, inferior, temporal, and nasal regions of the retina were cut into four equal pieces, and the RPE was separated from each section. Retinal sections were then immersed in 30% sucrose, embedded in OCT media, and flash frozen. Retinal pieces were sectioned transversely at 20 µm. During sectioning, DAPI staining and light microscopy was used to identify each nuclear layer in the retina, and the inner nuclear and ganglion cell layers were removed. DNA was extracted from samples using a Qiagen DNeasy blood and tissue kit, according to manufacturer's instructions.

AAV packaging

AAV libraries were constructed as previously described[7,8,15,16]. After each round of injection, capsid sequences were recovered by PCR from harvested cells using primers HindIII_F1 and NotI_R1, AscI_R1, or SpeI_R1 with reverse primers being specific to unique AAV backbones, in order to maintain separation of groups of libraries. PCR amplicons were then digested, and recloned into the AAV pRepSafeStop backbone. AAV packaging has been described previously[15,17]. AAV vectors with pRepSafeStop backbone were produced by triple transient transfection of HEK293T cells with the addition of the pRepIntronHelper plasmid in 5 times greater concentration than the library plasmid, purified via iodixanol density centrifugation, and buffer exchanged into PBS by Amicon filtration. DNase-resistant viral genomic titers were measured by quantitative real time PCR using a Biorad iCycler.

Deep sequencing of AAV libraries from rounds of selection

A ~75-85 base pair region containing the 7mer insertion or loop swap mutation site was PCR amplified from harvested DNA. Primers included Illumina adapter sequences containing

unique barcodes to allow for multiplexing of amplicons from multiple rounds of selection (**Supplementary Table 5**). PCR amplicons were purified and sequenced with a 100-cycle single-read run on an Illumina HiSeq 2500. Custom Python code was written to translate DNA sequences into amino acid sequences, and to identify and count reads containing unique 7mer insert sequences. Read counts were normalized by the total number of reads in the run. Python and Pandas were used to analyze dynamics of directed evolution and create plots.

Deep sequencing analysis

Best performing variants were selected by greatest fold increase in the final round of selection relative to the initial plasmid library (# reads in final round, normalized to total number of reads in the round / # of reads in library, normalized to total number of reads in the round). A pseudo-count of 1 was added before normalization to each individual variant to allow analysis of variants not appearing in sequencing of the plasmid library[13].

GFP barcode library construction

Unique 25 bp DNA barcodes were cloned behind an AAV ITR construct containing a self-complementary CAG promoter driving eGFP (CAG-GFP-Barcode-pA). Individual variants were packaged separately with constructs containing different barcodes. Variants were then titer matched and mixed in equal ratios before injection into mice, dogs and primates.

Deep sequencing of GFP-barcode libraries

Barcodes were PCR amplified directly from DNA or cDNA (created from mRNA using Superscript III reverse transcriptase), which was harvested from dog or primate retinal tissue. Samples were collected from areas across the retina, and from ONL or RPE. Primers amplified a ~50 bp region surrounding the GFP barcode and contained Illumina adapter sequences and secondary barcodes to allow for multiplexing of multiple samples (**Supplementary Table 5**). PCR amplicons were purified and sequenced with a 100-cycle single-read run on a MiSeq. Read counts were normalized by total number of reads in the run. Analysis of barcode abundance was performed using custom code written in Python, followed by creation of plots in Pandas. Best performing variants were selected based on the fold increase in the percent of total library, relative to the injected library (% of total in recovered sample / % of total in injected library). Analysis was performed on n=5 mice, n=3 dogs and n=1 primate.

Primers

Primer sequences are listed in **Supplementary Table 5**.

Animal studies

Mice: C57BL/6 mice from Jackson Laboratories were used for mouse experiments. All procedures were performed in accordance with the ARVO statement for the Use of Animals in Ophthalmic and Vision Research, and were approved by the University of California Animal Care and Use Committee AUP# R200-0913BC. Surgery was performed under anesthesia, and all

efforts were made to minimize suffering.

Dogs: Dogs, between the age of 7-17 months, except for bipolar-GFP-barcode injected animals, which were 6-7 years old at time of injection, were screened as previously described[19] for neutralizing antibodies to AAV2, AAV5 and AAV9. All dogs had titers <1:25. Intraocular injections of Triamcinolone Acetonide (0.1 ml of KENALOG 40) were performed at the time of injection. Animals were treated post-injection with topical application of prednisolone acetate and oral steroids. The animals were bred and/or maintained at the Retinal Disease Studies Facility (RDSF, New Bolton Center, University of Pennsylvania, Kennett Square, PA, USA), and all procedures were in compliance with the ARVO statement for the Use of Animals in Ophthalmic and Vision Research and with Institutional Animal Care and Use Committee approval. A summary of adverse events related to the procedures is summarized in **Supplementary Table 1**. AAV injections and fundus imaging were performed as previously described[20].

Primates: Cynomolgous monkeys between 4-10 years old were used for all studies. Intravitreal injections were made with methods described previously[21]. Monkeys received daily subcutaneous injections of cyclosporine at a dose of 6 mg/kg, and adjusted based on blood trough levels to within a 150-200 ng/ml target range. Confocal scanning laser ophthalmoscopic images (Spectralis HRA, Heidelberg Engineering) were obtained from the two retinas at 3 weeks after injection, with autofluorescence settings, leading to effective tdTomato and GFP visualization. For histology, the monkey was euthanized, both retinas were lightly fixed in 4% paraformaldehyde, and tissue was examined via confocal microscopy. At the conclusion of the experiment, euthanasia was achieved by administering an IV overdose of sodium pentobarbital (75 mg kg−1), as recommended by the Panel on Euthanasia of the American Veterinary Medical Association. Pieces of primate retina were then prepared in 30% sucrose, embedded in OCT media, flash frozen and sectioned at 20 μm for confocal microscopy imaging of native fluorophore expression. A summary of adverse events related to the procedures is summarized in **Supplementary Table 3**. The procedures were conducted according to the ARVO Statement for the Use of Animals and the guidelines of the Office of Laboratory Animal Care at the University of Rochester.

Acknowledgements

Leah Byrne led the canine studies with collaborators William Beltran, Gustavo Aguirre, Valerie Dufour and Luis Lima Pompeo Marinho at the University of Pennsylvania. Leah Byrne and Timothy Day led the primate studies with Deniz Dalkara and Meike Visel at UC Berkeley. Deep sequencing was performed at the UC Berkeley Vincent J. Coates sequencing facility. Confocal imaging was performed at the Berkeley Biological Imaging Facility. Injections for AAV selection in primates were performed at Valley Biosciences. William Merigan performed injection and fundus imaging of primate variants at the University of Rochester. We thank Yvonne Lin, Cécile Fortuny, and Jaskiran Mann for technical assistance. We thank Lydia Melnick for research coordination.

References

1. RetNet - Retinal Information Network. *https://sph.uth.edu/retnet/sum-dis.htm* Available at: (Accessed: 1st December 2017)
2. Kotterman, M. A. & Schaffer, D. V. Engineering adeno-associated viruses for clinical gene therapy. *Nat. Rev. Genet.* **15,** 445–451 (2014).
3. Klimczak, R. R., Koerber, J. T., Dalkara, D., Flannery, J. G. & Schaffer, D. V. A novel adeno-associated viral variant for efficient and selective intravitreal transduction of rat Müller cells. *PLoS ONE* **4,** e7467 (2009).
4. Dalkara, D. *et al.* In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous. *Science Translational Medicine* **5,** 189ra76 (2013).
5. Matsumoto, B., Blanks, J. C. & Ryan, S. J. Topographic variations in the rabbit and primate internal limiting membrane. *Investigative Ophthalmology & Visual Science* **25,** 71–82 (1984).
6. Heegaard, S. S., Jensen, O. A. O. & Prause, J. U. J. Structure and composition of the inner limiting membrane of the retina. SEM on frozen resin-cracked and enzyme-digested retinas of Macaca mulatta. *Graefes Arch. Clin. Exp. Ophthalmol.* **224,** 355–360 (1986).
7. Acland, G. M. *et al.* Gene therapy restores vision in a canine model of childhood blindness. *Nat Genet* **28,** 92–95 (2001).
8. Beltran, W. A., Cideciyan, A. V. & Lewin, A. S. Gene therapy rescues photoreceptor blindness in dogs and paves the way for treating human X-linked retinitis pigmentosa. in (2012). doi:10.1073/pnas.1118847109/-/DCSupplemental
9. Koerber, J. T., Maheshri, N., Kaspar, B. K. & Schaffer, D. V. Construction of diverse adeno-associated viral libraries for directed evolution of enhanced gene delivery vehicles. *Nat Protoc* **1,** 701–706 (2006).
10. Müller, O. J. *et al.* Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors. *Nat Biotechnol* **21,** 1040–1046 (2003).
11. Koerber, J. T., Jang, J.-H. & Schaffer, D. V. DNA Shuffling of Adeno-associated Virus Yields Functionally Diverse Viral Progeny. *Mol Ther* **16,** 1703–1709 (2008).
12. 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). doi:10.1016/j.ymthe.2017.09.006
13. Fowler, D. M., Stephany, J. J. & Fields, S. Measuring the activity of protein variants on a large scale using deep mutational scanning. *Nat Protoc* **9,** 2267–2284 (2014).
14. Lu, Q. *et al.* AAV-mediated transduction and targeting of retinal bipolar cells with improved mGluR6 promoters in rodents and primates. *Gene Ther* **23,** 680–689 (2016).
15. 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* (2013). doi:10.1038/gt.2013.64
16. Santiago-Ortiz, J. *et al.* AAV ancestral reconstruction library enables selection of broadly infectious viral variants. *Gene Ther* **22,** 934–946 (2015).
17. Koerber, J. T. *et al.* Molecular evolution of adeno-associated virus for enhanced glial gene delivery. *Mol Ther* **17,** 2088–2095 (2009).
18. Chaffiol, A. *et al.* A New Promoter Allows Optogenetic Vision Restoration with Enhanced Sensitivity in Macaque Retina. *Mol Ther* **25,** 2546–2560 (2017).

19. Kotterman, M. A. *et al.* Antibody neutralization poses a barrier to intravitreal adeno-associated viral vector gene delivery to non-human primates. *Gene Ther* **22,** 116–126 (2015).
20. Beltran, W. A. *et al.* rAAV2/5 gene-targeting to rods:dose-dependent efficiency and complications associated with different promoters. *Gene Ther* **17,** 1162–1174 (2010).
21. Yin, L. *et al.* Intravitreal injection of AAV2 transduces macaque inner retina. *Investigative Ophthalmology & Visual Science* **52,** 2775–2783 (2011).

Chapter 3: Screening for neutralizing antibodies against natural and engineered AAV capsids in non-human primate retinas

Timothy P. Day, Leah C. Byrne, John G. Flannery, David V. Schaffer

Submitted as a chapter to Methods in Molecular Biology

Abstract

Adeno-associated virus (AAV) has shown promise as a therapeutic gene delivery vector for inherited retinal degenerations in both preclinical disease models and human clinical trials. The retinas of nonhuman primates (NHPs) share many anatomical similarities to humans and are an important model for evaluating AAV gene delivery. Recent evidence has shown that pre-existing immunity in the form of neutralizing antibodies (NABs) in NHPs strongly correlates with weak or lack of AAV transduction in the retina when administered intravitreally, work with translational implications. This necessitates prescreening of NHPs before intravitreal delivery of AAV. In this chapter, we describe a method for screening NHP serum for pre-existing NABs.

Introduction

AAV-mediated gene therapy has enjoyed recent success ranging from multiple proof-of-concept experiments to human clinical trials [1-3]. Animal models are currently the best way to optimize application of AAV vectors, as they recapitulate many of the cellular barriers and tissue structures encountered in the final therapeutic application, in a manner unattainable *in vitro.* Small animal models, such as rodents, remain important to the development of human gene therapy; however, for several reasons including important structural and functional differences, they cannot predict transduction efficiency in humans. Large animals – such as dogs, pigs or non-human primates (NHPs) – more effectively evaluate the clinical potential of AAV for treatment of retinal degeneration patients and provide key insights. Specifically, the eye of the NHP has nearly identical anatomy compared to humans, with a cone-rich macula and cone-only fovea. The biological and immunological features of NHPs are also closer to humans, and the surgical procedures for the eye are similar. In the NHP, AAV can be delivered to the retina via subretinal or intravitreal injection. Subretinal injection involves creating a fluid filled space or 'bleb' between the apical side of the RPE and the photoreceptors, leading to a concentrated volume of virus in a small space, and resulting in a restricted area of transduction. Consequently, this localized injection does not treat the entire retina, and there are some associated risks associated with making the retinal detachment. By comparison, intravitreal injection is a less invasive procedure, as it does not require a through-retinal needle penetration (retinotomy) with the retinal detachment, and allows for the potential of panretinal expression [4]. That said, intravitreal injections create greater exposure of the viral capsid proteins to the immune system as they more readily exit the eye through the trabecular meshwork outflow pathway into the circulation. Vectors introduced into the vitreous must cross through the inner limiting membrane and multiple cell layers to reach either the photoreceptors or RPE [4, 5].

The tropism of AAV in the NHP retina depends upon several factors including the serotype, route of administration, promoter, and dose. Therefore, it is critical to choose the correct vector for the desired outcome. Subretinal injection of AAV2 in NHPs leads to photoreceptor and RPE transduction for at least 1 year [6, 7]. The ability to target foveal photoreceptors has been reported to be dose-dependent, with lower doses leading only to RPE transduction and higher doses leading to expression in both RPE and photoreceptors. The limited photoreceptor expression observed outside the fovea was determined to be predominantly rods [5, 8]. Subretinal injection of AAV2 has most notably been used in the treatment of Leber's congenital amaurosis (LCA2) in human clinical trails. Intravitreal injection of AAV2 (driving expression from a ubiquitous chicken beta actin, i.e. CBA, or CMV promoter) in NHPs leads to a disc of expression centered on the fovea with only Müller glia and especially retinal ganglion cells (RGCs) transduced. Variable expression in the periphery in Müller glia and RGCs was observed. In addition, RGC specific expression was achieved using the hCx36 promoter [9]. RGC expression has been used to restore light response to the NHP retina with channelrhodopsin expressed in RGCs [10, 11].

In addition to AAV2 vectors, subretinal delivery of AAV5 leads to transduction of both the photoreceptor cells and the RPE, with greater transduction efficiency observed in the photoreceptors [12]. Photoreceptor specificity can be achieved with hGRK1 promoter, which has been shown to have strong expression with no evidence of gross pathology after injection [13]. Most notably, subretinal delivery of an AAV5 vector was used to produce trichomatic vision in a primate model of red-green color blindness [14]. Subretinal delivery of AAV7 results in modest levels of photoreceptor transduction. AAV7 is outperformed by AAV8 for rod transduction and outperformed by AAV9 for cone transduction in the NHP retina [4]. Subretinal delivery of AAV8 leads to expression in RPE, photoreceptors, and Muller glia with the relative efficiency of transduction in the 3 classes of cells dependent upon the dose. Rod photoreceptors were transduced with higher efficiency than cones and foveal cones were transduced at a higher efficiency than extrafoveal cones in NHPs [8]. Subretinal delivery of AAV9 led to strong expression in cones both centrally and peripherally in NHP retinas with limited rod expression observed in the periphery [4]. This transduction pattern is thought to be due to the level of galactose found on the surface of cones, the receptor for AAV9 [15].

Despite the promise of subretinal delivery in clinical studies to date, it would be preferable to utilize intravitreal administration to avoid the risk associated with subretinal surgery and to potentially transduce the full surface area of the retina. However, natural evolution likely did not select AAV for its capacity to infect the retina, and as a result naturally occurring AAV variants isolated from healthy, extraocular tissues, would be expected to yield vectors with significant limitations in gene transfer to either healthy or diseased retina. To overcome this obstacle, viruses can be engineered for advantageous delivery properties for human ocular tissue. In particular, successful intravitreal injection will require the engineering of AAV variants to penetrate the physical barrier of the inner limiting membrane [5]. In general, two methodologies have been approached to engineer better AAV vectors for retinal gene therapy: rational design and directed evolution.

The rational design approach applies fundamental molecular biology knowledge to the improvement of the viral capsid. Zhong *et al*. showed that epidermal growth factor receptor protein tyrosine kinase-mediated phosphorylation of tyrosine residues on the AAV leads to

degradation of the virus via ubiquitination [16]. This insight resulted in the design of tyrosine to phenylalanine AAV mutants, which increase transduction efficiency by attenuating proteasome degradation [17, 18]. The utility of the tyrosine mutant vector was recently observed in a NHP model. For example, when a GFP transgene with a modified mGluR6 promoter was packaged in an AAV2 based tyrosine mutant, some transduction was observed in difficult-to-reach bipolar cells in NHP retinas after intravitreal injection [10]. In addition, low level transduction in a small number of foveal cones has been reported with tyrosine mutant AAV2 vector [17-19].
While rational design has led to improvements in AAV retinal transduction, in general rational engineering relies upon detailed mechanistic knowledge of the full gene delivery pathway – from point of administration to arrival in a target nucleus – to enable capsid modifications to overcome critical steps that limit transduction. Unfortunately, the requisite breadth and depth of knowledge is not available for most tissue and cellular targets in a primate system. As a result, a vector engineering strategy that does not rely on *a priori* mechanistic knowledge but can instead still lead to advantageous characteristics is required, and fortunately directed evolution is such an approach. Directed evolution is a vector engineering strategy in which large genetic libraries of AAV variants are generated through a variety of methods and selected via a selection pressure, such as successful retinal gene delivery. In this process, multiple iterations of selection lead to a convergence on novel AAV variants that have the selected properties [20]. For example, directed evolution in the mouse retina led to 7m8, a variant of AAV2 with an additional peptide sequence, that is capable of panretinal expression in the mouse when injected from the vitreous as well as improved expression in the NHP retina including increased expression across the retina and inside the fovea. Multiple studies have shown that 7m8 has increased transduction efficiency compared to tyrosine mutants [10, 21]. Additional engineering in larger animal models offers the promise of further improvement.

An alternate engineering approach involved the generation of hybrid AAV vectors that combined fragments of novel AAV capsid sequences isolated from primates mixed with AAV8. These novel hybrid recombinant vectors were screened, and two were shown to transduce ganglion cells in *ex vivo* macaque retinal explants. The majority of the transduced ganglion cells were found at the edges of the explant, with limited expression in the center [22]. This supports the idea that the inner limiting membrane acts as a barrier to AAV and that on the edges it was disrupted by the dissection [5].

These examples illustrate the importance of NHP models for evaluating AAV gene delivery in the retina. Despite the immune privilege of tissue in the eye both pre-existing and development of immunity post-injection are obstacles to AAV gene delivery.
Dose escalation studies of subretinally delivered AAV showed a dose-related increase in neutralizing antibodies in serum but not in the anterior chamber until high doses (10^{11} viral genomes) were delivered [8]. A biodistribution study of intravitreally delivered AAV2 showed high levels of vector DNA in the injected eye but low levels of vector DNA in the spleen and lymph nodes, as well as other organs [19]. After subretinal vector delivery, DNA has been found in lacrymal and nasal fluids for 3-4 days and in serum for up to 15-20 days [23]. This has been associated with a retinal hemorrhage during subretinal surgery [13].

Evidence in the literature demonstrates that the immune response to the viral capsid following subretinal injection does not impede the readministration of AAV subretinally to the

contralateral eye in both NHP studies and human clinical trials [24-26]. Intravitreal injection, on the other hand, led to NABs against the capsid and a humoral immune response that prevented transgene expression when the viral vector was subsequently delivered intravitreally (but not when administered subretinally) to the contralateral eye. Subretinal injection did not block transgene expression upon readministration in the contralateral eye with either subretinal or intravitreal injection [27]. A study of pre-existing immunity to AAV showed that there is both cross-reactivity of NABs against different AAV serotypes and that pre-existing NABs against AAV in primate serum samples correlate with weakened or no gene expression following intravitreal injection of AAV in the NHP retina [28]. Therefore, prescreening primates for NABs against AAV is specifically necessary for proper evaluation of intravitreally administered AAV vectors in the retina, both to characterize NABs before and after a subretinal surgery and especially for intravitreal injection. The following method describes a screening protocol for NABs against AAV in NHP serum samples.

The neutralizing antibody assay allows for the detection of NABs in the serum against AAV serotypes that could impede the expression of transgenes delivered by AAV via intravitreal injection. In this assay, serum from multiple NHPs is collected and then serially diluted. An AAV virus carrying a transgene for a fluorescent reporter is incubated with the serially diluted serum. This incubated AAV is then used to infect a cell line in a 96-well plate. The plate is then imaged for GFP expression. If expression of the fluorescent reporter is impeded by highly diluted serum, then it can be determined that there is a high level of NABs present in that animal. Conversely, if GFP expression is not reduced in high levels of serum then the animal has a low level of NABs present and is a suitable candidate for intravitreal injection.

Materials and Methods

Serum Processing

1. Primate serum in red top vacutainer
2. Collection tubes
3. Ice
4. Pipet
5. Cryovials
6. Benchtop centrifuge with swing-out rotor and carriers

Quantification of NAB titers

1. Greiner CELLSTAR 96-well polystyrene black plate with clear bottom
2. Greiner CELLSTAR 96-well plate polystyrene plate – clear
3. HEK 293 Cells ATCC #CRL-1573
4. Packaged AAV with a fluorescence reporter driven by a ubiquitous promoter
5. 1.5 mL Eppendorf tubes
6. Multichannel pipet
7. Disposable pipetting reservoir
8. DMEM media
9. Fetal bovine serum
10. Dulbecco’s phosphate-buffered saline (D-PBS)

11. Hoechst nuclear stain (2′ -(4-hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5′ -bi-1H-benzimidazole trihydrochloride hydrate; Sigma)
12. ImageXpress Micro Cellular Imaging and Analysis System

Methods

Serum Processing

1. A trained professional should perform blood collection with all proper safety measures in place, including an approved animal use protocol, necessary equipment for the animal's safety, and appropriate personal protective equipment. The blood sample is taken from a peripheral vein with blood collected in a red top Vacutainer.
2. Once blood is collected, place tubes upright in a rack and incubate at room temperature for 30 min. Do not exceed one hour.
3. Centrifuge the blood samples for 15 min at 1000 x *g* and do not use brakes to stop the centrifuge.
4. Use a pipette to aspirate off the serum and pool in a collection tube. Keep this collection tube on ice (Note 1).
5. If not using serum samples immediately, aliquot the serum into cryovials then freeze and store at -80 °C.

In vitro quantification of NAB titers

1. Before beginning the procedure, ensure that you have sufficient AAV with the capsid of interest that is packaged with a fluorescent reporter (usually eGFP) driven by a ubiquitous promoter and that has already been titered (Note 2).
2. 24 hours before the serum incubation, pass 293 cells into a 96-well black plate with clear bottom at a density of $1.5x10^4$ cells/well in 100µl DMEM + 10% FBS (Note 3). Incubate the cells for 24 hours at 37º C with 5% CO_2. The following day, cells should be approximately 70-80% confluent.
3. On the day of serum incubation, heat inactivate FBS by heating at 56º C for 30 minutes. Then mix DMEM + 1% heat inactivated FBS media.
4. Dilute serum samples in DMEM + 1% heat inactivated FBS using the following dilutions of primate serum: 1:2, 1:10, 1:25, 1:50, 1:100, 1:250, 1:500, 1:1000, 1:2500, and 1:5000. Make the dilutions in 1.5 mL Eppendorf tubes (See Notes 4-5). Each dilution should be tested in triplicate for each primate sample using 60 µL per well. Therefore, prepare at least 200 µL of each dilution (180 µL to be used and 20 µL to allow for pipetting error).
5. Transfer each dilution from the Eppendorf tube into the three replicate wells with 60 µl of dilution per well to a sterile 96-well plate (clear).
6. Include controls on the plate. Use 3 wells with virus but without serum for a positive GFP control and 3 wells without virus for a negative GFP control.
7. Make a stock solution of virus that will be incubated with serum. Dilute virus to ~2,000 MOI in DMEM + 1% heat inactivated FBS (Notes 6-7). Make 60 µL per dilution for each well and include overages for pipetting error. For example, testing one primate sample would require 1.8 mL of diluted virus (10 serum dilutions * 3 replicates of each dilution * 60 µL of diluted virus).
8. Add 60 µL of the virus stock solution to each well of the incubation plate. Each well will now have a total volume of 120 µL.
9. Mix by gently pipetting up and down.

10. Incubate at 37º C in 5% CO_2 for one hour.
11. Add 50 μL from each serum/virus mixture to each well of the plated cells in the first 96-well plate that was seeded the day before.
12. Incubate for 48 hours at 37ºC with 5% CO_2.
13. Dilute Hoechst nuclear stain 1:4000 in D-PBS
14. Remove media from each well of 96-well plate and add 100 μL of diluted Hoechst stain to each well
15. Image the plate using a high content fluorescence imager such as the ImageXpress Micro Cellular Imaging and Analysis System. This will allow for automated reading of the 96-well plate. The accompanying MetaXpress Image Analysis Software (or alternate software such as CellProfiler) can be used to determine the number of transduced cells in each well (Notes 8-9).
16. Average the results from the three replicates for each dilution. Validate that controls have high expression level, and determine the dilution at which less than 50% of cells have fluorescent expression. This is reported as the neutralizing antibody titer. See **Fig. 1** for an example plate image.

Specifically for intravitreally administered AAV vectors, there is a strong effect of pre-existing NABs on transgene expression. Neutralization of 50% of GFP expression at dilutions of 1:10 to 1:50 can negatively effect transgene expression in NHPs, with reports of neutralization at a dilution of 1:10 resulting in zero transgene expression. See Table 1 of reference [28]. It is ideal to only utilize NHPs with no detectable NABs present. In addition, NAB analysis is important for characterizing animals receiving subretinal injection, including seroconversion following administration. In summary, this protocol is useful for characterizing the potential for animal neutralization of administered AAV vector, both for intraocular injection and in general.

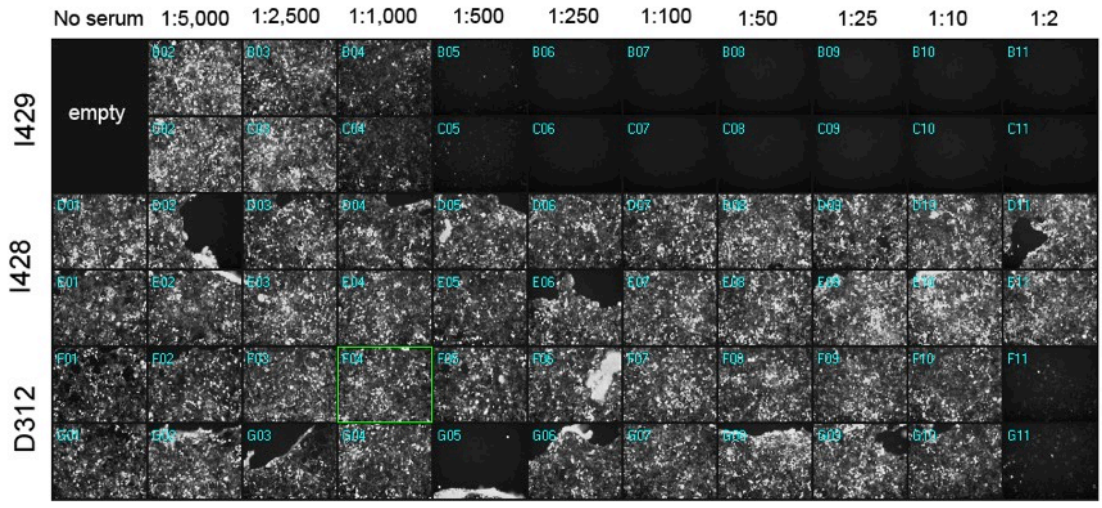


Figure 1. Example of imaged 96-well plate. Sera from three different animals were serially diluted and infected with AAV2-CMV-GFP. In this example, animal I428 is the best candidate for intravitreal injection of AAV with no loss of GFP expression at the highest serum concentration. The empty boxes indicate a negative control. The green box represents an ideal well for effective imaging.

Notes

1. When pipetting serum, be careful not to disturb the cell layer or aspirate any cells. After collecting the serum, hold the tube to the light and look for turbidity. If turbid, repeat the centrifugation and pipet serum to a new tube.
2. For an in-depth guide to AAV vector packaging and purification please refer to this reference [29].
3. The cell count, correct confluency, and consistency in each well are very important. Use a hemocytometer or other method to correctly count the number of cells. Using a multichannel pipet and a disposable pipetting reservoir will also increase accuracy and consistency.
4. Screen multiple NHPs at one time. Each primate sample requires 30 wells to screen for one AAV serotype or variant, so 3 primate samples can be screened in one plate, including controls. To ensure that a sufficient number of animals are seronegative against AAV, a rough guideline is that 3 NHPs should be screened to identify one for utilization in the AAV experiment. For example, if an experiment requires 3 animals then screen at least 9. However, this guideline may vary depending on the animal facility.
5. When transferring the serial dilution, the same pipet tip can be used throughout if pipetting is started with the most dilute sample to the least dilute sample. The carryover from the lower dilution will be negligible, and not changing tips will speed up the process.
6. Virus titer is very important to determine the correct MOI. In addition, the transduction efficiency can differ between serotypes, with different promoters, or between different cell types. If doing the experiment for the first time, consider testing the transduction efficiency of the AAV to be used at multiple MOIs to determine what would be optimal for that particular case.
7. MOI of 2,000 is determined as follows: Wells are seeded with 1.5 $x10^4$ cells. Assuming one cell division cycle there are roughly 3 $x10^4$ cells/well after 24 hours. Therefore, 6 $x10^7$ viral genomes in 60 μL of DMEM + 1% heat inactivated FBS are needed.
8. If NABs are present at a sufficiently high concentration, then AAV transduction will be inhibited, and there will be a reduction in GFP expression. Determine the lowest serum dilution that leads to less than 50% GFP expression in the well compared to positive control.
9. Cell imaging is one of many methods, and alternatives include flow cytometry or a luciferase assay, with the appropriate transgene payload, to determine the serum dilution which inhibits AAV transduction [13, 30].

Acknowledgments

We thank the members of both the Flannery and Schaffer labs including M.A. Kotterman for her work on the development of the serum screen. This work was funded by NIH 5R01EY022975.

References

1. A. M. Maguire, F. Simonelli, E. A. Pierce, E. N. Pugh Jr., F. Mingozzi, J. Bennicelli, S. Banfi, K. A. Marshall, F. Testa, E. M. Surace, S. Rossi, A. Lyubarsky, V. R. Arruda, B. Konkle, E. Stone, J. Sun, J. Jacobs, L. Dell'Osso, R. Hertle, J.-X. Ma, T. M. Redmond, X. Zhu, B. Hauck, O. Zelenaia, K. S. Shindler, M. G. Maguire, J. F. Wright, N. J. Volpe, J. W. McDonnell, A. Auricchio, K. A. High, and J. Bennett, Safety and Efficacy of Gene Transfer for Leber's Congenital Amaurosis, *N Engl J Med* 358, 2240 (**2008**).
2. J. W. B. Bainbridge, A. J. Smith, S. S. Barker, S. Robbie, R. Henderson, K. Balaggan, A. Viswanathan, G. E. Holder, A. Stockman, N. Tyler, S. Petersen-Jones, S. S. Bhattacharya, A. J. Thrasher, F. W. Fitzke, B. J. Carter, G. S. Rubin, A. T. Moore, and R. R. Ali, Effect of Gene Therapy on Visual Function in Leber's Congenital Amaurosis, *N Engl J Med* 358, 2231 (**2008**).
3. F. Simonelli, A. M. Maguire, F. Testa, E. A. Pierce, F. Mingozzi, J. L. Bennicelli, S. Rossi, K. Marshall, S. Banfi, E. M. Surace, J. Sun, T. M. Redmond, X. Zhu, K. S. Shindler, G.-S. Ying, C. Ziviello, C. Acerra, J. F. Wright, J. W. McDonnell, K. A. High, J. Bennett, and A. Auricchio, Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration., *Molecular Therapy* 18, 643 (**2009**).
4. L. H. Vandenberghe, P. Bell, A. M. Maguire, R. Xiao, T. B. Hopkins, R. Grant, J. Bennett, and J. M. Wilson, AAV9 Targets Cone Photoreceptors in the Nonhuman Primate Retina, *PLoS ONE* 8, e53463 (**2013**).
5. D. Dalkara, K. D. Kolstad, N. Caporale, M. Visel, R. R. Klimczak, D. V. Schaffer, and J. G. Flannery, Inner Limiting Membrane Barriersto AAV-mediated Retinal TransductionFrom the Vitreous, *Molecular Therapy* 17, 2096 (**2009**).
6. J. Bennett, A. M. Maguire, A. V. Cideciyan, M. Schnell, E. Glover, V. Anand, T. S. Aleman, N. Chirmule, A. R. Gupta, Y. Huang, G. P. Gao, W. C. Nyberg, J. Tazelaar, J. Hughes, J. M. Wilson, and S. G. Jacobson, Stable transgene expression in rod photoreceptors after recombinant adeno-associated virus-mediated gene transfer to monkey retina., *Proc. Natl. Acad. Sci. U.S.A.* 96, 9920 (**1999**).
7. S. G. Jacobson, S. L. Boye, T. S. Aleman, T. J. Conlon, C. J. Zeiss, A. J. Roman, A. V. Cideciyan, S. B. Schwartz, A. M. Komáromy, M. Doobrajh, A. Y. Cheung, A. Sumaroka, S. E. Pearce-Kelling, G. D. Aguirre, S. Kaushal, A. M. Maguire, T. R. Flotte, and W. W. Hauswirth, Safety in Nonhuman Primates of Ocular AAV2- RPE65, a Candidate Treatment for Blindness in Leber Congenital Amaurosis, *Human Gene Therapy* 17, 845 (**2006**).
8. L. H. Vandenberghe, P. Bell, A. M. Maguire, C. N. Cearley, R. Xiao, R. Calcedo, L. Wang, M. J. Castle, A. C. Maguire, R. Grant, J. H. Wolfe, J. M. Wilson, and J. Bennett, Dosage Thresholds for AAV2 and AAV8 Photoreceptor Gene Therapy in Monkey, *Science Translational Medicine* 3, 88ra54 (**2011**).
9. L. Yin, K. Greenberg, J. J. Hunter, D. Dalkara, K. D. Kolstad, B. D. Masella, R. Wolfe, M. Visel, D. Stone, R. T. Libby, D. Diloreto, D. Schaffer, J. Flannery, D. R. Williams, and W. H. Merigan, Intravitreal injection of AAV2 transduces macaque inner retina., *Investigative Ophthalmology & Visual Science* 52, 2775 (**2011**).
10. Q. Lu, T. H. Ganjawala, E. Ivanova, J. G. Cheng, D. Troilo, and Z.-H. Pan, AAV-mediated transduction and targeting of retinal bipolar cells with improved mGluR6

promoters in rodents and primates, 23, 680 (**2016**).
11. A. Sengupta, A. Chaffiol, E. Macé, R. Caplette, M. Desrosiers, M. Lampič, V. Forster, O. Marre, J. Y. Lin, J.-A. Sahel, S. Picaud, D. Dalkara, and J. Duebel, Red-shifted channelrhodopsin stimulation restores light responses in blind mice, macaque retina, and human retina, *EMBO Mol Med* 8, 1248 (**2016**).
12. A. J. Lotery, G. S. Yang, R. F. Mullins, S. R. Russell, M. Schmidt, E. M. Stone, J. D. Lindbloom, J. A. Chiorini, R. M. Kotin, and B. L. Davidson, Adeno-associated virus type 5: transduction efficiency and cell-type specificity in the primate retina., *Human Gene Therapy* 14, 1663 (**2003**).
13. S. E. Boye, J. J. Alexander, S. L. Boye, C. D. Witherspoon, K. J. Sandefer, T. J. Conlon, K. Erger, J. Sun, R. Ryals, V. A. Chiodo, M. E. Clark, C. A. Girkin, W. W. Hauswirth, and P. D. Gamlin, The Human Rhodopsin Kinase Promoter in an AAV5 Vector Confers Rod- and Cone-Specific Expression in the Primate Retina, *Human Gene Therapy* 23, 1101 (**2012**).
14. K. Mancuso, W. W. Hauswirth, Q. Li, T. B. Connor, J. A. Kuchenbecker, M. C. Mauck, J. Neitz, and M. Neitz, Gene therapy for red–green colour blindness in adult primates, *Nature* 461, 784 (**2009**).
15. C. L. Bell, L. H. Vandenberghe, P. Bell, M. P. Limberis, G.P. Gao, K. Van Vliet, M. Agbandje-McKenna, and J. M. Wilson, The AAV9 receptor and its modification to improve in vivo lung gene transfer in mice, *J. Clin. Invest.* 121, 2427 (**2011**).
16. L. Zhong, B. Li, G. Jayandharan, C. S. Mah, L. Govindasamy, M. Agbandje-McKenna, R. W. Herzog, K. A. Weigel-Van Aken, J. A. Hobbs, S. Zolotukhin, N. Muzyczka, and A. Srivastava, Tyrosine-phosphorylation of AAV2 vectors and its consequences on viral intracellular trafficking and transgene expression., *Virology* 381, 194 (**2008**).
17. H. Petrs-Silva, A. Dinculescu, Q. Li, W.T. Deng, J.-J. Pang, S.-H. Min, V. Chiodo, A. W. Neeley, L. Govindasamy, A. Bennett, M. Agbandje-McKenna, L. Zhong, B. Li, G. R. Jayandharan, A. Srivastava, A. S. Lewin, and W. W. Hauswirth, Novel Properties of Tyrosine-mutant AAV2 Vectors in the Mouse Retina, *Molecular Therapy* 19, 293 (**2009**).
18. H. Petrs-Silva, A. Dinculescu, Q. Li, S.H. Min, V. Chiodo, J.J. Pang, L. Zhong, S. Zolotukhin, A. Srivastava, A. S. Lewin, and W. W. Hauswirth, High-efficiency Transduction of the Mouse Retina by Tyrosine-mutant AAV Serotype Vectors, *Molecular Therapy* 17, 463 (**2008**).
19. G.J. Ye, E. Budzynski, P. Sonnentag, P. E. Miller, A. K. Sharma, J. N. Ver Hoeve, K. Howard, D. R. Knop, M. Neuringer, T. McGill, J. Stoddard, and J. D. Chulay, Safety and Biodistribution Evaluation in Cynomolgus Macaques of rAAV2tYF-CB-hRS1, a Recombinant Adeno-Associated Virus Vector Expressing Retinoschisin, *Human Gene Therapy Clinical Development* 26, 165 (**2015**).
20. M. A. Kotterman and D. V. Schaffer, Engineering adeno-associated viruses for clinical gene therapy, *Nat Rev Genet* 15, 445 (**2014**).
21. D. Dalkara, L. C. Byrne, R. R. Klimczak, M. Visel, L. Yin, W. H. Merigan, J. G. Flannery, and D. V. Schaffer, In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous., *Science Translational Medicine* 5, 189ra76 (**2013**).
22. P. Charbel Issa, S. R. De Silva, D. M. Lipinski, M. S. Singh, A. Mouravlev, Q. You, A. R. Barnard, M. W. Hankins, M. J. During, and R. E. MacLaren, Assessment of Tropism and Effectiveness of New Primate-Derived Hybrid Recombinant AAV Serotypes in the Mouse

and Primate Retina, *PLoS ONE* 8, e60361 (**2013**).
23. M. WEBER, Recombinant adeno-associated virus serotype 4 mediates unique and exclusive long-term transduction of retinal pigmented epithelium in rat, dog, and nonhuman primate after subretinal delivery, *Molecular Therapy* 7, 774 (**2003**).
24. D. Amado, F. Mingozzi, D. Hui, J. L. Bennicelli, Z. Wei, Y. Chen, E. Bote, R. L. Grant, J. A. Golden, K. Narfstrom, N. A. Syed, S. E. Orlin, K. A. High, A. M. Maguire, and J. Bennett, Safety and efficacy of subretinal readministration of a viral vector in large animals to treat congenital blindness., *Science Translational Medicine* 2, 21ra16 (**2010**).
25. J. Bennett, M. Ashtari, J. Wellman, K. A. Marshall, L. L. Cyckowski, D. C. Chung, S. McCague, E. A. Pierce, Y. Chen, J. L. Bennicelli, X. Zhu, G.-S. Ying, J. Sun, J. F. Wright, A. Auricchio, F. Simonelli, K. S. Shindler, F. Mingozzi, K. A. High, and A. M. Maguire, AAV2 gene therapy readministration in three adults with congenital blindness., *Science Translational Medicine* 4, 120ra15 (**2012**).
26. P. J. B. MD, J. W. MSc, K. A. M. COT, S. M. BA, M. A. PhD, J. D.-P. BA, O. U. E. PhD, D. C. C. DO, J. S. MSc, J. F. W. PhD, D. R. C. MPH, P. A. BA, L. L. C. BA, J. L. B. PhD, F. M. PhD, A. A. MD, E. A. P. MD, J. R. MD, P. B. P. L. MD, P. F. S. MD, K. A. H. MD, and P. A. M. M. MD, Safety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by, *The Lancet* 388, 661 (**2016**).
27. Q. Li, R. Miller, P.Y. Han, J. Pang, A. Dinculescu, V. Chiodo, and W. W. Hauswirth, Intraocular route of AAV2 vector administration defines humoral immune response and therapeutic potential., *Mol. Vis.* 14, 1760 (**2008**).
28. M. A. Kotterman, L. Yin, J. M. Strazzeri, J. G. Flannery, W. H. Merigan, and D. V. Schaffer, Antibody neutralization poses a barrier to intravitreal adeno-associated viral vector gene delivery to non-human primates, *Gene Ther* 22, 116 (**2014**).
29. J. G. Flannery and M. Visel, Adeno-Associated Viral Vectors for Gene Therapy of Inherited Retinal Degenerations, in *Methods in Molecular Biology*, Totowa, NJ, Humana Press (**2012**), pp. 351–369.
30. K. Rapti, V. Louis-Jeune, E. Kohlbrenner, K. Ishikawa, D. Ladage, S. Zolotukhin, R. J. Hajjar, and T. Weber, Neutralizing Antibodies Against AAV Serotypes 1, 2, 6, and 9 in Sera of Commonly UsedAnimal Models, *Molecular Therapy* 20, 73 (**2009**).

Chapter 4: Directed evolution to increase the carrying capacity of AAV

Abstract

Inherited retinal degenerations are a major source of blindness, which historically have had very few therapeutic options. Gene therapy has long held the promise of a treatment to address the underlying cause of genetic diseases. The retina offers many advantages for gene therapy treatments including the relatively small tissue size, accessibility, internal control, and immune privilege. Among the gene delivery vehicles, adeno-associated virus (AAV) has risen to the top as the vector of choice and found clinical success in several clinical trials, most notably for the treatment of Leber congenital amaurosis type 2. AAV is one of the smallest viruses in terms of genome and particle size, so consequently, a major hurdle to treating a larger range of genetic eye diseases is the limited packaging capacity of ~5 kb. Past approaches implemented to overcome this hurdle included dividing the transgene between multiple vectors or utilizing proteasome inhibitors. Both solutions result in truncated transgenes and a loss in efficiency, which renders achieving meaningful clinical outcomes challenging and therefore additional strategies, must be developed. Here we explore a strategy to increase the packaging capacity of AAV by utilizing the process of directed evolution, which involves the generation of variant libraries and subsequent iterative selection of variants with the desired property. Specifically, three different sizes (5.4 kb, 5.7 kb, and 6.4 kb) were tested with several different AAV libraries. Focus was placed on introducing mutations into the β-barrel region, which is internal to the capsid and has the possibility of changing the steric and electrostatic interactions between the capsid proteins and packaged ssDNA. This library was created using synthetic nucleotide analogs to induce mutations. The selection was accomplished by packaging an oversized transgene that encoded a particular variant, serving as a "barcode" for that capsid, as well as GFP 3' of the *cap* sequence. The expression of GFP was then used for *in vitro* selection. In the case of the 5.7 kb and 6.4 kb vectors, a high number of false positives were detected leading to a convergence of AAV2 sequences. In order to increase the likelihood of selecting a functional library variant and decrease the number of false positives, the vector was redesigned to 5.4 kb and the GFP sequence was moved to 5' of the *cap* sequence and inverted. In order to accommodate the GFP sequence within the reduced the size of 5.4 kb, the *rep* sequence was removed and *cap* expression was driven by a CMV promoter. Performing selections with this construct led to a convergence of AAV2 variants with six mutations resulting in four different amino acid changes. When packaged with an oversized construct the variants failed to express GFP demonstrating that the mutations did not lead to a functional vector carrying an oversized transgene. This is could be due to the relatively small percentage of the evolutionary landscape explored in the libraries tested or the possibility that mutations resulting in increased packaging capacity are too destabilizing. It is therefore suggested that alternative strategies be implemented to increase the packaging capacity or the use of novel approaches that fit within the natural packaging capacity of AAV.

Introduction

Inherited retinal degenerations are a heterogeneous group of genetic diseases, the majority of which converge on the phenotypic presentation of progressive degeneration of photoreceptor cells leading to blindness[1,2]. To date, more than 250 mutations have been identified[3]. AAV has been utilized in a number of clinical trials for retinal degenerations[4-8]. The largest milestone was accomplished in 2008, with multiple reports of the efficacy of *RPE65* treatment for LCA2, especially in younger patients[9-12]. Despite its many advantages as a vector and its clinical success, the relatively small packaging capacity of AAV is still a major drawback to treating many forms of retinal degenerations. The AAV genome is 4.7 kb and has a total packaging capacity of approximately 5 kb, after which titers drastically decrease[13,14]. Packaging and expression of a transgene requires two ITRs (145 bp) and a promoter further limiting the size of the therapeutic gene that can be delivered to usually < 4.0 kb. The current packaging capacity hinders the use of AAV as a viral vector for the treatment of many retinal degenerations that necessitates larger cDNA including Stargardt disease (*ABCA4* with cDNA of ~6800 bp), Usher syndrome type 1B (MYO7A with cDNA ~6600 bp), Usher syndrome type 2 (USH2B with cDNA ~4640 bp), retinitis pigmentosa type 12 (CRB1 with cDNA of ~4200 bp), Leber congenital amaurosis (CEP290 with cDNA ~7400 bp), recessive congenital stationary night blindness (*TRPM1* with cDNA ~4800 nt), X-linked congenital stationary night blindness (*CACNA1F* with cDNA ~5900 nt), and other retinal degenerations. In addition, spCas9 genome editing strategies with the guide RNA packaged in a single vector can only be accomplished to-date with suboptimal shortened promoters pushing the limit of the carrying capacity of AAV and would benefit from an increased packaging limit[15]. It has previously been reported in the literature that AAV2 is capable of carrying 6 kb genetic cargo when accompanied by a proteasome inhibitor and that AAV5 is capable of carrying up to 8.9 kb[16,17]. These findings failed to be replicated and later several groups subsequently showed that the oversized constructs were not being fully packaged[18-20]. Instead, truncated products were carried in separate vectors and then through either second strand synthesis of genomes of opposite polarity or through homologous recombination the full genome was generated after delivery. The studies also demonstrated that 5.2 kb is the maximum size transgene that can detected via Southern blot for wild-type AAVs.

While the crystal structures of multiple AAV serotypes have been published, the viral genome is not visualized in the structure[21-25]. A rational design of improved vectors is especially challenging in this case due to the relatively large size of these regions and the unfortunate lack of mechanistic understanding of structure-function relationships. In order to overcome this challenge, a directed evolution approach was taken to increase the carrying capacity of AAV for the treatment of retinal degenerations requiring an oversized transgene. By engineering the capsid of natural variants, we hypothesized that the packaging capacity could be increased. This was with a specific focus on mutating amino acids that are both internal to the capsid and interact with the ssDNA genome. There is literature that has demonstrated a strong correlation between the number of positively charged amino acids in the interior of the capsid and the size of the genome packaged[26]. Thus, increasing the overall positive charge of the internal capsid could lead to tighter coiling of DNA and interactions with the negatively charged phosphate backbone of the transgene. In addition, increasing the basic residues in the β-barrel region or the N-terminals of VP1 and VP2 (which are located on the interior of the capsid) could modulate the steric

effects of the cDNA and increase the packageable genome size[21,27]. This motif is a highly conserved jelly-roll β-barrel that is made of two anti-parallel β-sheets oriented with respect to the icosahedral symmetry axes, which is common to parvoviruses and two to four times larger than in small RNA viruses[21]. The observation that other viruses in the *Parvoviridae* family, such as B19 and minute virus of mouse, which have a similar diameter of 25 nm but are capable of packaging over 5 kb of genetic cargo also supported the idea that this could be applied to AAV[26,28].

Directed evolution was applied to AAV at three different transgene sizes (5.4kb, 5.7kb, and 6.4kb) with a number of capsid libraries. Error-prone PCR is an effective way to generate libraries. Using promiscuous nucleotide analogs that are able to base-pair with specific subsets of nucleotides, mutations were targeted to the β-barrel region. Specifically the triphosphate derivatives of 6-(2-deoxy-b-D-ribofuranosyl)-3,4-dihydro-8H-pyrimido-[4,5-C][1,2]oxazin-7-one (dPTP) and 8-oxo-2'deoxyguanosine (8-oxo-dGTP) were used. The nucleotide analog dPTP basepairs both with A and G while 8-oxo-dGTP pairs with A and C. The analogs are mixed with normal nucleotides and periodically incorporated into the nascent DNA strand based on the sequence of the original template during a normal cycle of PCR with Taq polymerase. During subsequent rounds of PCR, the nascent strand becomes a template strand and the analogs will promiscuously pair leading to a mutation at that position. After a small number of rounds of PCR, the mutagenized template undergoes a new round of PCR reactions using only native nucleotides in order to amplify the mutations that were generated while significantly diluting the strands that contain the nucleotide analogs to reduce that chance that a library member with the analog is packaged in AAV. This method has multiple advantages over traditional error-prone PCR including mutation tunability, ability to mutate short nucleotide sequences, decrease likelihood of stop codons, and decrease indel generation. Because the analogs are biased against dTTPs, the percentage of stop codons is reduced increasing the overall percentage of viable variants in the library. The number of mutations introduced can be tuned both via the concentration of nucleotide analogs as well as the number of PCR cycles[29,30]. This strategy was successfully utilized to generate libraries targeted at the β-barrel core in AAV2.

Directed evolution was attempted with each of the three different constructs. For the 5.7 kb and 6.4 kb constructs, there was a convergence on wild-type AAV2 likely due to the selection being too challenging for AAV and resulting in false positives. The 5.4 kb construct directly led to strong convergence on AAV2 library members but the selected variants were unable to package oversized cargo.

Results

AAV based libraries have been generated previously for directed evolution applications, which much success[31-34]. We utilized the AAV2 based error-prone library, the AAV9 based error-prone library, and shuffle library, which is comprised of multiple AAV serotypes representing multiple clades[32]. In addition, we engineered a unique library focused on the β-barrel region in order to increase the likelihood of variants that are capable of packaging oversized cargo. Because the β-barrel region is highly conserved among AAV serotypes and *Parvoviridae* in general as well as the fact that it is found internally on the capsid we sought to target this specific sequence. In order to specifically target the β-barrel core, six different

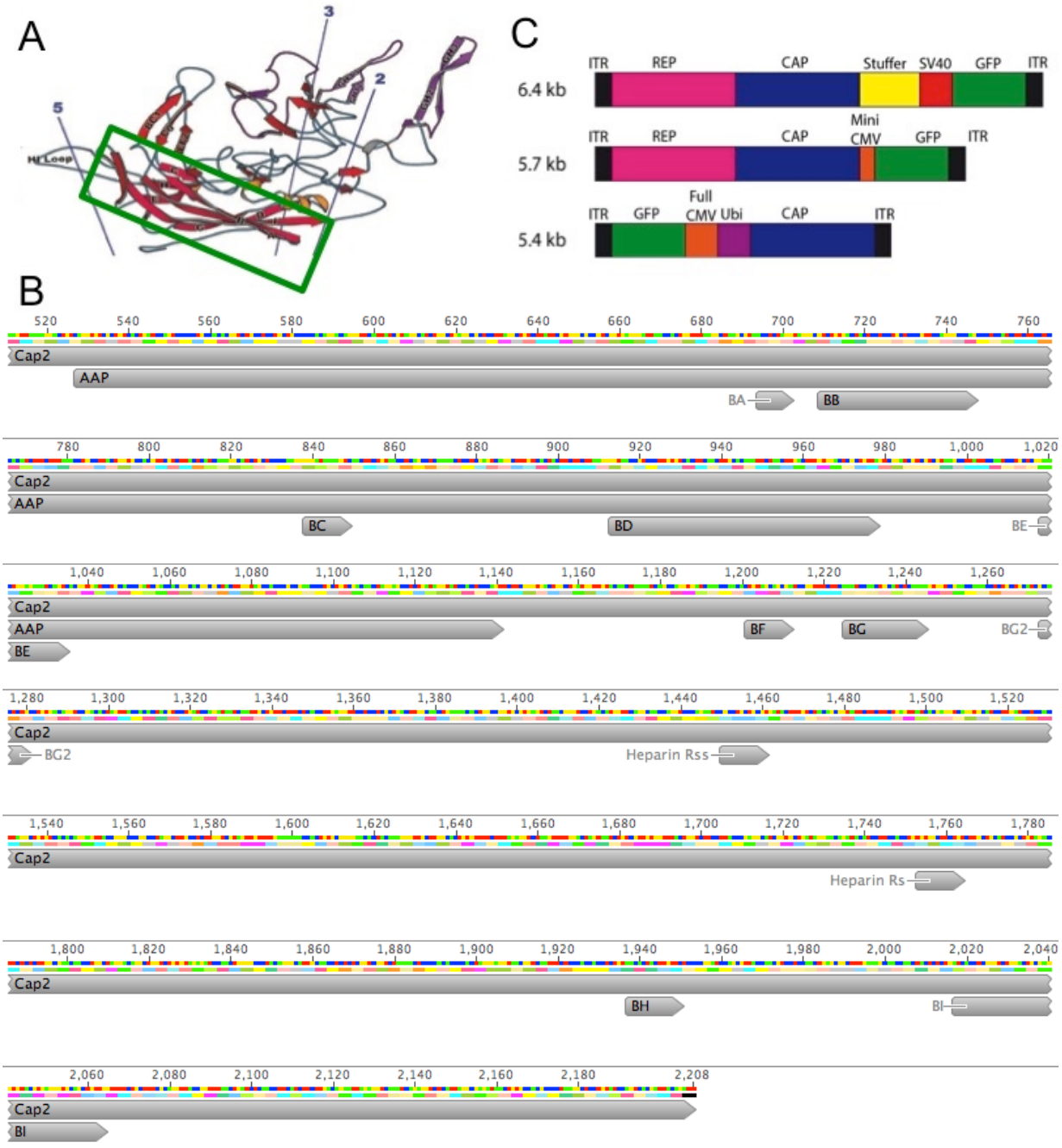


Figure 1. Constructs and β-barrel library for directed evolution. **A)** Ribbon drawing of AAV2 subunit showing the axes of symmetry adapted from Xie *et al.*[21]. The β-barrel region lies on the inner surface of the capsid and is designated by the green rectangle. **B)** Selected sequence of *cap* showing the location of the nine sequences that encode the β-barrel region (BA-BI) as well as the assembly activating protein (AAP). The regions were mutated to create the β-barrel library**. C)** The *cap* libraries were cloned into three different vectors with different sizes. The 6.4 kb construct required a stuffer sequence in order to reach the 6.4 kb. The 5.7 kb vector was optimized to be the smallest vector while still keeping the entire AAV genome by using a miniaturized CMV 323 bp promoter. The 5.4 kb construct eliminated the *rep* sequence in order to decrease the overall size and a Full CMV promoter was utilized to maximize GFP expression. The promoter-GFP sequence was reversed and moved to 5' of the cap sequence in order to reduce the potential for false positives.

restriction sites were created in the AAV *cap* region via multiple rounds of site-directed mutagenesis to serve as handles for cloning. Four of the changes led to silent mutations, one led to a Tyr to Phe substitution (maintaining the hydrophobic side chain and aromatic ring), and the last led to a Leu to Ile substitution (maintaining the hydrophobic side chain). By utilizing the six new restriction sites and the endogenous *BsiW*I, four different regions of approximately 150 – 230 bp were mutagenized (**Fig. 1 A-B**). Mutagenic PCR using the nucleotide analogs dPTP and 8-oxo-dGTP was necessary to induce mutations in the small β-barrel coding region because traditional error-prone PCR does not produce mutations at a high enough frequency. Each library contained 10^6 – 10^7 variants and prepackaging yielded ~80% of variants with amino acid changes in the amplified regions (**Table 1**). The target β-barrel library along with the error-prone AAV2, error-prone AAV9, and shuffle libraries were then cloned into the 6.4 kb, 5.7 kb, and 5.4 kb constructs.

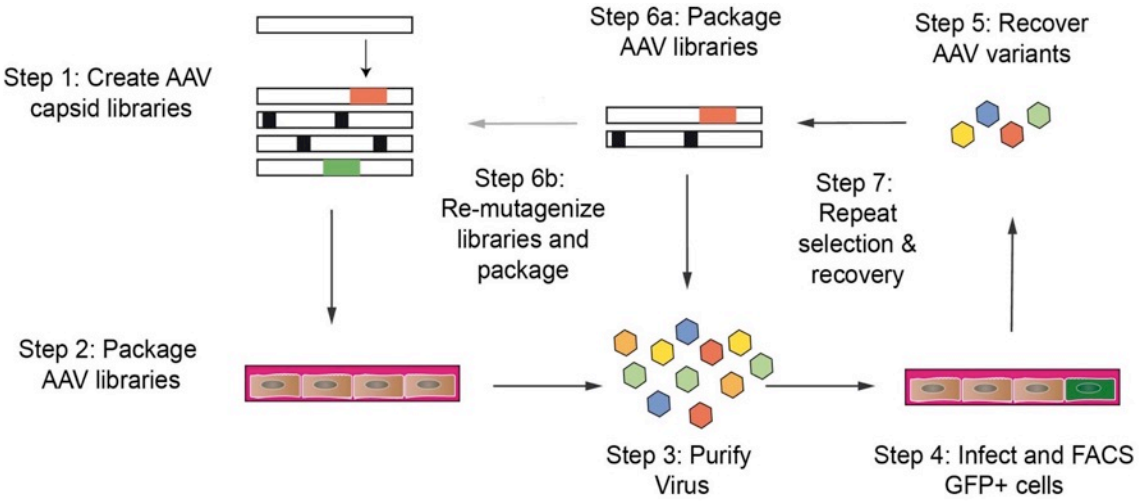


Figure 2. Directed evolution to increase the packaging capacity. Step 1 is the introduction mutations into the cap sequence of AAV library variants. Step 2 and 3 is to package the different library variants and purify the virions. Step 4 is the introduction of the selection pressure of increased carrying capacity, which results in GFP expression. In step 5, via FACS, GFP+ cells were isolated and cap sequences were recovered through PCR. In step 6a, these variants were repackaged and underwent selection pressure again. Step 6b allows for the introduction of new mutations for additional rounds of mutations. In Step 7, additional rounds of selection are performed until convergence is reached.

To initially challenge the packaging capacity of AAV, a 6.4 kb construct was cloned with a SV40 promoter and GFP placed 3' of the *cap* sequence (**Fig. 1 C**). Because AAV truncates the packaged transgene once capacity has been reached, it was hypothesized that only variants capable of packaging oversized cargo would express GFP when later infected. *In vitro* directed evolution was performed first on libraries cloned in the 6.4 kb construct. To perform directed evolution for increased carrying capacity, libraries were pooled and infected at an MOI of 1000 in a near confluent plate of HEK293T cells. After 72 hours, fluorescence-activated cell sorting (FACS) of 10^6 cells was performed and ~4000 GFP+ cells were recovered each round. DNA was extracted and from the GFP+ cells and two additional round of selection were performed (**Fig. 2**). After three rounds of selection, there was a strong convergence onto the wild-type AAV2 *cap* sequence. A 5.7 kb construct was generated by cloning a 323 bp CMV promoter driving GFP to reduce the selection pressure and lead to a larger number of variants capable of packaging larger

genes. The 5.7 kb construct was then tested with the same libraries. After three rounds of selection there was a similar convergence onto the wild-type AAV2 sequence.

The 5.7 kb construct was the smallest vector that could be cloned while retaining the entire wild-type AAV genome. To lower the selection pressure below 5.7 kb, components of the AAV genome needed to be removed. Nathwani *et al.* demonstrated the successful packaging of AAV utilizing a CMV promoter to drive cap expression[35]. Utilizing this strategy a 5.4 kb construct was cloned by eliminating the *rep* sequence, opting to drive *cap* expression via a exogenous promoter. To mitigate the possibility of recombination and truncation of the construct two different promoters were utilized and the promoter-GFP sequence was inverted. This inversion reduces the possibility of false positives because truncations would either occur in the GFP sequence or the *cap* sequence (**Fig. 1A**). A Ubi-C promoter and a Full CMV promoter were selected to drive GFP and *cap*. Both were tested for driving GFP expression and for packaging with the 5.4 kb construct. The Ubi-C promoter was not as strong in HEK293T cells resulting in both lower titers when driving *cap* and lower GFP expression than the Full CMV promoter (**Fig. 3**). This supported using the construct with GFP driven by CMV to maximize GFP expression and increase the number of cells recovered by FACS. The selected library variants were cloned into this backbone and one round of selection was performed leading to very strong convergence of AAV2-based variants with two to five point mutations (**Table 2**).

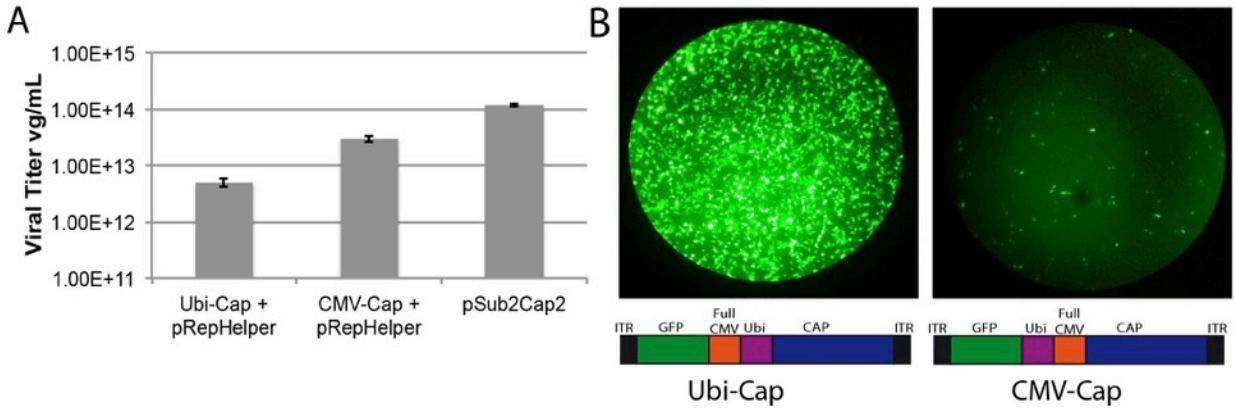


Figure 3. Packaging and transfection of 5.4 kb backbones. A) Titering of 5.4 kb constructs with *cap* expression driven by either Ubi promoter or CMV promoter with *rep* supplied in *trans*. Titer is reduced when using Ubi promoter to 5e12 vg/mL, titer is 3e13 vg/mL when driven by Full CMV promoter, and the titer of the positive control is 1e14 vg/mL. **B)** Transfection of 1 µg of either construct in a 6-well plate of HEK293T cells imaged after 48 hours. GFP driven by CMV in the Ubi-Cap construct is significantly stronger than GFP driven by the Ubi promoter in the CMV-Cap construct. The Ubi-Cap construct was used as the 5.4 kb construct for directed evolution because of the strong GFP expression and the titer was high enough for *in vivo* selections.

Discussion

The data demonstrates the enormous challenge of increasing the carrying capacity of AAV. We tested backbones that were 200-1200 bp greater than highest reported packaging capacity of 5.2 kb. It is not surprising that the 6.4 kb vector did not result in a successful variant because that is a large size jump that is greater than other *Parvoviridae* of a similar size. This backbone along with the 5.7 kb backbone both had the GFP construct 3' of *cap* which led to a high probability of false positives and a convergence of the natural AAV2 sequence. This is due to the fact that AAV preferentially packages from the 3' end, so consequently when transgenes

are larger than the packaging capacity the DNA is truncated on the 5' end. In this case the positive strand would contain the full SV40-GFP or miniCMV-GFP sequence as well as the *cap* sequence. Truncation would likely occur in the *rep* sequence eliminating one ITR. There is evidence in the literature that single ITR transgenes would be capable of low levels of expression[36]. This would explain the low level of GFP positive cells observed through FACS. It is possible that the mutations in the library were at times destabilizing leading to lower packaging efficiencies of the library variants and convergence of AAV2. When designing the 5.4 kb backbone, *rep* was eliminated, the CMV-GFP sequence was moved to the 5' end of the construct (and inverted so that the GFP sequence was at the very 5'end) and *cap* expression was driven by the Ubi promoter. This lowers the chance of false positives because the truncations would occur either in the GFP sequence preventing selection or in the cap sequences preventing packaging. By reducing the size of the construct and therefore reducing the selection pressure as well as mitigating the chance of false positives there was a convergence onto six different mutations in AAV2. Unfortunately, when the variants were packaged with the 5.4 kb transgene and infected in HEK293T cells there was no GFP expression. Notably, only one point mutation was in the anticipated β-barrel region and another point mutation resulted in conversion of a negatively charged amino acid to a neutral charge. An error-prone PCR was performed on the selected variants to add diversity to the library building upon the selected mutations and a round of selection was performed. We were unable to obtain a successful round of selection with these variants because there was no GFP expression. This suggests that the convergence likely led to variants that are able to package well but still incapable of packaging oversized transgenes. It also validated moving and inversing the GFP sequence 5' of *cap* because fewer false positives were observed and wild-type AAV2 was avoided with selections using the 5.4 kb transgene.

The libraries generated explored a relatively small portion of the evolutionary landscape. So, it is still possible that amino acid changes would lead to increased packaging capacity but due to the random nature of the library generation, none of the necessary mutations occurred. Additional strategies that could be applied in the future would be to use site-directed mutagenesis to introduce specific mutations to increase positive charges in the barrel region (i.e. mutating negatively charged amino acids to neutral amino acids or neutral amino acids to positively charged amino acids). This would ensure that there is a strong overall change in charge on the internal capsid. Ling *et al.* demonstrated that replacing one ITR D-sequence with a substitute sequence eliminates inhibition and leads to greater transgene expression as well as only the positive strand of AAV being packaged[37]. This strategy could be implemented to both increase expression as well as directing all truncations of the transgene to the GFP sequence to help further mitigate the possibility of false positives. To tackle specific diseases that would require oversized transgenes other treatment strategies that fit with in the packaging capacity of AAV could be applied. For example, CRISPR editing of the cryptic splice site of CEP290 could be used for the treatment of LCA10. Dual vectors could also be used although the transfection efficiency would be reduced since both vectors must express and homologous recombination must occur for expression of the full gene. Steps must be taken to mitigate the expression of truncated products as well, which could lead to toxicity. The assembly-activating protein would likely be mutated as well since it shares the same coding region as cap. This could reduce the packaging efficiency of mutant variants.

Methods

Cloning β-barrel library

The *cap* sequence was cloned into a TOPO vector by digesting both the TOPO backbone and a pSub2Cap2G with *Hind*III and *Not*I and ligated with T4 Ligase (Promega) generating the pTOPO-Cap2. Then the following mutations were made systematically using site-directed mutagenesis (SDM; Quikchange) for the following sites: G674A (silent, *Bsu*36I), C909T (silent, *Mfe*I), G1188A and A1190T (Y>F mutation, *EcoR*I), G1335A (silent, *Afl*II), A1857T A1861C T1863A (I>L mutation, *Blp*I), and C2088G (silent, *Acc*III) using the SDM primers found in **Table 3**. The pTOPO-Cap2 mutant was digested with *Hind*III and *Not*I and ligated with T4 Ligase (Promega) back into the AAV-L backbones.

Synthetic nucleotide analog PCR to generate error-prone library

The pTOPO-Cap2Mut construct was used as the template for four different regions that contain the β-barrel sequences. For region 1, pTOPO-Cap2 was digested with *Bsu36*I and *Mfe*I. For region 2, pTOPO-Cap2 was digested with *Mfe*I and *BsiW*I. For region 3, pTOPO-Cap2 was digested with *EcoR*I and *Afl*II. For region 4, pTOPO-Cap2 was digested with *Blp*I and *Acc*III. The corresponding amplification primers for each region can be found in **Table 3**. Each PCR was performed as follows: First the 10x mutagenic buffer 70 mM $MgCl_2$, 500 mM KCl, 100 mM Tris-HCl (pH adjusted to 8.3), 0.1 g/L gelatin was prepared. Next the 10x dNTP mix was prepared as follows: 5 mM dATP, 5mM dGTP, 5 mM dCTP, 5 mM dTTP, 0.5 mM dPTP (Trilink Biotech N-2037), 0.5 mM 8-oxo-dGTP (Trilink Biotech N-2034). Eight parallel PCR reactions were prepared as follows: 5 µL 10x Analog buffer. 5 µL 10x dNTP mix, 1 µL template DNA (10 ng/µL), 2.5 µL 10 µM Forward primer, 2.5 µL 10 µM Reverse primer, 1 µL Taq Polymerase, and 33 µL water. Using a thermocycler. The following protocol was utilized: 1. 2 min 92º C; 10 cycles: 1 min 92ºC, 1.5 min 55º C, 5 min/kb 72º C; 3. 1 cycle: 10 min 72º C. The 8 reactions were pooled and used as the template for a second PCR with natural nucleotides only. The master mix was the same as the first PCR but 5 µL 2 mM dNTP (dATP, dGTP, dCTP, dTTP only) was used instead of the PCR mix. Instead of 10 cycles of PCR amplification, 20 rounds of amplification was performed. The PCR reactions were then digested with the same restriction enzymes as the backbone it was cloned into. Each library was packaged and titered as previously described [33,38].

Cloning oversized AAV backbones

The pSub2Cap2G construct is a modified version of the pSub2Cap2 plasmid which contains both a NotI site and SpeI site 3' of the cap sequence but before the 3' ITR using site-direct mutagenesis (Quikchange) and served as 6.4 kb backbone. Ligation independent cloning was utilized to create the 5.7kb backbone. The pSub2Cap2G backbone was digested with *Not*I and *Spe*I. A 113bp mini CMV promoter derived from Ostedgaard *et al.* driving GFP expression was cloned using Gibson Assembly (New England BioLabs) utilizing the 5.7Fwd and 5.7Rev primers[39]. A UbiC-GFP plasmid from Emory University was digested with *Age*I and *Sph*I, the *cap* sequence from pSub2Cap2G was amplified with IFAUbiCap2+/- primers (note that the forward primer also eliminated a *Not*I site from the sequence by removing a "c" basepair), and

Gibson Assembly was used to generate pUbi-Cap. A plasmid from kindly supplied by the University of North Carolina (UNC) with CMV-GFP with a 50 bp polyA was PCR amplified with InverseUNCpA50+/- primers, and Gibson cloning was used to create the 5.4kb backbone with an inverted CMV-GFP.

AAV vector production

HEK293T cells (ATCC; Manassas, VA) were procured and cultured using Dulbecco's Modified Eagle's Medium (DMEM, Gibco) with 10% fetal bovine serum (Invitrogen) and 2 mM L-Glutamine (Thermo Fisher Scientific) at 37º C with 5% CO_2. AAV libraries were packaged in HEK293T cells as previously described [34,38,40]. In brief, transient transfection of the library, pHelper, and Rep Helper (if rep was removed from the backbone) was performed, vectors were purified by iodexanol density centrifugations, and Amicon filtration was then used to buffer exchange into PBS. Titering of viral genomes was determined by quantitative real time PCR using an iCycler (Bio-Rad).

Library selection and evolution

HEK 293T cells were seeded into a 6-well plate. Each well was infected with the pooled libraries at an MOI of 1000. After 72 hours, cells were harvested and FACS (Cytopeia Influx Cell Sorter, BD Biosciences) to isolate GFP+ positive cells. DNA was then isolated using the DNeasy Blood and Tissue kit (Qiagen) and *cap* sequences were recovered using the vgDNA primers (**Table 3**). The amplified DNA was digested with *Hind*III and *Not*I and ligated overnight using T4 ligase (Promega) into the 6.4 kb, 5.7 kb, or 5.4 kb backbones. The recovered variants were then repackaged for two more iterative rounds of selection with 10 colonies sequenced each round to evaluate convergence.

Table 1. Sequencing of β-barrel library members

β-barrel Region 1	Mutation 1	Mutation 2	Mutation 3	Mutation 4	Mutation 5	Mutation 6	Mutation 7	Mutation 8	Mutation 9
1	Asp283Asn BC	Pro293Leu							
2	Thr251Ala	Ile260Ile							
3	Ile240Ser BB	Gly265Gly	Asn270Asp	Asp295Asp					
4	Asn254DELETION	Asn268Asp	Asp269Gly						
5	Thr245Thr BB	Asn254Asp	Asp296Gly						
6	No mutations								
7	No mutations								
8	Val239Ala BB	Gln259STOP	Asn268Asn	Gly274Asp	Pro278Ser	Gly280Glu BC	Tyr281His BC	His290Arg	Pro293Leu
9	Thr244Ala BB	Asn253Asp	Leu256Pro	Tyr272Cys	Phe273Ser	Asn285Ser			
10	No mutations								
β-barrel Region 2									
1	Thr324Ala	Ile332Ile							
2	Val320Ala	Lys321Gln	Leu350Pro						
3	Gln325STOP								
4	Phe306Ser BD	Asp327Asp							
5	Phe316Phe	Thr330Ala							
6	Thr329Met	Va342Ala BE							
7	Phe306Ser BD								
8	Phe313Leu BD	Phe343Phe BE	Thr344Ile BE						
9	No mutations								
10	No mutations								
β-barrel Region 3									
1	Asn408Ser	Phe415Phe	Phe420Ser	His421His	Ala425Val BG	Met434Ile	Pro436Leu	Leu437Pro	
2	Ala425Thr BG	Asn435Ser							
3	No mutations								
4	Phe420Leu	Ser422Gly	Leu433His						
5	Arg404Arg BF	Leu433Pro							
6	Met402Val BF	His426Arg BG							
7	Tyr424Cys	Asn435Asn							
8	No mutations								
9	No mutations								
10	No mutations								
β-barrel Region 4									
1	Met634Val	Lys665Arg	Trp684Arg BI	Gln687Pro BI					
2	Gln645Pro	Phe661Ile	Lys688Glu BI						
3	No mutations								
4	Met634Val								
5	No mutations								
6	Gly636Gly	Ile648Ile	Lys649Lys	Ser662Asn	Ala664Ala	Ser674Phe BI	Leu686Leu BI		
7	Asn656Ser	Thr675Met BI	Val680Met BI	Gln687STOP BI					
8	His627His	Asn656Asn							
9	Phe628Leu	Ile646Thr							
10	Lys640Arg	Ala663Ala							

Table 2. Sequence convergence of AAV2-based library variants

Variant	Mutation 1	Mutation 2	Mutation 3	Mutation 4	Mutation 5	Mutation 6
1		Ser85Gly		Pro250Ser		
2	Asp41Gly	Ser85Gly		Pro250Ser	Pro644Pro	Val708Ile
3	Asp41Gly	Ser85Gly		Pro250Ser		
4	Asp41Gly	Ser85Gly			Pro644Pro	Val708Ile
5		Ser85Gly		Pro250Ser	Pro644Pro	Val708Ile
6	Asp41Gly	Ser85Gly		Pro250Ser	Pro644Pro	
7	Asp41Gly	Ser85Gly		Pro250Ser	Pro644Pro	Val708Ile
8		Ser85Gly		Pro250Ser	Pro644Pro	Val708Ile
9		Ser85Gly	Gly130Gly	Pro250Ser	Pro644Pro	Val708Ile
10		Ser85Gly		Pro250Ser	Pro644Pro	Val708Ile
11		Ser85Gly		Pro250Ser	Pro644Pro	Val708Ile
12	Asp41Gly	Ser85Gly		Pro250Ser	Pro644Pro	Val708Ile

Table 3. Primers used for cloning

Primers	Sequence
5.7Fwd	tacctgactcgtaatctgtaattgcggccgcaaaatcaacgggactttccaaaatgtcg
5.7Rev	ccatgctctagagcatggaaactagtaaaaacctcccacacctccc
IFAUbiCap2Fwd	aggggttcctgcggcgcatgcttgactagtaacccgtgtcggctcca
IFAUbiCap2Rev	cggcagccataccggtaattcgaagcttgagctcgacctg
InverseUNCpA50 Fwd	ggggttcctgcggcgGCCGCATCTAGAACAATTGCACACAA AA
InverseUNCpA50 Rev	gggttactagtcaag AAGATCGGAATTCGCCCTTaagctag
vgDNAFwd	GCGGAAGCTTCGATCAACTACG
vgDNARev	TCAGTTGAACTTTGGTCTCTGCG
Forward ITR primer	GGAACCCCTAGTGATGGAGTT
Reverse ITR primer	CGGCCTCAGTGAGCGA
AAV2 ITR probe	FAM-CACTCCCTCTCTGCGCGCTCG-BHQ
SDM Bsu36I Fwd	ggagtgggtaattcctcaggaaattggcattgcga
SDM Bsu36I Rev	tcgcaatgccaatttcctgaggaattacccactcc
SDM MfeI Fwd	tggcaaagactcatcaacaacaattggggattccga
SDM MfeI Rev	tcggaatccccaattgttgttgatgagtctttgcca
SDM EcoRI Fwd	tcttcattttactgcctggaattctttccttctcagatgctgc
SDM EcoRI Rev	gcagcatctgagaaggaaagaattccaggcagtaaaatgaaga
SDM AflII Fwd	catcgaccagtacctgtattacttaagcagaacaaacactc
SDM AflII Rev	gagtgtttgttctgcttaagtaatacaggtactggtcgatg
SDM BlpI Fwd	gggcccatctgggctaagctaccacacacggacgg
SDM Blp Rev	ccgtccgtgtgtggtagcttagcccagatgggccc
SDM AccIII Fwd	agcaaacgctggaatccggaaattcagtacacttc
SDM AccIII Rev	gaagtgtactgaatttccggattccagcgtttgct
B-barFwd1	tgggtaattcctcaggaaattggcat
B-barRev1	tccccaattgttgttgatgagtctt
B-barFwd2	ggcaaagactcatcaacaacaattg
B-barRev2	aggacgtacgggagctggta
B-barFwd3	cattttactgcctggaattctttcc
B-barRev3	tctgcttaagtaatacaggtactggtc
B-barFwd4	tctgggctaagctaccacacacg
B-barRev4	gaatttccggattccagcgtttgc

References

1. Sullivan, L. S. & Daiger, S. P. Inherited retinal degeneration: exceptional genetic and clinical heterogeneity. *Mol Med Today* **2,** 380–386 (1996).
2. Sahel, J.-A., Marazova, K. & Audo, I. Clinical characteristics and current therapies for inherited retinal degenerations. *Cold Spring Harbor Perspectives in Medicine* **5,** a017111 (2014).
3. Sohocki, M. M. *et al.* Prevalence of mutations causing retinitis pigmentosa and other inherited retinopathies. *Hum. Mutat.* **17,** 42–51 (2001).
4. Constable, I. J. *et al.* Phase 2a Randomized Clinical Trial: Safety and Post Hoc Analysis of Subretinal rAAV.sFLT-1 for Wet Age-related Macular Degeneration. *EBioMedicine* **14,** 168–175 (2016).
5. MD, P. J. B. *et al.* ArticlesSafety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by. *The Lancet* **388,** 661–672 (2016).
6. MacLaren, R. E. *et al.* Retinal gene therapy in patients with choroideremia: initial findings from a phase 1/2 clinical trial. *Lancet* **383,** 1129–1137 (2014).
7. Simonelli, F. *et al.* Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration. *Molecular Therapy* **18,** 643–650 (2009).
8. Ghazi, N. G. *et al.* Treatment of retinitis pigmentosa due to MERTK mutations by ocular subretinal injection of adeno-associated virus gene vector: results of a phase I trial. *Hum. Genet.* **135,** 327–343 (2016).
9. Bainbridge, J. W. B. *et al.* Effect of Gene Therapy on Visual Function in Leber's Congenital Amaurosis. *N Engl J Med* **358,** 2231–2239 (2008).
10. Cideciyan, A. V. *et al.* Human gene therapy for RPE65 isomerase deficiency activates the retinoid cycle of vision but with slow rod kinetics. *Proceedings of the National Academy of Sciences* **105,** 15112–15117 (2008).
11. Hauswirth, W. W. *et al.* Treatment of Leber Congenital Amaurosis Due to RPE65Mutations by Ocular Subretinal Injection of Adeno-Associated Virus Gene Vector: Short-Term Results of a Phase I Trial. *Human Gene Therapy* **19,** 979–990 (2008).
12. Maguire, A. M. *et al.* Safety and Efficacy of Gene Transfer for Leber's Congenital Amaurosis. *N Engl J Med* **358,** 2240–2248 (2008).
13. Dong, J.-Y., Fan, P.-D. & Frizzell, R. A. Quantitative Analysis of the Packaging Capacity of Recombinant Adeno-Associated Virus. *Human Gene Therapy* **7,** 2101–2112 (1996).
14. Hermonat, P. L., Quirk, J. G., Bishop, B. M. & Han, L. The packaging capacity of adeno-associated virus (AAV) and the potential for wild-type-plus AAV gene therapy vectors. *FEBS Lett.* **407,** 78–84 (1997).
15. Ruan, G.-X. *et al.* CRISPR/Cas9-Mediated Genome Editing as a Therapeutic Approach for Leber Congenital Amaurosis 10. *Molecular Therapy* **25,** 331–341 (2017).
16. Grieger, J. C. & Samulski, R. J. Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. *J. Virol.* **79,** 9933–9944 (2005).
17. Allocca, M. *et al.* Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. *J. Clin. Invest.* **118,** 1955–1964 (2008).
18. Dong, B., Nakai, H. & Xiao, W. Characterization of Genome Integrity for Oversized Recombinant AAV Vector. *Molecular Therapy* **18,** 87–92 (2009).

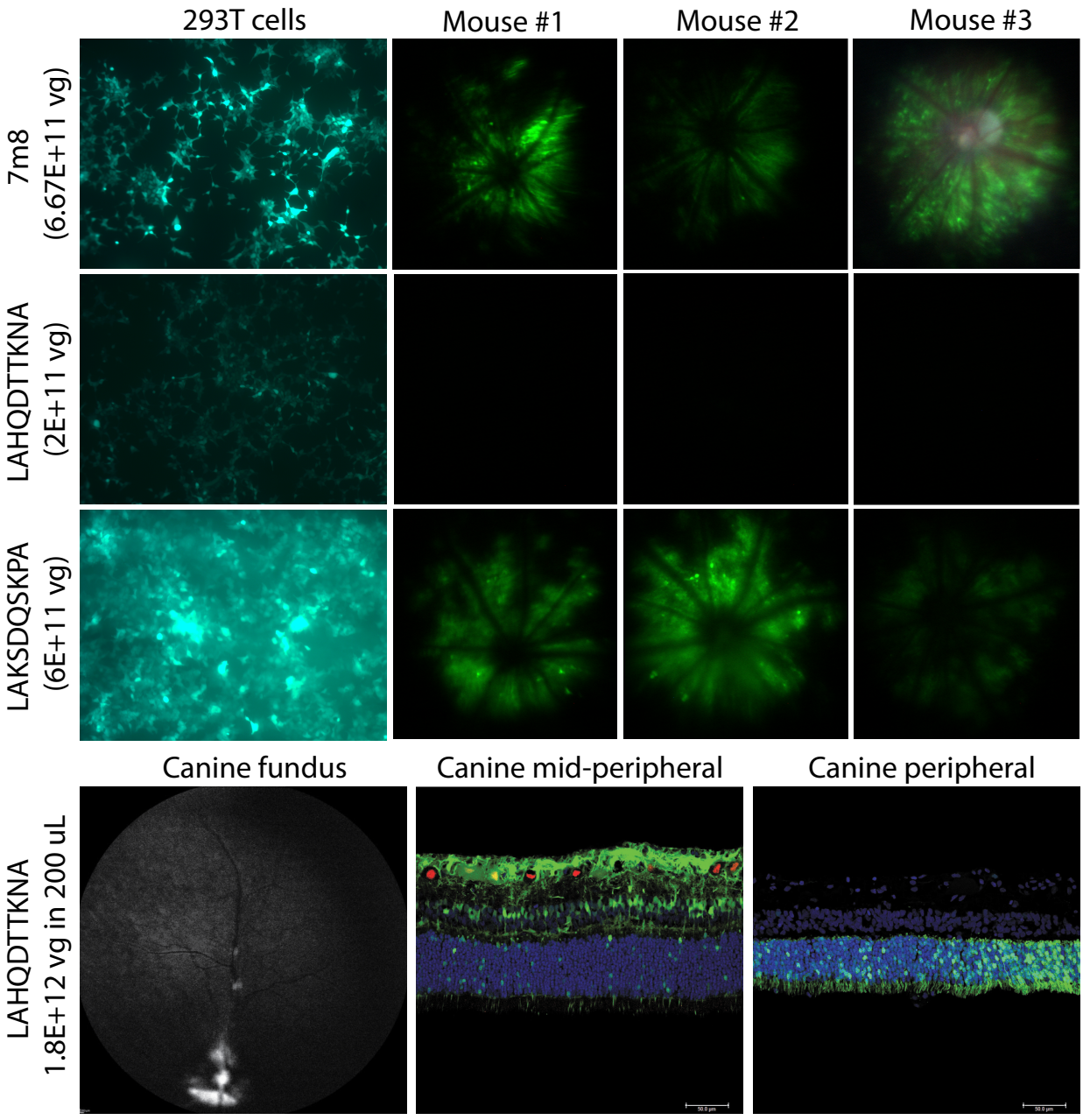
19. Wu, Z., Yang, H. & Colosi, P. Effect of Genome Size on AAV Vector Packaging. *Molecular Therapy* **18,** 80–86 (2009).
20. Lai, Y., Yue, Y. & Duan, D. Evidence for the Failure of Adeno-associated Virus Serotype 5 to Package a Viral Genome ⋀8.2 kb. *Molecular Therapy* **18,** 75–79 (2009).
21. Xie, Q. *et al.* The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 10405–10410 (2002).
22. Ng, R. *et al.* Structural Characterization of the Dual Glycan Binding Adeno-Associated Virus Serotype 6. *J. Virol.* **84,** 12945–12957 (2010).
23. Nam, H.-J. *et al.* Structure of adeno-associated virus serotype 8, a gene therapy vector. *J. Virol.* **81,** 12260–12271 (2007).
24. Govindasamy, L. *et al.* Structurally mapping the diverse phenotype of adeno-associated virus serotype 4. *J. Virol.* **80,** 11556–11570 (2006).
25. Govindasamy, L. *et al.* Structural Insights into Adeno-Associated Virus Serotype 5. *J. Virol.* **87,** 11187–11199 (2013).
26. Belyi, V. A. & Muthukumar, M. Electrostatic origin of the genome packing in viruses. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 17174–17178 (2006).
27. Kronenberg, S., Böttcher, B., Lieth, von der, C. W., Bleker, S. & Kleinschmidt, J. A. A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini. *J. Virol.* **79,** 5296–5303 (2005).
28. Deiss, V., Tratschin, J.-D., Weitz, M. & Siegl, G. Cloning of the human parvovirus B19 genome and structural analysis of its palindromic termini. *Virology* **175,** 247–254 (1990).
29. Zaccolo, M., Williams, D. M., Brown, D. M. & Gherardi, E. An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* **255,** 589–603 (1996).
30. Zaccolo, M. & Gherardi, E. The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. *J. Mol. Biol.* **285,** 775–783 (1999).
31. Klimczak, R. R., Koerber, J. T., Dalkara, D., Flannery, J. G. & Schaffer, D. V. A novel adeno-associated viral variant for efficient and selective intravitreal transduction of rat Müller cells. *PLoS ONE* **4,** e7467 (2009).
32. Koerber, J. T., Jang, J.-H. & Schaffer, D. V. DNA Shuffling of Adeno-associated Virus Yields Functionally Diverse Viral Progeny. *Molecular Therapy* **16,** 1703–1709 (2008).
33. Koerber, J. T., Maheshri, N., Kaspar, B. K. & Schaffer, D. V. Construction of diverse adeno-associated viral libraries for directed evolution of enhanced gene delivery vehicles. *Nat Protoc* **1,** 701–706 (2006).
34. Dalkara, D. *et al.* In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous. *Science Translational Medicine* **5,** 189ra76 (2013).
35. Nathwani, A. C. *et al.* Efficient gene transfer into human cord blood CD34+ cells and the CD34+CD38- subset using highly purified recombinant adeno-associated viral vector preparations that are free of helper virus and wild-type AAV. *Gene Ther* **7,** 183–195 (2000).
36. Wang, X. S., Ponnazhagan, S. & Srivastava, A. Rescue and replication of adeno-associated virus type 2 as well as vector DNA sequences from recombinant plasmids containing deletions in the viral inverted terminal repeats: selective encapsidation of viral genomes in progeny virions. *J. Virol.* **70,** 1668–1677 (1996).
37. Ling, C. *et al.* Enhanced Transgene Expression from Recombinant Single-Stranded D-

Sequence-Substituted Adeno-Associated Virus Vectors in Human Cell Lines In Vitroand in Murine Hepatocytes In Vivo. *J. Virol.* **89,** 952–961 (2015).
38. Flannery, J. G. & Visel, M. in *Methods in Molecular Biology* **935,** 351–369 (Humana Press, 2012).
39. Ostedgaard, L. S. *et al.* A shortened adeno-associated virus expression cassette for CFTR gene transfer to cystic fibrosis airway epithelia. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 2952–2957 (2005).
40. Maheshri, N., Koerber, J. T., Kaspar, B. K. & Schaffer, D. V. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat. Biotechnol.* **24,** 198–204 (2006).

Appendix A: Chapter 2 Supplementary Figures

Round	ID	DOB	Age at injection	Virus injected	Amount virus injected	Notes
1	I-430	9/16/11	11 months	prepacked Shuffle, AAV2-7mer, EP2 and Loopswap libraries	~4E+13 vg total, 1E+13 vg each library	No adverse events
2	I-429	9/16/11	14 months	recovered variants from round 1	~1E+13 vg	No variants recovered; no immune response noted
3	I-428	8/20/11	17 months	recovered variants from round 1 (repeated)	~1E+13 vg	No adverse events (Repeat of previous round)
4	CDJCCL	10/5/12	12 months	recovered variants from round 2	~6.7E+12 vg	OS uveitis
5	CEDCZF	4/26/13	7 months	recovered EP PCR variants from round 3, SCHEMA library	~5E+12 vg	Error prone PCR conducted, No adverse events
6	CEDCVS	4/21/13	9 months	recovered variants from round 4	~5E+12 vg	Uveitis, no detachment, retina grossly normal, no variants recovered
7	CEJCDN	10/3/13	13 months	recovered variants from round 4 (repeated)	~5E+11 vg	No adverse events (Repeat of previous round)
GFP-barcode	CGBCAN	2/1/15	8 months	GFP-barcode library	6E+11 vg each of 21 variants In 250 uL	No adverse events
GFP-barcode	CGBCDI	2/5/15	8 months	GFP-barcode library	6E+11 vg each of 21 variants In 250 uL	No adverse events
GFP-barcode	CGBCGS	2/7/15	8 months	GFP-barcode library	6E+11 vg each of 21 variants In 250 uL	No adverse events
BiP-GFP-barocode	AS318	6/30/10	6 years 9 months	Bipolar-GFP-barocde library	1.56E+13 vg in 400 uL	crd2 carrier
BiP-GFP-barocode	GI94	4/15/11	6 years	Bipolar-GFP-barocde library	1.56E+13 vg in 400 uL	crd3 carrier
Variant validation				AAV2-LAHQDTTKNA	1.8E+12 vg in 200 uL	No adverse events
Variant validation	I575	11/25/13	21 months	AAV2- PAPQDTTKKA	2.0E+11 vg	No adverse events
Variant validation	D354	7/8/16	6.5 months	AAV2- PAPQDTTKKA	6.25E+12 vg	No adverse events
Variant validation	D357	7/8/16	6.5 months	AAV2- PAPQDTTKKA	6.25E+12 vg	No adverse events

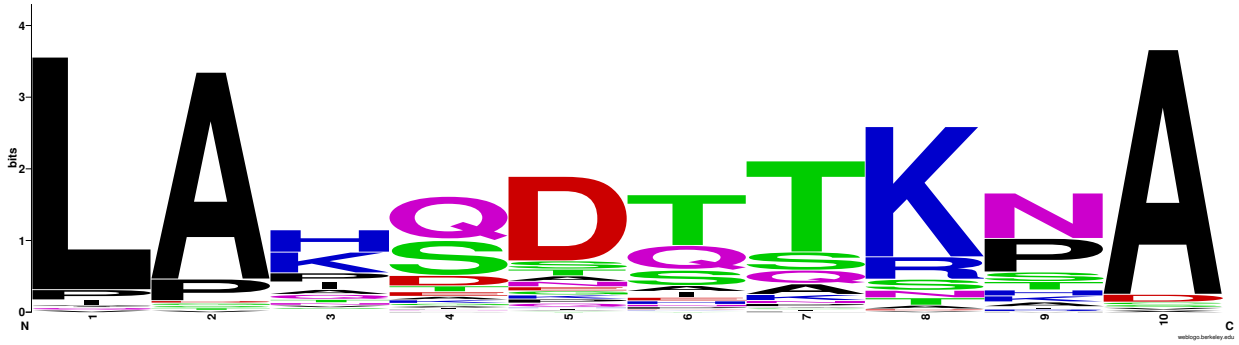
Supplementary Table 1. Summary of the rounds of selection performed in canines. Rounds 3 and 7 were repeated selections of the previous rounds (rounds 2 and 6, respectively), which did not result in the amplification of any AAV variants.



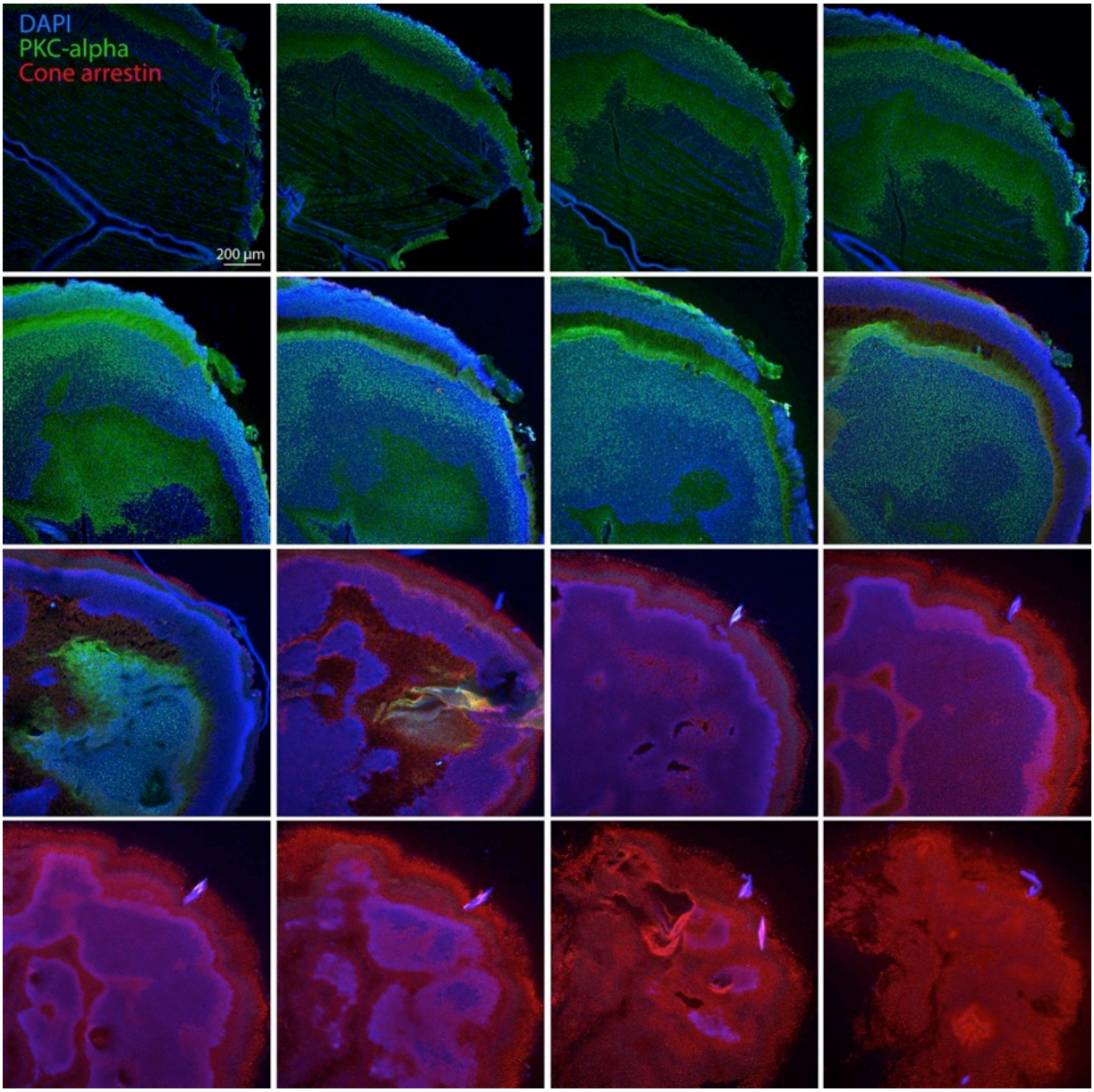
Supplementary Figure 1. AAV2-LAHQDTTKNA infects dog retina but not mouse retina. Similar titers of 7m8-scCAG-eGFP, AAV2-LAHQDTTKNA-scCAG-eGFP, or AAV2-LAKSDQSKPA-scCAG-eGFP (another prominent canine-derived variant) were used to infect HEK293T cells and injected intravitreally in mice (top panel). 7m8 and AAV2-LAKSDQSKPA infect HEK293T cells and mouse retina efficiently, while AAV2-LAHQDTTKNA infects HEK293T cells moderately, but does not infect the mouse retina. In contrast, the canine-derived variant AAV2-LAHQDTTKNA infects the canine retina, resulting in GFP expression in inner and outer retinal layers.

Rank	insert	plasmid	library	r1	r2	r3	r4	r5	r5_v_plasmid
1	LAKDATKTIA	0.10078644	0.003085845	0.021005508	0.145484468	0.147075345	0.527669213	1.429339739	14.18186558
2	LSHQDTTKNA	0.001527067	0.000114291	0.000764258	0.003344013	0.021273508	0.033825853	0.019409833	12.71052945
3	LPHQDTTKNA	0.013743605	0.000236555	0.007306767	0.017078352	0.058747374	0.148242874	0.145093112	10.55713607
4	LAISDQTKHA	0.297778117	0.08908549	0.944321507	5.375660044	2.103716932	0.804977775	2.772114709	9.30932983
5	LAKSDQSKPA	0.45812018	0.005246733	0.085840041	0.07108018	0.115165083	3.235199985	3.742974955	8.170290498
6	LATTSQNKPA	0.077880431	0.027828418	0.101993671	0.076155914	0.214283076	0.285989689	0.59415906	7.629118827
7	LAKDATKNA	0.001527067	5.32E-06	1.16E-05	1.99E-05	4.60E-05	0.000252563	0.011375946	7.449537919
8	LAVSDSTKAA	0.001527067	5.32E-06	1.16E-05	0.000119429	0.006115367	0.001333299	0.011365984	7.443014681
9	LGHQDTTKNA	0.001527067	1.06E-05	0.000544244	0.001612292	0.003984951	0.009849966	0.01105718	7.2407943
10	LAPDSTTRSA	0.26265557	0.260537249	0.442435759	0.829792926	1.733438352	1.005195813	1.797623523	6.844033512
11	LAHQDTTKNA	11.10941437	0.138352687	2.930048415	7.706616209	27.37567834	71.79901793	73.91492705	6.653359448
12	LDHQDTTKNA	0.009162404	0.000194028	0.002246455	0.010589374	0.044554817	0.059898602	0.057661699	6.293293965
13	PAPQDTTKNA	0.001527067	1.33E-05	0.000440027	0.000517526	0.004689981	0.009209748	0.008950339	5.86112944
14	LANSTDQTRA	0.054974422	0.001671831	0.000926373	0.034236323	0.660873468	0.405598939	0.318665934	5.796621863
15	LAHQDTTKKA	0.010689471	0.000159475	0.002362251	0.007046313	0.05116064	0.05761966	0.057049072	5.336940664
16	LAQSEHQRPA	0.129800718	0.032070457	0.277159231	0.362984643	0.646221109	0.521396247	0.677556096	5.219971878
17	LAKQQSASTA	0.303886386	0.082430059	0.093471039	0.208144901	0.194143747	2.314736202	1.548657668	5.096173237
18	LAKGTELKPA	0.013743605	0.000247187	8.11E-05	0.001015147	0.074365316	0.035552681	0.068469844	4.981941904
19	LAIIDATKNA	0.030541345	0.00965888	0.101391528	0.360416919	0.15883095	0.094476272	0.139723905	4.574909966
20	PAHQDTTKNA	0.006108269	0.000101001	0.00137798	0.003144965	0.068954979	0.036451336	0.027428779	4.490434033
21	VAIEDHTKSA	0.001527067	1.06E-05	4.63E-05	0.005852023	0.026407964	0.022149209	0.00664925	4.354261437
22	LAANQPSKPA	0.018324807	0.005982977	0.032689388	0.095981133	0.033596202	0.083175535	0.079731233	4.350999818
23	QAHQDTTKNA	0.001527067	1.86E-05	0.000254753	0.000796194	0.010238259	0.00672523	0.006604424	4.324906866
24	LAVDGAQRSA	0.21226235	0.025388448	0.254382034	0.564680376	1.45459907	1.469941539	0.909726991	4.285861297
25	LAQADTTKNA	0.041230816	0.003489848	0.033534704	0.058281368	0.272156826	0.328602393	0.175156688	4.248198418
26	LAPDQTTRNA	0.024433076	0.005613526	0.000127376	0.032683746	0.14141978	0.056650522	0.103758186	4.246628008
27	LAKSDQSNPA	0.001527067	1.33E-05	0.000173695	9.95E-05	0.000275881	0.005832449	0.006280677	4.112901627
28	LTHQDTTKNA	0.007635336	6.11E-05	0.00101901	0.002965821	0.011617665	0.027241588	0.030850528	4.040493684
29	LAHQDTTKND	0.009162404	9.57E-05	0.001435878	0.010370421	0.037994975	0.040275026	0.036981784	4.036253579
30	LAHETSPRPA	0.16186913	0.273087945	0.732019968	0.686060084	0.509215429	0.658279648	0.652049872	4.028253382
31	LATTPIAKPA	0.323738261	0.04488163	0.116306134	0.201516589	0.40341498	1.907028653	1.281656656	3.958928592
32	LAASDSTKAA	0.222951821	0.01395142	0.260901384	0.376898125	3.160112112	1.397173958	0.847826706	3.802735059
33	LAIEDHTKSA	0.186302207	0.033208047	0.336064976	1.776904973	1.500701888	0.955035579	0.70004899	3.757599027
34	LPPQDTTKNA	0.003054135	3.46E-05	0.000590563	0.001592387	0.005440991	0.011582668	0.011470579	3.755754341
35	IARGVAPSSA	0.004581202	7.44E-05	4.63E-05	0.000199048	0.000183921	0.017626565	0.017058938	3.723681753
36	LAKDKQASNA	0.058028556	0.038911942	0.04363217	0.233165283	0.312956593	0.221462718	0.212148434	3.655931631
37	LAAQGTAKPA	0.035122547	0.001950913	0.010282741	0.044606744	0.13315867	0.209979901	0.11855588	3.375492102
38	LAHQDTTKM	0.010689471	3.72E-05	0.000521085	0.002249247	0.137496136	0.014807254	0.035985642	3.366456809
39	LASKDITKTA	0.076353363	0.002942317	0.032689388	0.10008153	0.342613824	0.420506044	0.248552456	3.255291513
40	LAVDGSQRSA	0.001527067	2.92E-05	4.63E-05	0.000258763	0.000781663	0.000446391	0.004935885	3.232264483
41	LANKASISNA	0.003054135	0.002594129	0.060885867	0.027966299	0.087852843	0.014642794	0.009846867	3.224110435
42	LAHQDTTKKC	0.007635336	2.13E-05	0.000162115	0.002448295	0.044846025	0.007735483	0.023100541	3.025477835
43	LAKQSQTTPA	0.018324807	0.022541816	0.000138956	0.015525774	0.065337869	0.05440682	0.053462959	2.917518244
44	LAHQDTTKNV	0.006108269	3.99E-05	0.000567403	0.002965821	0.052509393	0.014243392	0.01721832	2.818854267
45	LASKTTQASA	0.068718027	0.272715836	0.109508872	0.262903113	0.230038959	0.359227154	0.187782792	2.732656924
46	LATSDSTKAA	0.001527067	1.86E-05	0.000104217	0.002189532	0.004444753	0.002502138	0.004089164	2.677789243
47	LAKSPNKDNA	0.196991677	0.300717019	0.375018962	0.805110926	0.350599053	0.90244369	0.522999632	2.654932626
48	LVHQDTTKNA	0.004581202	1.86E-05	0.000463187	0.001054956	0.007402813	0.012011438	0.012133014	2.648434672
49	LAISDQTNHA	0.001527067	0.000164791	0.001470617	0.007902221	0.004214852	0.001221701	0.003999511	2.6190801
50	LPISDQTKHA	0.001527067	0.000106317	0.001435878	0.006628311	0.002881426	0.001251069	0.003924801	2.570155815

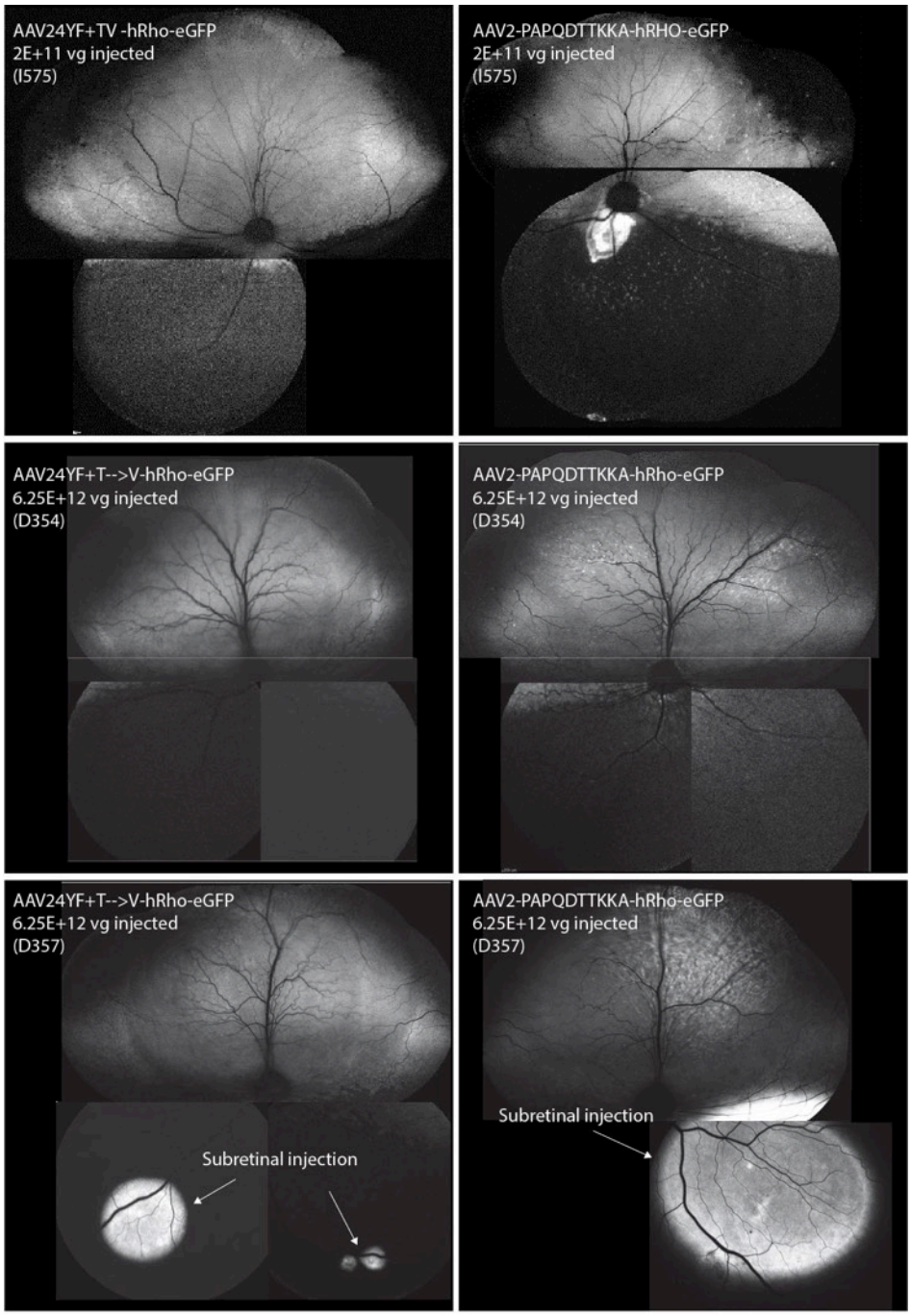
Supplementary Table 2. Top 50 variants from canine screen, ranked by fold increase in the final round of selection. 'Plasmid', 'library', 'r1', 'r2', 'r3', 'r4', and 'r5' show the percent of the total library for the indicated variant in each round. 'r5_v_plasmid' indicates the fold increase in the percent representation of the variant in the final round of selection relative to the original library.



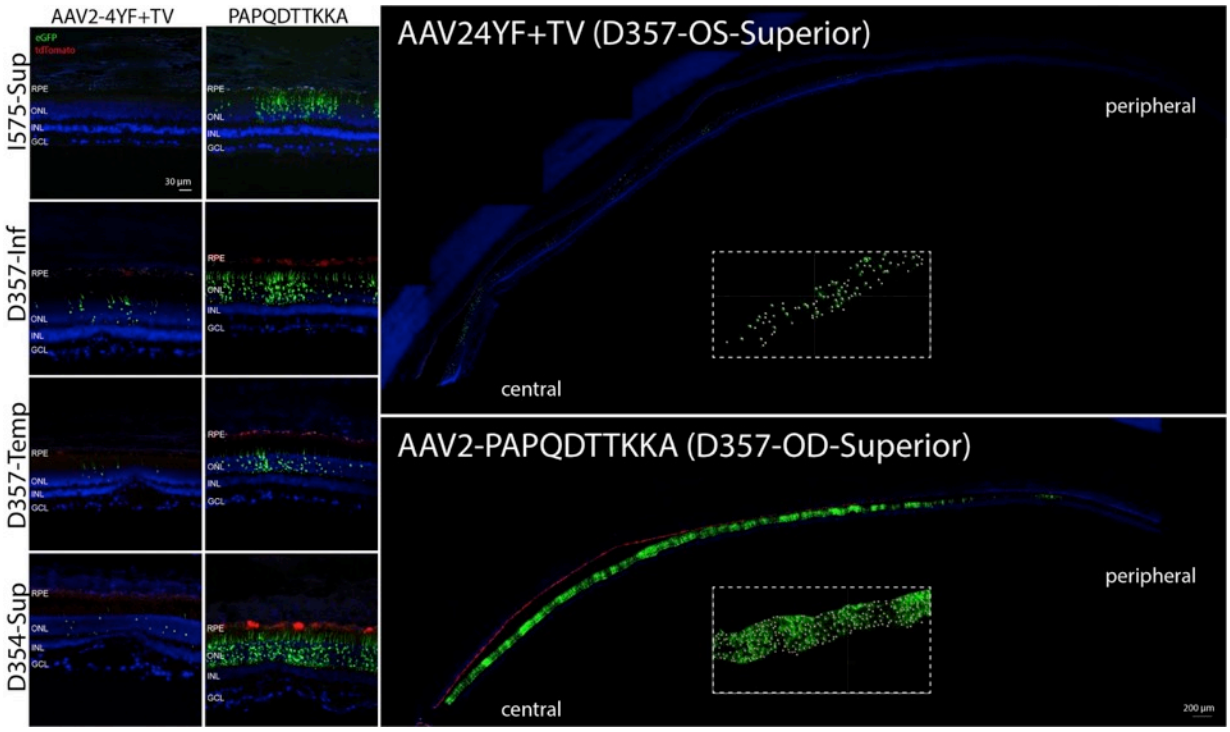
Supplementary Figure 2. Sequence logo (http://weblogo.berkeley.edu/) illustrating a consensus sequence of the top 100 canine-derived AAV2-7mer variants. LA...A (in black) are linkers present in the library. The height of the stack indicates the amino acid sequence conservation at that position, and the height of symbols within the stack indicates the relative frequency of each amino acid at that position.



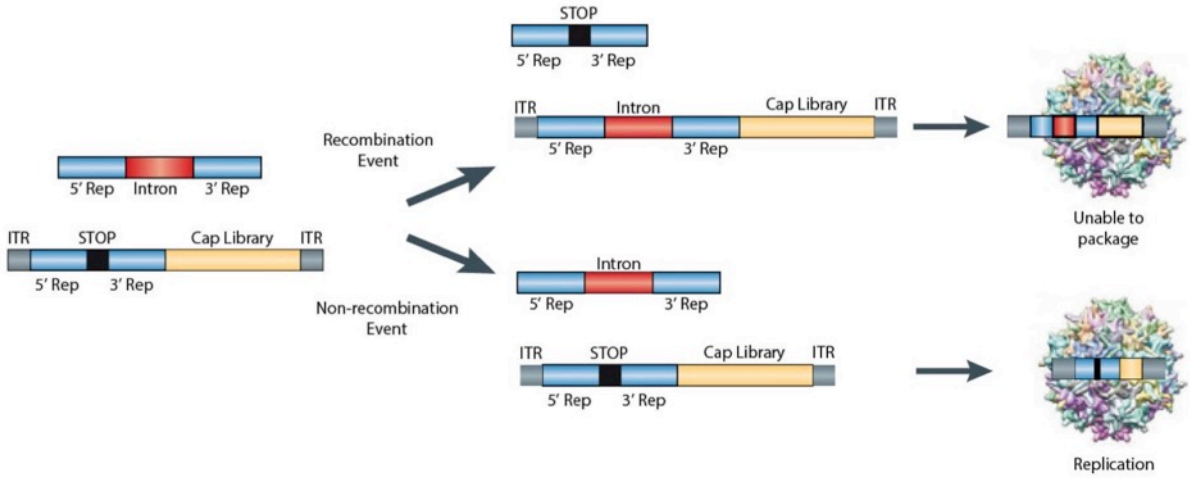
Supplementary Figure 3. Transverse cryosectioning of mouse retinal punch to isolate the outer retina. A punch of retina was flatmounted, embedded in cutting medium, and flash frozen, then mounted in a cryostat and sectioned at 20 µm sections through retinal layers. Layers were then stained for PKC-alpha (a marker of bipolar cells in the inner nuclear layer) and cone arrestin (a marker of photoreceptors in the outer retina). Labeling shows successful isolation of outer retinal tissue from inner retinal cells, which was then used for amplification of libraries. RPE was peeled away prior to sectioning.



Supplementary Figure 4. GFP fundus imaging in canine retinas injected with AAV24YF+TV or AAV2-PAPQDTTKKA-hRho-eGFP. Top row: a canine injected intravitreally with 2E+11 vg showed no fluorescence in the eye treated with AAV24YF+TV. The contralateral eye, injected with AAV2-PAPQDTTKKA showed low levels of GFP expression in superior and inferior retina. In two dogs injected intravitreally with 6.25E+12 vg, AAV24YF+TV retinas had no GFP expression visible by fundus imaging, while AAV2-PAPQDTTKKA-injected eyes had expression in superior and inferior retina. In contrast, subretinal injection of both viruses resulted in bright GFP expression (lower images).



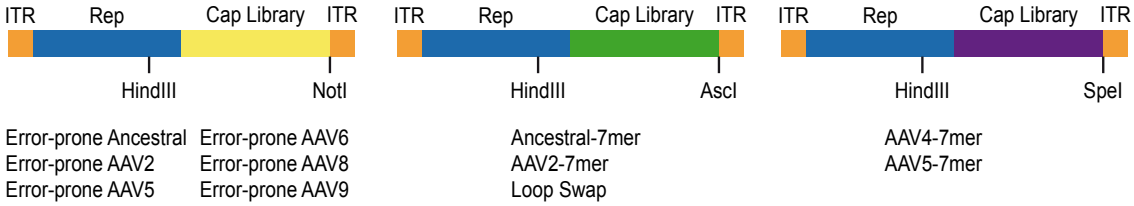
Supplementary Figure 5. Native fluorophore expression in canine retinas injected with AAV24YF+TV or AAV2-LAHQDTTKNA. Dogs were injected intravitreally with a mixture of viruses encoding hVMD2-tdTomato driving expression in RPE cells and hRho-eGFP driving expression in rods. Images show expression from 3 dogs injected with either 5E+11 (dog I575) or 6E+12 (D357 and D354) viral particles. Quantification of labeled cells in a montage created from a section of the superior quadrant, performed using Imaris software (insets), revealed that AAV24YF+TV injection resulted in 6,855 GFP-expressing rods, while injection of AAV2-PAPQDTTKKA resulted in 29,371 GFP-expressing rods.



Supplementary Figure 6. Construction of replication incompetent libraries. In order to mitigate the possibility of replication of the AAV libraries in the primate retina a *rep* in *trans* strategy was utilized where the *rep* sequence in the library transgene is mutated (pRep-SafeStop) with stop codons preventing the expression of the four Rep proteins and the full *rep* sequence is supplied on pRepIntronHelper. The ITRs in AAV are highly recombinogenic leading to the possibility of recombination and subsequent replication in the presence of a helper virus when using this strategy, so an intron was inserted in the *rep* supplied in *trans*. Schematic depicting both recombination and non-recombination events with the pRepIntronHelper. If recombination event does occur then the intron sequence (Nebulin intron 8, 774 bp) will push the transgene over the packaging capacity of AAV leading to incomplete packaging. If recombination does not occur then the mutated rep sequence will be packaged, mitigating the possibility of replicating.

Round	NHP ID	Age/Weight	Libraries injected	Amount Virus Injected	Notes
1	V002278	4-10 years 5.98 kg	AAV2 error-prone AAV5 error-prone AAV6 error-prone AAV8 error-prone AAV9 error-prone Loop Swap Ancestral-7mer	2.0E+11 vg per library	No error-prone libraries were recovered
2	V002262	4-10 years 6.35 kg	Recovered variants from round 1 AAV2-7mer	2.5E+11 vg ONL 2.5E+11 vg RPE 5.0E+10 vg AAV2-7mer	No variants were recovered from this round
3	V002148	4-10 years 6.48 kg	Recovered variants from round 1 AAV2-7mer	1.3E+11 vg ONL 1.3E+11 vg RPE 5.0E+10 vg AAV2-7mer	Repeat of previous round
4	V002265	8 years 5 months 4.92 kg	Recovered variants from round 3	4.3E+12 vg ONL 3.7E+12 vg RPE	Error prone PCR conducted No adverse events
5	V002540	6 years 9 months 4.59 kg	Recovered variants from round 4 AAV4-7mer AAV5-7mer	~1E+12 vg per library	No adverse events
6	V002861	6 years 6 months 6.60 kg	Recovered variants from round 5	2.4E+12 vg ONL 6.3E+12 vg RPE	No adverse events
GFP-barcode	V002361	9 years 5 months 6.00 kg	Barcoded individual variants	~1.0E+10 vg each variant	Hyphema in left eye resolved in 12 days
Variant validation	106	9 years 5 months 14.5 kg	7m8 and NHP9	7m8 = 3.4E+12 vg NHP9 = 7.0E+12 vg	No adverse events

Supplementary Table 3. Summary of the rounds of selection performed in primates, indicating the age and weight of the primates injected, the virus and titer injected at each round, and notes on the rounds of selection completed. The third round was a repeat of the 2nd round of selection, which did not result in PCR amplification of any variants.



Supplementary Figure 7. Backbones used. Cloning of libraries in individual backbones. In order to better evaluate groups of libraries and to be able to tune representation of the library groups, different restriction sites were used for cloning 3' of cap. The restriction sites allowed for separate cloning and also acted as handles for unique amplification primers. *Not*I was used for the error-prone libraries and *Asc*I was used for the initial 7mer and Loop Swap libraries. The AAV4-7mer and AAV5-7mer libraries, which were introduced in later rounds, were cloned into the *Spe*I backbone.

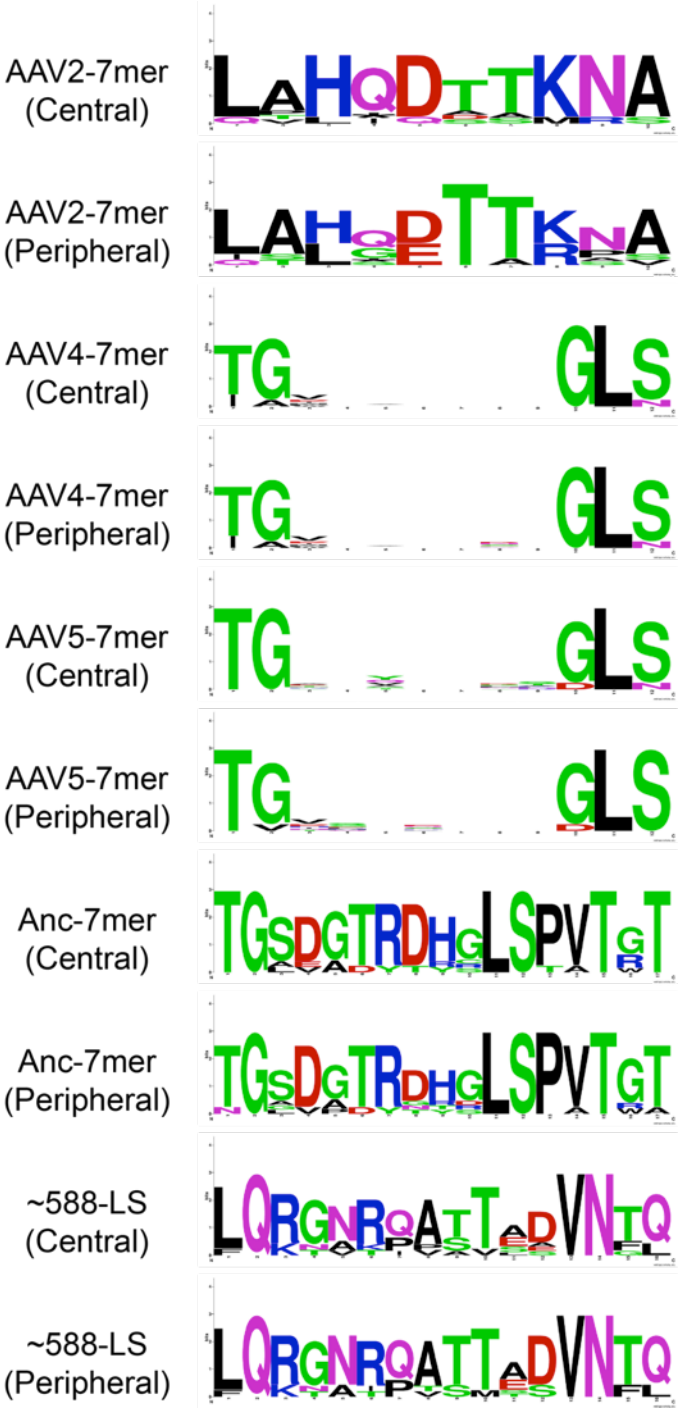
Central ONL

Rank	AAV2-7mer	Fold change	AAV4-7mer	Fold change	AAV5-7mer	Fold change	Ancestral-7mer	Fold change	Loop Swap 588	Fold change
1	LTHQDTTKNA	341	TGVMRSTNSGLN	702	TGASYLDNSGLS	36	TGLDATRDHGLSPVTGT	557	LQKNARPASTESVNFQ	13272
2	QAHQDTTKNA	329	TGEVDLAGGGLS	503	TGMHVTMMAGLN	31	TGSDGTRDHGLSPVTWT	475	LQRGVRIPSVLEVNGQ	5003
3	LAH*DTTKNA	189	TGWLTPDDMGLS	488	TGERTAYDKGLS	29	TGSDGTRDHRLSPVTGT	438	LQKTARPASTEAVNFQ	3622
4	LAHQDTAKNA	174	IGVLHQSDLGLS	480	TGSGQMGGSGLS	27	TGAVADYTRGLSPATRT	368	LQRGNRPVTTADVNTQ	1808
5	LAHQDTTKNS	173	IGMAAHDSVGLS	455	TGTVYSDQTGLS	25	TGSDGTRDHGLSTVTGT	367	LQRGNKQATTADVNTQ	1734
6	LVHQDTTKNA	155	TAVPGLENGLS	434	TGRAQMELSGLS	22	TGSDGTRDHGLSPVTRT	363	LQRGNRQATTEDVNTQ	1674
7	LALIQDSMRA	150	TGPEQIISNGLS	433	TGAEAFSGTDLS	21	TGSDGTRDHSLSPVTGT	341	LQRGNTQATTADVNTQ	1376
8	LAHQDATKNA	149	TGDKATLESGLS	414	TGFYGESLNGLS	21	TGSDGTRDYGLSPVTGT	317	LQRGNRQAATSDVNTL	1192
9	LAHQDSTKNA	148	TGISAETQAGLS	397	TGTQYNQSLGLS	21	TGSDGTRDHCLSPVTGT	301	FQRGNRQATTADVNTL	1156
10	LPHQDTTKNA	143	TGVMEVKPDGLS	375	TGELVANDRGLS	21	TGSEGTRDHGLSPVTRT	294	LQRGNRQATTSDVNTQ	1114

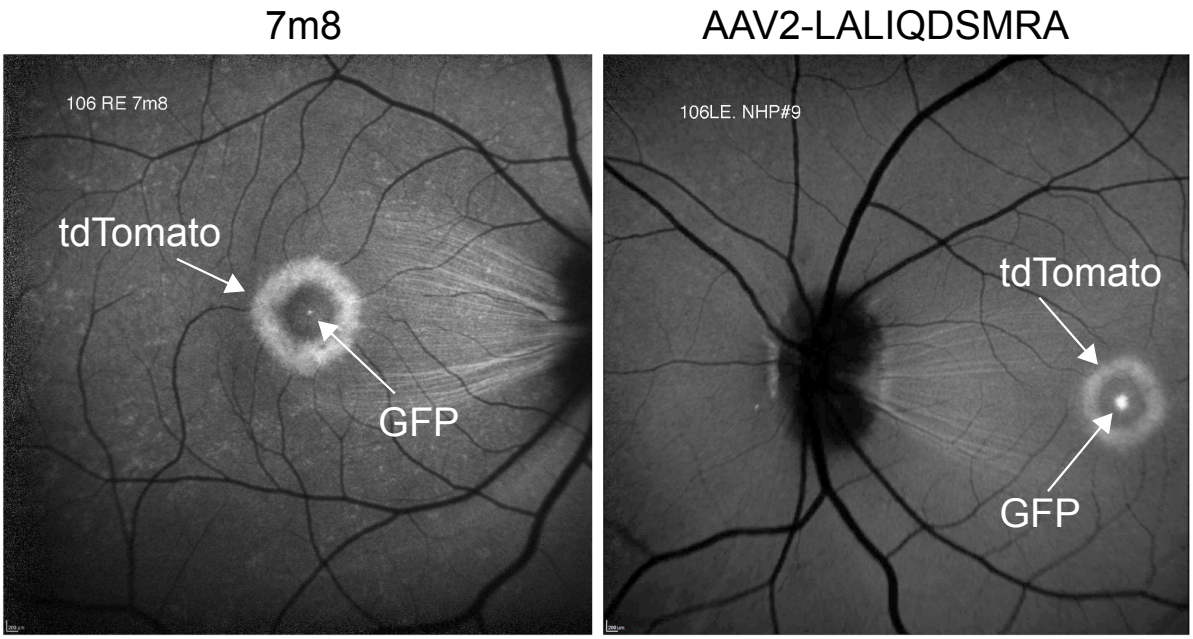
Peripheral ONL

Rank	AAV2-7mer	Fold change	AAV4-7mer	Fold change	AAV5-7mer	Fold change	Ancestral-7mer	Fold change	Loop Swap 588	Fold change
1	QAHQDTTKNA	357	TGPVEGPSSGLS	1703	TVVSTQAGIGLS	40	TGSDGTRNHGLSPVTGT	766	LQRGNRPVTTADVNTQ	10425
2	LTHQDTTKNA	251	TGGHDSSLDGLS	1581	TGVMHSQASGLS	39	TGGDPTRGTGLSPVTGA	645	LQKNARPASTESVNFQ	10277
3	LSLGETTRPA	228	TSPYSGSSDGLS	1569	TGTSAHRDTGLS	36	NGAVADYTRGLSPATGT	630	LQKTARPASMESVNFQ	6906
4	LAHQDTTKNV	223	TGYDEVLSEGLS	1243	TGVGSSIPGGLS	35	TGLDATRDHGLSPVTGT	623	LQRGNRQTTTADVNTQ	4387
5	IALSETTRPA	222	TGPMLSLEKGLS	1150	TGEGHESTSGLS	35	TGSDGTRDHGLSPVTWT	499	LQRGNTQATTADVNTQ	3991
6	LAHQDTAKNA	215	TGYNQALHSGLS	1130	TGDHDDRSQGLS	33	TGSDGTRDHRLSPVTGT	469	LQRGNRQATTSDVNTQ	3962
7	LALGETTRSA	192	TGGQMPNNEGLS	1107	TGVELPMREDLS	33	TGSDGTRDYGLSPVTGT	393	LQRGNRQATTTDVNTL	3538
8	LAH*DTTKNA	189	TGSLEIRDDGLS	1102	TGIFNEGFPGLS	30	TGSDGTRDHGLSPVTRT	391	LQRGNRQATTEDVNTQ	3217
9	LAHQDTTKNS	176	TGSSPVEMPGLS	1024	TGNQEKANVGLS	29	TGSDGTRDHSLSPVTGT	385	FQRGNRQATTADVNTL	3085
10	LALGETTRAA	167	TGGTQLSLEGLS	1000	TGHSRQNLDGLS	28	TGSDGTRDHDLSPVTGT	342	LQRGNIQATTADVNTQ	2236

Supplementary Table 4. For each library tested, the top 10 variants that were recovered from either the central (top) or peripheral (bottom) retinal regions are listed, along with their respective fold changes in the percent of the library from the final round of selection relative to the original library (% of total in round 5 / % of total in library).



Supplementary Figure 8. Sequence logos (http://weblogo.berkeley.edu/) illustrating consensus sequences of the top 10 primate-derived variants collected from central or peripheral areas from each library tested. LA...A are linkers present in the AAV2-7mer library. TG…GLS are linkers present in the AAV4, AAV5 and Anc-7mer libraries. The height of the stack indicates the amino acid sequence conservation at that position, and the height of symbols within the stack indicates the relative frequency of each amino acid at that position.



Supplementary Figure 9. Fundus imaging of a primate injected intravitreally with 7m8 in the one eye and AAV2-LALIQDSMRA in the contralateral eye. Both viruses were packaged with SNCG-tdTomato (driving expression in retinal ganglion cells) or PR2.1-eGFP (driving expression in cones), which were then mixed in equal titers before injection. AAV2-LALIQDSMRA resulted in reduced RGC expression and increased cone expression compared to 7m8.

Primer	Sequence
SDM1	GACCTTAATCACAATCTTTTAAAACCCCGGCATGGCGGCT
SDM2	GGCTCGTGGACAAGTAAAGGGATTACCTCGGA
Neb Genomic_F	GTAAGGGTCTGCTCCATTGCCACTT
Neb Genomic_R	CTAAATCAAAAAAGAGTGAAAAGTTAGGAGG
IFA_F	TGGCTCGTGGACAAGGTAAGGGTCTGCTCCATTGC
IFA_R	CTCCGAGGTAATCCCCTAAATCAAAAAAGAGTGAAAAGTT
HindIII_F1	GACGTCAGACGCGGAAGCTTC
NotI_R1	GGTTTATTGATTAACAAGCGGCCG
AscI_R1	TGGCGGACTTTATAGGCGCG
SpeI_R1	GCCCAGTTCGAATAGCGAGT
LS588_Forward adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNGT TCTGTATCTACCAACCTCCA**
LS588_rev_index1	CAAGCAGAAGACGGCATACGAGAT**CGTGAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGACCATGCCTGGAAGAACGCC
LS453_Forward adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNAT CGACCAGTACCTGTATTACT**
LS453_rev_index1	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGGTCCCGAATGTCACTCGCTC
Anc_Forward adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNCC TGCAGTCGTCTAACACCGC**
Anc_rev_index1	CAAGCAGAAGACGGCATACGAGAT**CTCTAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNTAAGGCTCCCTGGCTGTTGAC
5_Forward adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNAT GGCCACCAACAACCAGAGC**
5_rev_index1	CAAGCAGAAGACGGCATACGAGAT**ATCAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGAGGTTGTACGTGCCGGTCGC
4_Forward adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNAC CTACCTGGCGGTGACCAGA**
4_rev_index1	CAAGCAGAAGACGGCATACGAGAT**TAGTTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNTCCACGGTCGGCAGGTTGCTG
LS588_rev_index2	CAAGCAGAAGACGGCATACGAGAT**ACATCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGACCATGCCTGGAAGAACGCC
LS588_rev_index3	CAAGCAGAAGACGGCATACGAGAT**GCCTAA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGACCATGCCTGGAAGAACGCC
LS588_rev_index4	CAAGCAGAAGACGGCATACGAGAT**TGGTCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGACCATGCCTGGAAGAACGCC
LS588_rev_index5	CAAGCAGAAGACGGCATACGAGAT**CACTGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGACCATGCCTGGAAGAACGCC
LS588_rev_index6	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGACCATGCCTGGAAGAACGCC
LS588_rev_index7	CAAGCAGAAGACGGCATACGAGAT**GATCTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGACCATGCCTGGAAGAACGCC
LS588_rev_index8	CAAGCAGAAGACGGCATACGAGAT**TCAAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGACCATGCCTGGAAGAACGCC
LS453_rev_index2	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGGTCCCGAATGTCACTCGCTC
LS453_rev_index3	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGGTCCCGAATGTCACTCGCTC
LS453_rev_index4	CAAGCAGAAGACGGCATACGAGAT**TACAAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGGTCCCGAATGTCACTCGCTC
LS453_rev_index5	CAAGCAGAAGACGGCATACGAGAT**TTGACT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGGTCCCGAATGTCACTCGCTC
LS453_rev_index6	CAAGCAGAAGACGGCATACGAGATGGAACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGGTCCCGAATGTCACTCGCTC
LS453_rev_index7	CAAGCAGAAGACGGCATACGAGAT**TGACAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGGTCCCGAATGTCACTCGCTC
LS453_rev_index8	CAAGCAGAAGACGGCATACGAGAT**GGACGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGGTCCCGAATGTCACTCGCTC
Anc_rev_index2	CAAGCAGAAGACGGCATACGAGAT**GCGGAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNTAAGGCTCCCTGGCTGTTGAC
Anc_rev_index3	CAAGCAGAAGACGGCATACGAGAT**TTTCAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNTAAGGCTCCCTGGCTGTTGAC
Anc_rev_index4	CAAGCAGAAGACGGCATACGAGAT**GGCCAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNTAAGGCTCCCTGGCTGTTGAC
Anc_rev_index5	CAAGCAGAAGACGGCATACGAGAT**CGAAAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNTAAGGCTCCCTGGCTGTTGAC
Anc_rev_index6	CAAGCAGAAGACGGCATACGAGATCGTACGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN

	NNNNTAAGGCTCCCTGGCTGTTGAC
Anc_rev_index7	CAAGCAGAAGACGGCATACGAGAT**CCACTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNTAAGGCTCCCTGGCTGTTGAC
Anc_rev_index8	CAAGCAGAAGACGGCATACGAGAT**GCTACC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNTAAGGCTCCCTGGCTGTTGAC
5_rev_index2	CAAGCAGAAGACGGCATACGAGAT**GCTCAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGAGGTTGTACGTGCCGGTCGC
5_rev_index3	CAAGCAGAAGACGGCATACGAGAT**AGGAAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGAGGTTGTACGTGCCGGTCGC
5_rev_index4	CAAGCAGAAGACGGCATACGAGAT**CTTTTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGAGGTTGTACGTGCCGGTCGC
4_rev_index2	CAAGCAGAAGACGGCATACGAGAT**CCGGTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNTCCACGGTCGGCAGGTTGCTG
4_rev_index3	CAAGCAGAAGACGGCATACGAGAT**ATCGTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNTCCACGGTCGGCAGGTTGCTG
4_rev_index4	CAAGCAGAAGACGGCATACGAGAT**TGAGTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNTCCACGGTCGGCAGGTTGCTG
2-7mer Forward adapter	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNTC TACCAACCTCCAGAGAGG**
rev_index1	CAAGCAGAAGACGGCATACGAGAT**CGTGAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGTTGACATCTGCGGTAGCTG
rev_index2	CAAGCAGAAGACGGCATACGAGAT**ACATCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGTTGACATCTGCGGTAGCTG
rev_index3	CAAGCAGAAGACGGCATACGAGAT**GCCTAA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGTTGACATCTGCGGTAGCTG
rev_index4	CAAGCAGAAGACGGCATACGAGAT**TGGTCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGTTGACATCTGCGGTAGCTG
rev_index5	CAAGCAGAAGACGGCATACGAGAT**CACTGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGTTGACATCTGCGGTAGCTG
rev_index6	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGTTGACATCTGCGGTAGCTG
rev_index7	CAAGCAGAAGACGGCATACGAGAT**GATCTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGTTGACATCTGCGGTAGCTG
rev_index8	CAAGCAGAAGACGGCATACGAGAT**TCAAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGTTGACATCTGCGGTAGCTG
rev_index9	CAAGCAGAAGACGGCATACGAGAT**CTGATC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNGTTGACATCTGCGGTAGCTG
rev_index10	CAAGCAGAAGACGGCATACGAGAT**AAGCTA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGTTGACATCTGCGGTAGCTG
rev_index11	CAAGCAGAAGACGGCATACGAGAT**GTAGCC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGTTGACATCTGCGGTAGCTG
rev_index12	CAAGCAGAAGACGGCATACGAGAT**TACAAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT NNNNNGTTGACATCTGCGGTAGCTG
F_adapter_GFPBC	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNGG CCATCAAGCTTATCGATACC
R_adapter_GFPBC	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNCTGATCAGCGAGCTCTAGTCG

Supplementary Table 5. Primers used in the study.

Appendix B: Supplying rep in trans to generate replication-safe AAV libraries for in vivo directed evolution in primate retina

Background

Adeno-associated virus is the vector of choice for gene therapy of the retina. Significant improvements in AAV as a gene delivery vehicle have been engineered through the process of directed evolution, which has been discussed in detail in Chapters 1-4 of this dissertation. For gene therapy applications, the viral genome is replaced by the transgene of choice but when AAV is undergoes directed evolution the wild-type AAV genome is present. This is necessary because the packaged *cap* sequence acts as a barcode for the successful capsid variants. AAV is a *Dependovirus*, so when in the presence of a helper virus, such as herpes simplex virus or adenovirus, the AAV libraries for directed evolution could replicate[1]. Rhesus macaques and other primates are carriers for the helper virus herpes B[2] raising the question of the safety for directed evolution in primates as compared to the relative safety of rodent models[3,4]. It is possible that replication could occur in the retina due to the link between the trabecular meshwork and the general blood system leading to an immune response that would be dangerous for the primate and confound the selection pressure in the directed evolution process. In order to overcome this challenge, we initially deleted 80 bp in the *rep* sequence to create a replication-safe backbone called pRepSafeΔ80. A previous study developing *rep*-expressing cell lines showed that *rep* could be supplied in *trans* and allow for AAV packaging[5]. During packaging, an additional pRepHelper plasmid was used but when libraries were challenged with adenovirus, replication still occurred likely because there was recombination between the pRepSafeΔ80 and the pRepHelper plasmids. ITRs are highly recombinogenic and could lead to a higher likelihood of recombination between transfected plasmids. A pRepIntronHelper plasmid was then created in order to reduce the viability of library variants that underwent a recombination event. Because if the *rep* sequence containing the intron recombined with the mutated *rep* sequence on the library backbone then the sequence would be too large for AAV to package rendering it inviable. A pRepSafeStop plasmid was created which introduced a stop codon at the same as site the intron insertion and a second stop codon at the beginning of the *rep* sequence to completely eliminate *rep* expression. This resulted in a system that reduced the ability of AAV libraries to replicate in the presence of adenovirus by 10-fold compared to wild-type AAV while retaining similar titers. This system was then used to increase safety in the primate directed evolution study.

Methods and Results

The pRepSafeΔ80 construct was created by digesting pSub2Cap2[6] with *Sgra*I and *BamH*I, Klenow reaction, and blunt-end ligation. This resulted in an 80 bp deletion and frameshift in the coding sequence that led expression of truncated, non-functional Rep78 and Rep68 while eliminating the start sites for Rep52 and Rep40 (**Fig. 1A**). The frameshift further prevents the possibility of expression by introducing additional stop codons. Eliminating the *cap* sequence in the packaging plasmid pAAV2/rh10 was accomplished by digesting with *Pme*I and *Bsm*I, Klenow reaction, and blunt-end ligation generating the pRepHelper plasmid (**Fig. 1A**). Triple transfection as previously described was used to package the AAV2 *cap* sequence in

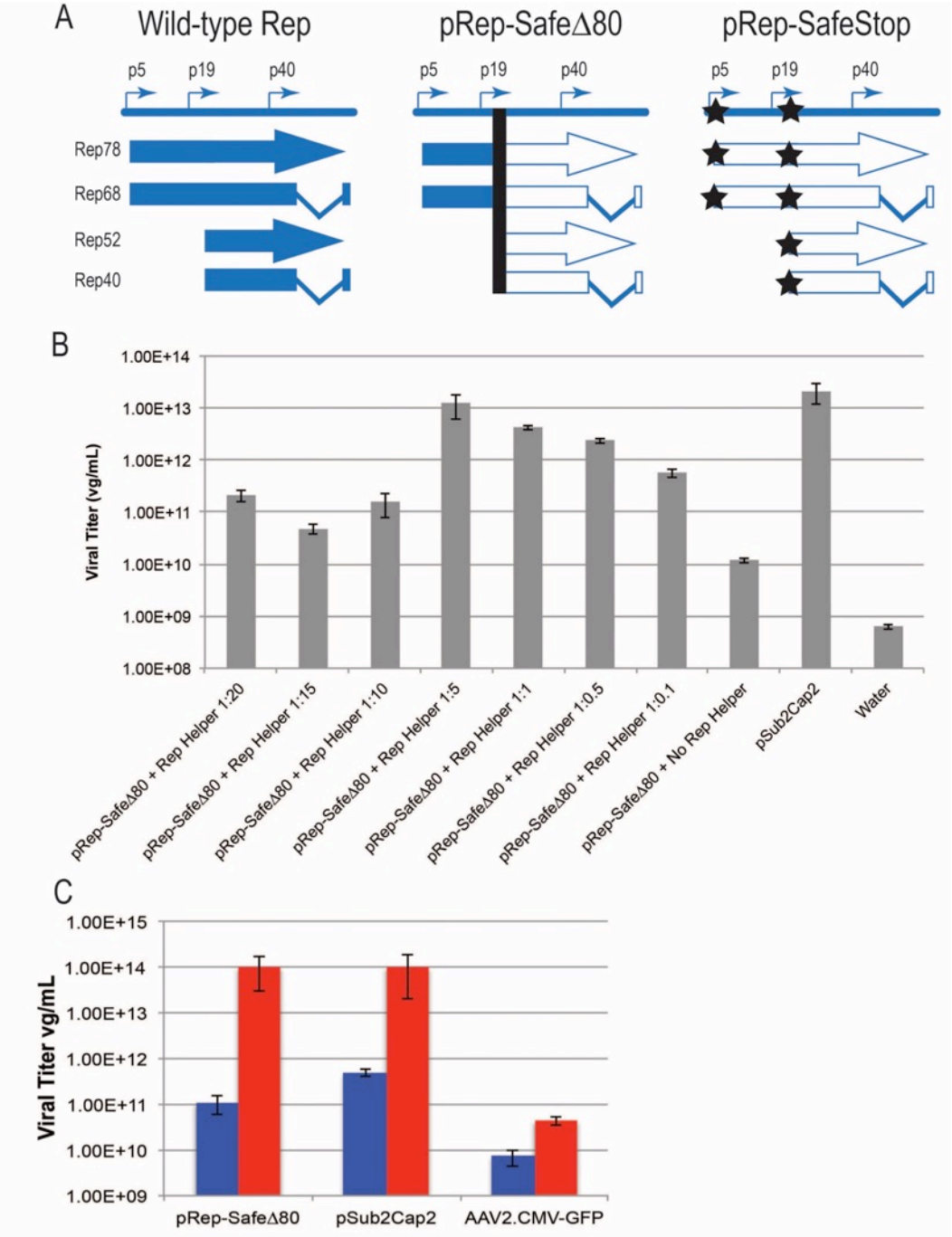


Figure 1. *Rep* in *trans* system using pRep-SafeΔ80 and pRepHelper. A) Schematic of expression of *rep*. Wild-type *rep* leads to expression of all four Rep proteins. The pRepSafeΔ80 construct has an 80 bp deletion indicated by the black bar in the *rep* sequence leading to the truncated expression of the larger Rep proteins and prevents expression of the smaller Rep proteins by eliminating the start codon. The pRepSafeStop construct utilizes two stop codons indicated by the black stars to completely prevent the expression of all four Rep proteins. **B)** Titering of the pRepSafeΔ80 with an AAV2 *cap* sequence with different ratios of pRepHelper compared to controls. A 5:1 pRepHelper to pRepSafeΔ80 ratio was the optimal ratio of pRepHelper to yield a titer similar to the pSub2Cap2 control. Without pRepHelper there is little packaging. **C)** An adenovirus rescue experiment with the pRepSafeΔ80 resulted in a similar amount of replication as the positive control demonstrating that the *rep* in *trans* system using pRepSafeΔ80 and pRepHelper does not decrease replication in the presence of a helper virus. Blue bars indicate no adenovirus infection while red bars indicate adenovirus infection.

pRepSafeΔ80 with different ratios of RepHelper[7]. The system was able to package virus and the 5:1 pRepHelper to pRepSafeΔ80 was identified as the optimal ratio of pRepHelper to yield a similar titer of 2E+13 vg/mL to pSub2Cap2 which contains *rep*. When pRepHelper was not used, the titer was significantly reduced showing that it is necessary for packaging (**Fig. 1B**). To determine if the possibility of replication had been reduced a rescue experiment using human adenovirus 5 (Ad5) was performed. A 6-well plate of HEK293T cells were infected with AAV at an MOI of 10^5 and then incubated at 37º C and 5% CO_2 for 48 hours. Next, 10 µL of Ad5 was infected and the plates were incubated at 37º C and 5% CO_2 for 48 hours. This infection level of Ad5 led to an observable cytopatheic effect in the cells that were then harvested, spun at 1000 rpm for 3 min, and the pellet was resuspended in 100 µL of lysis buffer (0.15M NaCl, 50 mM Tris HCl, 0.05% Tween, pH 8.5). Finally, 5 µL of the crude lysate was used for titering as previously described[7]. The titering showed that the pRepSafeΔ80 was as replication competent as the pSub2Cap2 control (**Fig. 1C**). This is likely due to recombination between *rep* sequences of the pRepSafeΔ80 and the pRepHelper plasmids during packaging. This necessitated modifications in the system in order to significantly reduce the replication ability of the libraries.

Table 1.

pSub2xTAAA SDM1 Fwd:	gaccttaatcacaatcttttaaaaccccggcatggcggct
pSub2xTAAA SDM1 Rev:	agccgccatgccggggttttaaaagattgtgattaaggtc
pSub2xTAAA SDM2 Fwd:	ggctcgtggacaagtaaagggattacctcgga
pSub2xTAAA SDM2 Rev:	tccgaggtaatccctttacttgtccacgagcc
NI8 Genomic Fwd:	gtaagggtctgctccattgccactt
NI8 Genomic Rev:	ctaaatcaaaaaagagtgaaaagttaggagg
NI8 IFA Fwd:	tggctcgtggacaaggtaagggtctgctccattgc
NI8 IFA Rev:	ctccgaggtaatcccctaaatcaaaaaagagtgaaaagtt

The pRepIntronHelper was developed so that *rep* could be supplied in *trans* but in the case of a recombination event the intron would prohibit proper packaging of the transgene because it would be larger than the packaging capacity of AAV. Along with the new helper plasmid, the pRepSafeStop plasmid was developed to ensure strong termination of *rep* as well as position the stop codon in between the same bases as the intron (**Fig. 1A**). The pRepSafeStop plasmid was created by Quikchange site-directed mutagenesis using pSub2xTAAA SDM primers (**Table 1**). This created stop codons at 5th amino acid position and the 235th amino acid position. The first codon maintained the reading frame while the second stop codon inserted two bases leading to a frameshift mutation. The TAAA sequence was used as the stop codon sequence because evidence has demonstrated that TAA is the strongest stop codon and having A following the stop codon leads to the highest likelihood of a termination event[8]. To create the pRepIntronHelper, the first AGG sequence after the Rep52/Rep40 start codon was used as the site for intron insertion based on a computational analysis of splice signal motifs[9]. Nebulin encodes for a large protein in muscle that is important for muscle stabilization, regulating the precise assembly of myosin and actin filaments, and the sequence contains many large introns[10] so it was ideal for identifying larger intron sequences. In particular, intron 8 from the nebulin sequence was used to increase the size of the *rep* sequence by 774 bp. In a recombination event the packaged genome would be nearly 5.5 kb and would not appropriately package. This would significantly reduce the ability of AAV to replicate, increasing the overall safety profile. Nebulin intron 8 was amplified from genomic DNA extracted from HEK293T cells, Taq polymerase PCR

using the NI8 Genomic primers was performed, and the amplicon was cloned into the pRepHelper plasmid using Infusion Assembly cloning resulting in the pRepIntronHelper using IFA NI8 primers (**Table 1**).

The system utilizing a 5:1 ratio of pRepIntronHelper and pRepSafeStop with an AAV2 *cap* sequence titered within an order of magnitude to the wild-type control pSub2Cap2 with titers of 3E+13 vg/mL and 2E+14 vg/mL, respectively. Again, when the Rep Helper plasmid was not used the titer crashed demonstrating that pRepIntronHelper is effective and required for packaging (**Fig. 2A**). An adenovirus rescue study with pRepIntronHelper and pRepSafeStop with an AAV2 *cap* sequence resulted in no rescue while the wild-type control pSub2Cap2 increased 104-fold (**Fig. 2B**). This demonstrated that the *rep* in *trans* system is capable of significantly reducing the ability of AAV to replicate while still maintaining high packaging efficiencies. Two libraries that were used for primate directed evolution, error-prone AAV2 and 7mer-Ancertral libraries, were cloned into pSubSafeStop and demonstrated a 23- and 20-fold increase, respectively, in an adenovirus rescue study. The wild-type control showed a 243-fold increase (**Fig. 2C-D**). It is possible that truncations of the pRepIntronHelper were packaged and could still provide *rep* expression in the rescue studies especially at the high MOI of 10^5.

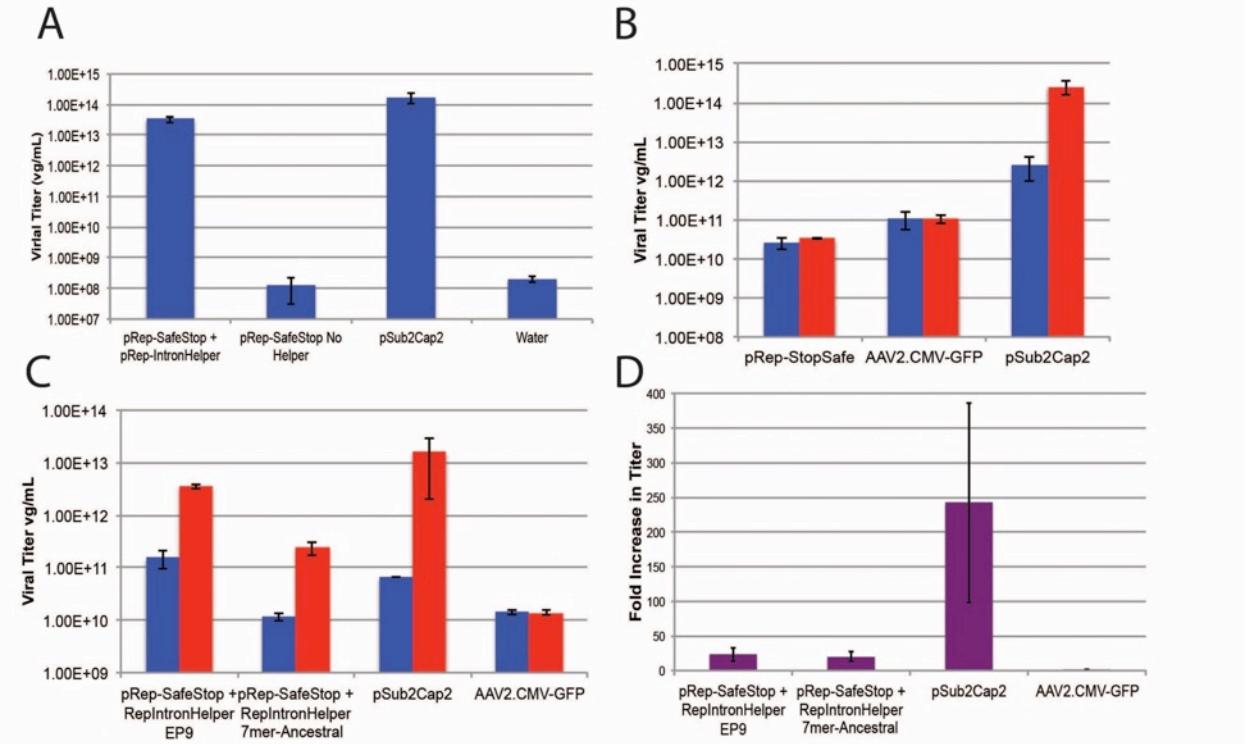


Figure 2. *Rep* in *trans* system using pRep-SafeStop and pRepIntronHelper reduces replication by 10-fold. **A)** Titering of the 5:1 ratio of pRepIntronHelper with pRepSafeSafe and AAV2 *cap* packages at a similar titer as the wild-type control. There is no packaging when pRepSafeSafe is packaged without RepIntronHelper. **B)** Adenovirus rescue with 5:1 ratio of pRepIntronHelper with pRepSafeSafe and AAV2 *cap* shows no replication. The AAV2.CMV-GFP control shows no replication and the pSub2Cap2 control shows a 104-fold increase in titer. Blue bars indicated no adenovirus infection while red bars indicate adenovirus infection. **C)** Adenovirus rescue study with two different libraries packaged for primate directed evolution compared to controls. Blue bars indicated no adenovirus infection while red bars indicate adenovirus infection. **D)** The fold difference in titer demonstrating that the control pSub2Cap2 virus compared two packaged libraries has a 10-fold increase in titer.

Even under these difficult conditions the libraries still demonstrated greater than 10-fold less replication than the control. This demonstrated that the *rep* in *trans* system reduces the likelihood of libraries replicating in the presence of a helper virus leading to safer libraries for the primate directed evolution study.

Table 2.

Nebulin Intron 8 sequence (between exon 8 and 9)
gtaa gggtctgctc cattgccact tctgtcttcc tatgtcatcc catcatatat ttcatgtaat acttggtgaa ctcttgatca tctggaatac aattccccaa tctttttatg agtagccttg agacatgtga gactgcagggatagggcttg gcaaaaagga actgagagga aagctcctag gcagagtggg gagagggtcc tttgggaacc tgagcaacct gggggcagag gggagtgtct gtggagaata aaggtgctct ggagaaccag ctgagagagc cacaaagtaa tatctcaaaa ttgtctgtag acccttggtatactttgga tgtccctggc tattactcta attagaggtt tattttgtgc cccttcttgc aagtgtccta tattgccctg aactatgtga agtgcttcct ggctccttac attgtgttag cacttgtaag tgctattaat aaacaatgac aaataatgac caaaaaaaaa aaaaaaaaaa aaaaacaaag aaggcaggta gagtcacaga gaaaagttga aacaaaatca gacacatctgcccctctgaa aataagaaaa agcagggcat gctatcagta aaaatgactt gggttggatg gaaagtggga agttatctgt gaatttcatt tacctttgct ctcacctggc ccttggttct ttaggttatt aatggatgaa ttcaaaacta tgtattttgt gtaaacgtga ccctacttcc tcctaacttt tcactctttt ttgatttag

References

1. Weitzman, M. D. & Linden, R. M. Adeno-associated virus biology. *Methods Mol. Biol.* **807,** 1–23 (2011).
2. Weigler, B. J. Biology of B virus in macaque and human hosts: a review. *Clin. Infect. Dis.* **14,** 555–567 (1992).
3. Dalkara, D. *et al.* In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous. *Science Translational Medicine* **5,** 189ra76 (2013).
4. Klimczak, R. R., Koerber, J. T., Dalkara, D., Flannery, J. G. & Schaffer, D. V. A novel adeno-associated viral variant for efficient and selective intravitreal transduction of rat Müller cells. *PLoS ONE* **4,** e7467 (2009).
5. Hölscher, C. *et al.* Cell lines inducibly expressing the adeno-associated virus (AAV) rep gene: requirements for productive replication of rep-negative AAV mutants. *J. Virol.* **68,** 7169–7177 (1994).
6. Maheshri, N., Koerber, J. T., Kaspar, B. K. & Schaffer, D. V. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat. Biotechnol.* **24,** 198–204 (2006).
7. Flannery, J. G. & Visel, M. in *Methods in Molecular Biology* **935,** 351–369 (Humana Press, 2012).
8. McCaughan, K. K., Brown, C. M., Dalphin, M. E., Berry, M. J. & Tate, W. P. Translational termination efficiency in mammals is influenced by the base following the stop codon. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 5431–5435 (1995).
9. Lim, L. P. & Burge, C. B. A computational analysis of sequence features involved in recognition of short introns. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 11193–11198 (2001).
10. Trinick, J. Titin and nebulin: protein rulers in muscle? *Trends Biochem. Sci.* **19,** 405–409 (1994).

Appendix C: Exploring the role of the CRB1 extracellular domain in protein-protein interactions

Background

Mutations in the crumbs homologue-1 (*CRB1*) gene, initially identified in *Drosophila*, are associated with Leber congenital amaurosis (LCA) and autosomal recessive retinitis pigmentosa (RP) [1-3]. LCA is the most severe and is characterized by impairment to vision at birth. RP is milder with a later stage of onset. In patients, mutations in *CRB1* lead to thickening and disorganization of the retina[4] with a wide range of additional presentations including cystoid fluid collections and cone-rod dystrophy. There has been little evidence for a geneotype-phenotype correlation between mutations suggesting that genetic background influences severity. There are three CRB proteins found in the mammalian retina[5]. *CRB1* is expressed only in the brain and retina. It is comprised of a large extracellular domain that makes up 95% of the protein and contains a signal peptide, 19 epidermal growth factor (EGF)-like domains, and three laminin AG-like domains. There is a single transmembrane domain and a small intracellular domain of 37 amino acids. There are multiple isoforms but the prototypical isoform is 1406 amino acids. The majority of mutations occur in the extracellular domain. Missense mutations often affect residues in the laminin AG-like domains or the EGF-like domains and are predicted to disrupt disulfide bridges leading to irregular secondary structures[6]. Specific NMR studies on a N1317H mutation demonstrated altered calcium binding and affects on protein folding[7]. CRB2 expresses in the brain, retina, RPE, kidney, heart, and lung. CRB2 is 1285 amino acids with 15 EGF-like domains and 3 laminin AG-like domains[8]. CRB3 has a wide range of tissue expression although it lacks an extracellular domain. The three proteins come together to form the crumbs protein complex. PALS1 binds to the crumbs complex and recruits PATJ and MUPP1[9]. Alternatively, PAR6 can bind to the PDZ domain and recruit PAR3, aPKC (atypical protein kinase C), and CDC42 at the sub apical region (SAR) just apical of the adherens junction[10]. The adherens junction, found in photoreceptors and Müller glia, separates the apical and basolateral membrane. The recruitment of these proteins is important for maintaining cell polarity and for cell signaling[11].

The role of the extracellular domain in CRB1 and CRB2 has been understudied but there is evidence from motifs found in the structure for protein-protein interactions and roles in adhesion. Cell-cell interactions provide the foundation for cell cohesion and adhesion and the interactions are mediated by the binding of transmembrane proteins. Example proteins of homophilic binding are cadherins. E-cadherin (also known as cadherin 1) is the classical cadherin and has five extracellular cadherin repeats, a single transmembrane domain and a cytoplasmic domain, which binds to p120-catenin and β-catenin. It is a calcium-dependent protein that promotes tissue organization, suppresses cancer and blocks apoptosis through a combination of biophysical adhesion and intracellular signaling[12]. An atypical cadherin, Cadherin EGF LAG seven-pass G-type receptor 1 (CELSR1), does not interact with catenins but is important for cell adhesion. CELSR1 contains, nine cadherin domains, seven EGF-like repeats, and two laminin G-like domains[13]. Both CRB1 and CRB2 contain EGF-like domains and laminin AG-like domains[14]. Suggesting the potential for cell-cell adhesion.

Additional evidence for protein-protein interaction of the CRB1 and CRB2 extracellular domains come from CRB in *Drosphilia* and CRB2a and CRB2b in zebrafish. The extracellular domain of fly CRB contains multiple EGF-like and laminin G-like motifs. *In vitro* experiments demonstrate that the extracellular domain is necessary for aggregation through CRB-CRB interactions[15]. In zebrafish, there are 5 proteins in the CRB family: CRB1, CRB2a, CRB2b, CRB3a, CRB3b. There is cell-specific expression of CRB2a and CRB2b at the cell membranes of photoreceptor inner segments and Müller glia apical processes. The cell-cell interactions between the extracellular domains of these two proteins mediate cell-cell adhesion and the particular cone mosaic pattern found in zebrafish retina. Again, *in vitro* studies showed that the extracellular domain is required for cell aggregation[16]. These two studies set a precedent that would allow for the potential of CRB1 and CRB2 to be part of a new class of transmembrane cell-cell adhesion molecules.

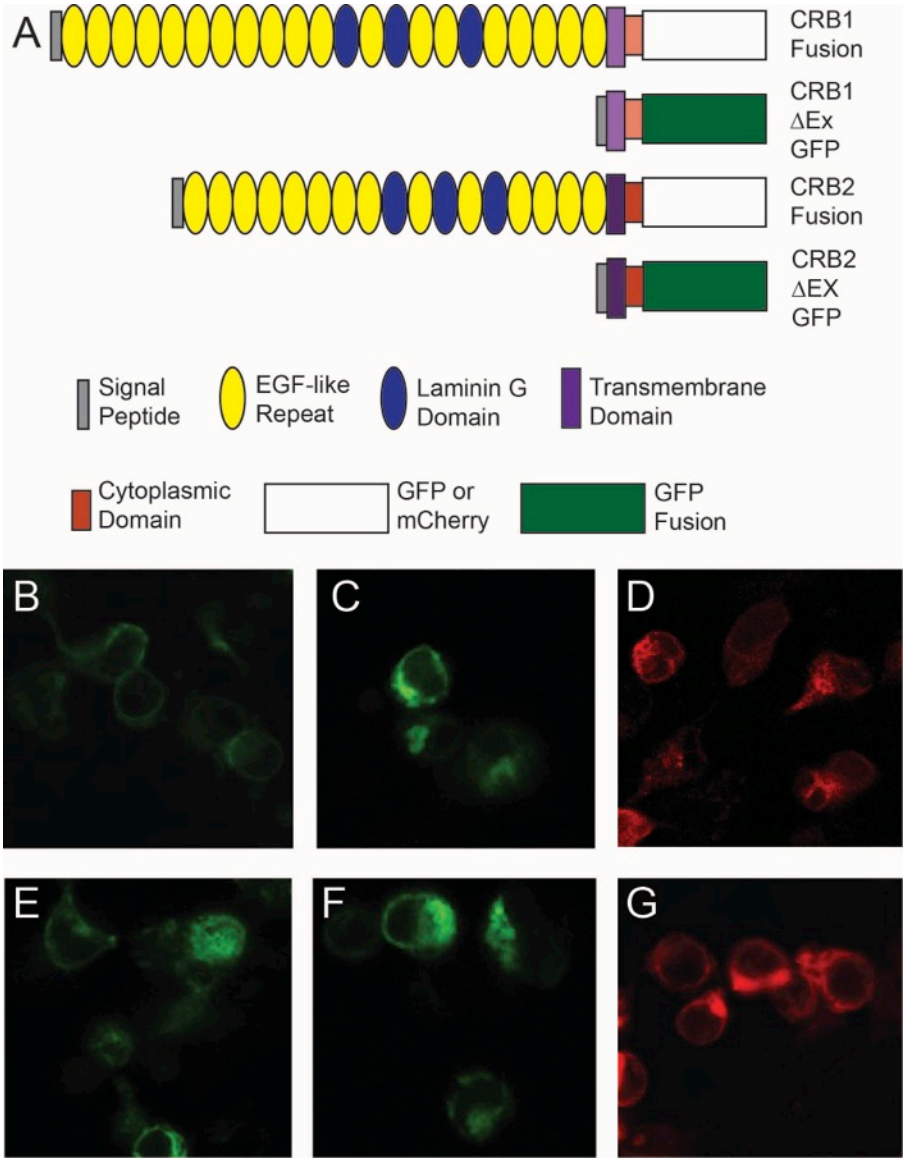


Figure 1. CRB fusion proteins and expression. A) Cartoon representation of the structure of CRB1 and CRB2 fusion constructs as well as the truncated versions of the proteins where the extracellular domain has been eliminated. B-G) Expression of constructs in HEK293T cells showing membrane localization. **B)** CRB1-GFP **C)** CRB1-ΔEx GFP **D)** CRB1-mCherry **E)** CRB2-GFP **F)** CRB2-ΔEx GFP **G)** CRB2-mCherry. The CRB2-mCherry construct showed strong clustering of the protein in the membrane.

In order to address the question of protein-protein interactions the single-molecule pull-down (SiMPull) assay was utilized. This is a methodology that takes advantage of the principles of traditional pull-down assays combined with single-molecule fluorescence microscopy. An imaging surface is coated with biotinylated antibodies against the protein of interest. In the case of CRB1 and CRB2, fusion proteins were created to allow for identification. Cell lysate containing the protein of interest is flowed into the chamber and the bound antibody captures that protein as well as its binding partner. The protein complexes can then be directly visualized through total internal reflection fluorescence (TIRF) microscopy. The SiMPull assay was used to probe the interactions of CRB1 and CRB2 fusion proteins. The hypothesis is that the extracellular domain of CRB1 participates in protein-protein interactions and that mutations found in CRB1 patients could contribute to the loss of this interaction leading to disease through loss of cell adhesion.

Results

In order to determine if the extracellular domains of CRB1 and CRB2 are necessary for protein-protein interactions, fusion constructs with both mCherry and GFP were constructed so that CRB1 and CRB2 could be visualized. In addition, truncated CRB1 and CRB2 constructs fused to GFP were generated where the entire extracellular domain was removed but the N-terminal membrane signal sequence was retained (**Fig. 1A**). These constructs were then transfected into HEK293T cells to evaluate expression. All of the constructs expressed in HEK293T cells and appeared to be localized to the cell membrane (**Fig. 1B-G**). Strong clustering was observed for the truncated ΔEx constructs and the CRB2-mCherry. A SiMPull assay was used to determine if CRB1 or CRB2 acts as a monomer or clusters. TIRF microscopy, which allows for the visualization of individual fusion proteins, was performed on chips coated with anti-GFP antibodies to capture either CRB1-GFP or CRB2-GFP. CRB1-GFP acted as a monomer, which was validated by bleaching events of individual GFP molecules where 77.8% indicated single bleaching events (**Fig. 2A-B**). CRB2-GFP showed greater clustering in TIRF images (**Fig. 2C**).

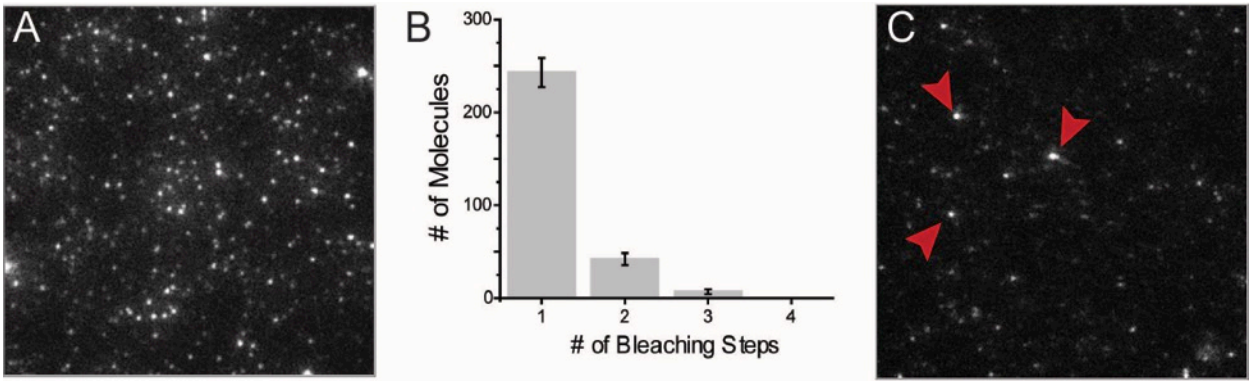


Figure 2. SiMPull assay with CRB1-GFP and CRB2-GFP. A) Representative image of TIRF for GFP pull down from cells expressing CRB1-GFP. **B)** Bleaching counts of CRB1-GFP suggesting that CRB1 is a monomer. **C)** Representative image of TIRF for GFP pull down from cells expressing CRB2-GFP. Red arrowheads indicate sites of strong clustering.

To determine the interaction of CRB1 with CRB2, a SiMPull assay was performed with CRB1-GFP pulling down CRB2-mCherry (**Fig. 3A-B**). There was strong colocalization of

CRB1-GFP with CRB2-mCherry with 83.9% of the mCherry signal colocalizing with GFP. This indicates protein-protein interaction. Since there was evidence that CRB1 acts as a monomer, further studies focused on the necessity of an extracellular domain for CRB1-CRB2 and CRB2-CRB2 interactions. In both cases when the extracellular domain was removed there was complete loss of protein-protein interactions (**Fig. 3C-D**).

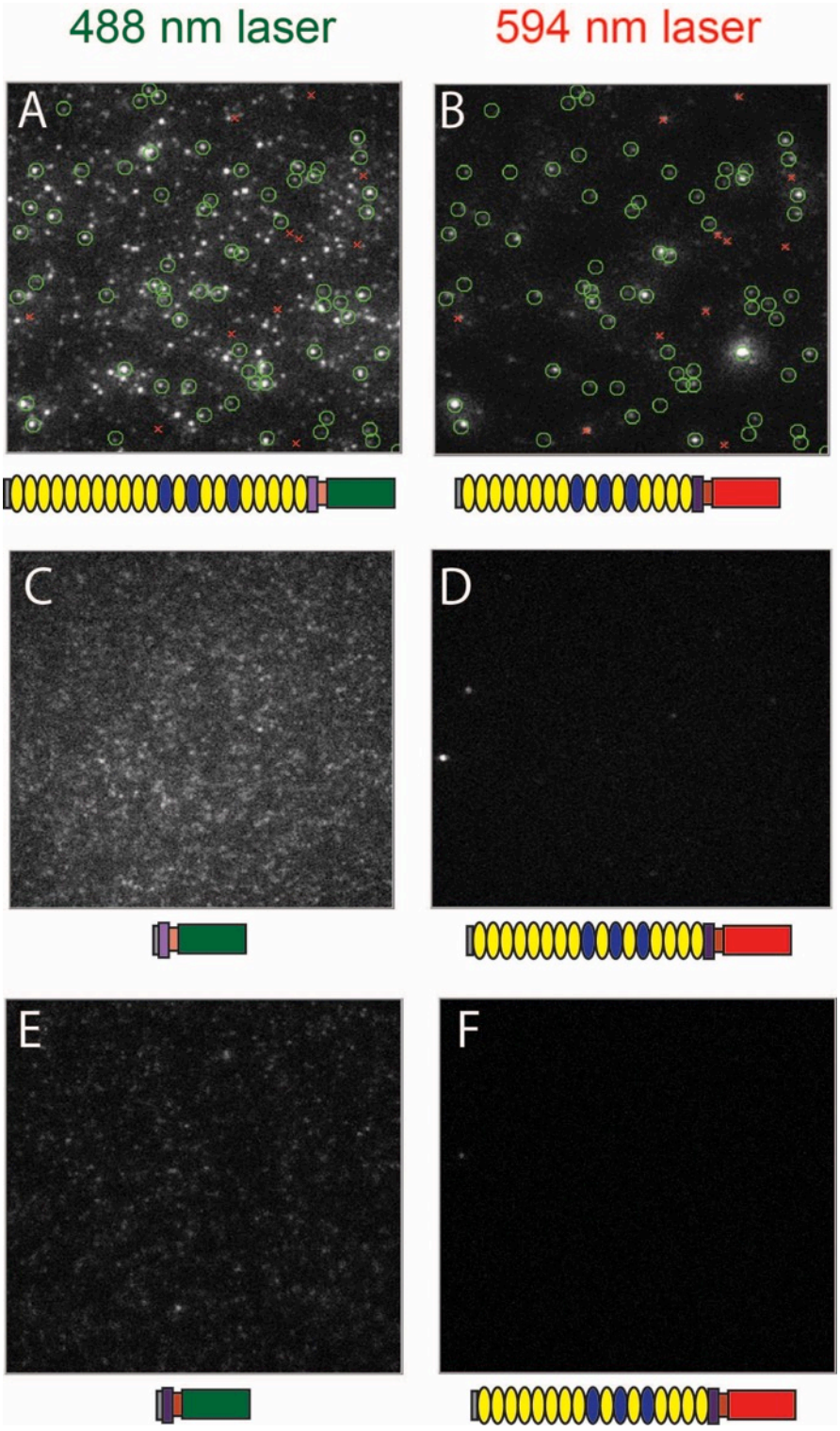


Figure 3. SiMPull Assay demonstrating that the extracellular domain of CRB1 and CRB2 are required for protein-protein interaction with CRB2. Representative images of TIRF for GFP pull downs from cells expressing CRB1 and CRB2 fusion proteins with the representative cartoon under each image. **A-B)** CRB1-GFP **(A)** pull down of CRB2-mCherry **(B)**. Green circles represent colocalization of the two fusion proteins. Red Xs represent only mCherry. There was 83.9% colocalization of GFP and mCherry. **C-D)** CRB1-ΔEx GFP **(C)** pull down of CRB2-mCherry **(D)**. Virtually no mCherry pull down was observed even with over saturating the SiMPull chip when the extracellular domain is removed. **E-F)** CRB2-GFP **(E)** pull down of CRB2-ΔEx mCherry **(F)**. No mCherry pull down when extracellular domain is removed.

Discussion

The majority of the literature has focused on the protein-protein interactions of the cytoplasmic domain and the downstream signaling of these interactions. These preliminary experiments begin to suggest that the extracellular domains play an important role in protein-protein interactions. CRB1 and CRB2 form a complex with CRB3 so it was not surprising to see protein-protein interactions between CRB1-GFP and CRB2-mCherry. It was very interesting that the extracellular domain was a necessary condition for this interaction. Also, loss of the extracellular domain prevented protein-protein interactions of both CRB1-CRB2 and CRB2-CRB2. The methodology used does not allow for discrimination between *cis* or *trans* protein interactions but does validate that this is an important path of research to be continued. An alternative explanation could be that the extracellular domain is only required for complex assembly and not for cell-cell interaction. Future experiments will involve Western blots to validate the CRB1 and CRB2 fusion expression. FRET could be used to elucidate the *cis* and *trans* interactions. *In vitro* assays could be used as well to determine changes in Wnt, Notch, or Hippo signaling of which there is evidence of influence of the CRB complex. The strength of interactions could also be tested to determine if CRB1 or CRB2 plays a role in cell-cell adhesion. Further experimentation will help to inform the best treatment strategy for patients who suffer from inherited retinal degenerations caused by *CRB1* mutations.

Methods

Cloning CRB fusion constructs

CRB1 and CRB2 cDNA was amplified from genomic DNA extracted from HEK293T cells. The cDNAs were then cloned into TOPO vectors. A CMV promoter and eGFP sequence containing a GSGSGS linker was cloned using Gibson cloning and digesting with *Eag*I. Gibson cloning was also used to add mCherry by digesting CRB1-GFP or CRB2-GFP with *Bgl*II and *Eag*I. The extracellular domain of CRB1 comprises amino acids 26-1347. Gibson cloning was designed to remove amino acids 34-1343 to ensure that the N-terminal signaling sequence would remain and the transmembrane domain would be minimally disrupted. The extracellular domain of CRB2 comprises amino acids 29-1224. Gibson cloning was designed to remove amino acids 33-1229 to ensure that the N-terminal signaling sequence would remain and the transmembrane domain would be minimally disrupted. Primers can be found in **Table 1**.

Transfection and imaging

A 6-well plate of HEK293T cells grown on Poly-D-Lysine (1.0 mg/mL) coated 18 mm coverslips was transfected with 0.75 µg of the GFP and constructs and 0.5 µg of the mCherry constructs using 2 µL of JetPrime transfection agent. The plates were incubated at 37º C for 48 hours. Confocal microscopy using a Zeiss LSM 880 NLO AxioExaminer was then performed.

SiMPull Assay

Constructs were transfected in the same way as for imaging. The plates were incubated at 37º C for 48 hours. Cells were detached using 700 µL of Calcium Free PBS for 20 min. Cells

were then pelleted and the supernatant was removed. The pellet was then resuspended in lysis buffer (150 mM NaCl, 1 mM EDTA, protease inhibitor) with 1.5% IGEPAL. Cells were lysed for 1 hour at 4ºC. The lysate was pelleted and the supernatant was collected. Slides were prepared and coated with a biotinylated anti-GFP antibody. A TIRF microscope was used to acquire the single-molecule images. The supernatant was serially diluted between 1:10-1000 in lysis buffer immediately before being flowed through the chamber and allowed to incubate for two minutes. Additional lysis buffer was applied to wash unbound protein and then samples were imaged.

Table 1.

CRB1cDNA+	ATGGCACTTAAGAACATTAACTACCTTCTCATCTTCTAC
CRB1cDNA-	AATCAGTCTCTCCATTGCAGGGGGT
CRB2cDNA+	ATGGCGCTGGCCAGGCCTG
CRB2cDNA-	CTAGATGAGTCTCTCCTCCGGTG
EagCRB1GFP+	tctctccacaggtgtccaggc
EagCRB1GFP-	cctcgcccttgctcacggccggaTCCAGATCCAGATCCAATCAGTCTCTCC
C1mCherryGib1+	ctccacaggtgtccaggcg
C1mCherryGib1-	ctcacTCCAGATCCAGATCCAATCAGTC
C2mCherryGib1+	ctccacaggtgtccaggcggccgccATGGCGCTGGCCAGGCCT
C2mCherryGib1-	ctcacTCCAGATCCAGATCCGATGAGTCTCTCCTCCGGTGGCAC CTTGA
C2mCherryGib2+	gagactcatcggatctggatctggaGTGAGCAAGGGCGAGGAG
C2mCherryGib2-	GAAGGCACAGTCGAGGCAGATCTttacttgtacagctcgtccatgcc
C2dExGFP1-	ATAGCCACCACTCTCGGTAG
C2dExGFP1-	CGTCCCAGCCAGGAGGGAAA
C2dExGFP2+	CTACCGAGAGTGGTGGCTAT GTGTCTGAATGTATCTGCAAT GCCAGA
C2dExGFP2.2+	TTTCCCTCCTGGCTGGGACG TTCCCACTGCTGGAGGTG

References

1. Hollander, den, A. I. *et al.* Mutations in a human homologue of Drosophila crumbs cause retinitis pigmentosa (RP12). *Nat. Genet.* **23,** 217–221 (1999).
2. Richard, M. *et al.* Towards understanding CRUMBS function in retinal dystrophies. *Human Molecular Genetics* **15 Spec No 2,** R235–R243 (2006).
3. Hollander, den, A. I. *et al.* Isolation of Crb1, a mouse homologue of Drosophila crumbs, and analysis of its expression pattern in eye and brain. *Mech. Dev.* **110,** 203–207 (2002).
4. Aleman, T. S. *et al.* Human CRB1-Associated Retinal Degeneration: Comparison with the rd8 Crb1-Mutant Mouse Model. *Investigative Ophthalmology & Visual Science* **52,** 6898–6910 (2011).
5. Hollander, den, A. I. *et al.* CRB1 mutation spectrum in inherited retinal dystrophies. *Hum. Mutat.* **24,** 355–369 (2004).
6. Gosens, I., Hollander, den, A. I., Cremers, F. P. M. & Roepman, R. Composition and function of the Crumbs protein complex in the mammalian retina. *Experimental Eye Research* **86,** 713–726 (2008).
7. Davis, J. A., Handford, P. A. & Redfield, C. The N1317H substitution associated with Leber congenital amaurosis results in impaired interdomain packing in human CRB1 epidermal growth factor-like (EGF) domains. *J. Biol. Chem.* **282,** 28807–28814 (2007).
8. Bujakowska, K. *et al.* CRB1 mutations in inherited retinal dystrophies. *Hum. Mutat.* **33,** 306–315 (2011).
9. van Rossum, A. G. S. H. Pals1/Mpp5 is required for correct localization of Crb1 at the subapical region in polarized Muller glia cells. *Human Molecular Genetics* **15,** 2659–2672 (2006).
10. Pellissier, L. P. *et al.* CRB2 acts as a modifying factor of CRB1-related retinal dystrophies in mice. *Human Molecular Genetics* (2014). doi:10.1093/hmg/ddu089
11. van de Pavert, S. A. *et al.* Crb1 is a determinant of retinal apical Müller glia cell features. *Glia* **55,** 1486–1497 (2007).
12. van Roy, F. & Berx, G. The cell-cell adhesion molecule E-cadherin. *Cell. Mol. Life Sci.* **65,** 3756–3788 (2008).
13. Strutt, D., Schnabel, R., Fiedler, F. & Prömel, S. Adhesion GPCRs Govern Polarity of Epithelia and Cell Migration. *Handb Exp Pharmacol* **234,** 249–274 (2016).
14. Thompson, B. J., Pichaud, F. & Röper, K. Sticking together the Crumbs — an unexpected function for an old friend. *Nat. Rev. Mol. Cell Biol.* **14,** 307–314 (2013).
15. Letizia, A., Ricardo, S., Moussian, B., Martín, N. & Llimargas, M. A functional role of the extracellular domain of Crumbs in cell architecture and apicobasal polarity. *Journal of Cell Science* **126,** 2157–2163 (2013).
16. Zou, J., Wang, X. & Wei, X. Crb apical polarity proteins maintain zebrafish retinal cone mosaics via intercellular binding of their extracellular domains. *Dev. Cell* **22,** 1261–1274 (2012).