TRANSCRIPTION-DEPENDENT DIRECTED EVOLUTION OF THE ADENO-

ASSOCIATED VIRUS CAPSID

By

Samuel Jonathan Huang

A DISSERTATION

Presented to the Department of Physiology & Pharmacology

and the Oregon Health & Science University

School of Medicine

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

June 3rd, 2021

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

_____

This is to certify that the PhD dissertation of

Samuel J. Huang

has been approved

_____
Hiroyuki Nakai, MD, PhD, Mentor/Advisor

_____
Robert Duvoisin, PhD, Committee Chair

_____
Daniel Marks, MD, PhD, Member

_____
Cary Harding, MD, Member

_____
Mark Pennesi, MD, PhD, Member

_____
John Brigande, PhD

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AAV	adeno-associated virus
BBB	blood brain barrier
bp	base pair
CREATE	Cre recombination-based AAV targeted evolution
CSF	cerebrospinal fluid
ELISA	enzyme-linked immunosorbent assay
FDR	false discovery rate
GFP	green fluorescent protein
HEK	human embryonic kidney
hSYN1	human synapsin 1
ITR	inverted terminal repeat
MVM	minute virus of mice
N2A	neuro-2A
NHP	non-human primate
NS	non-splicing
nt	nucleotide
PEI	polyethyleneimine
RT-PCR	reverse transcription polymerase chain reaction
SA	splice acceptor
SD	splice donor
SMA	spinal muscular atrophy
SMN	survival motor neuron
TRACER	Tropism redirection of AAV by cell-type-specific expression of RNA
TRADE	transcription-dependent directed evolution

ACKNOWLEDGMENTS

First and foremost, I would like to thank Dr. Hiroyuki Nakai who enthusiastically took me into his lab knowing that I had never done PCR. He has demonstrated a seemingly endless enthusiasm for science and the pursuit of knowledge, epitomized by 4-hour long lab meetings discussing the orientation of gradients in a fixed angle centrifuge. Furthermore, he has introduced me to the world of academic-industry relationships and the role that they play in translating discoveries. He has modeled for me the pursuit of scientific rigor and has conceptually redefined the concept of "hardwork."

I am also grateful to my dissertation advisory committee including Drs. Robert Duvoisin, Cary Harding, Mark Pennesi, Dan Marks, and Hiroyuki Nakai, who offered support for both scientific and non-scientific challenges.

Many thanks are owed to current and former members of the lab: Zhen, Xiaolan, Helen, Anusha, Jessie, Courtney, Kazu, Teru, Taisuke, Yoshi1, Yoshi2, Yanhua, John, Rika, Allison, Swapna, and Qing. Who have I not asked to pull a plate or freeze a reaction? In particular, I offer my thanks to Dr. Kei Adachi for being a consistent source of practical information, technical advice, and a reassuring chuckle.

Thank you to collaborators and the larger OHSU scientific community. Special mention to Dr. Renee Ryals for always including me in projects and conversations and sharing her journey to independence. NHP work could not be carried out without Drs. Greg Dissen and Jodi McBride, as well as Will and the entire veterinary and pathology teams, including Dr. Anne Lewis for sharing her pathology perspective. Thank you to nearby labs, especially Suh in the Hayflick lab. Also, thank you to the Mandel lab and John. Thanks to the Grompe lab, in particular Bin, Amita, and Jeffrey. Imaging could not

ABSTRACT

In recent years, *in vivo* adeno-associated virus (AAV)-mediated gene therapy has demonstrated remarkable safety and efficacy, culminating in the approval of two directly administered AAV vector products for central nervous system (CNS) disease. AAV vectors offer hope for transforming a wide variety of clinical practices and addressing many diseases at the gene level, in particular inherited diseases. This is evidenced by the hundreds of ongoing clinical trials.

While remarkable, the amazing clinical success may only represent the most basic potential of this vector. Diseases addressed by AAV-mediated gene therapy have been carefully selected, in part, due to the readily available naturally-occurring capsid serotypes, which mediated sufficient gene transfer to target cell types. However, future applications, and perhaps improved efficacy for current applications, can be achieved by engineering the AAV vectors. Indeed, the field of AAV capsid engineering is rapidly developing. Yet a major obstacle is the ability to develop next-generation AAV vectors with clinical relevance.

In the first part of this dissertation (Chapter 2), we focus on addressing the technical limitations imposed by current directed evolution selection strategies in large, clinically-relevant animal models (i.e., non-human primates), and develop a novel platform termed TRAnscription-dependent Directed Evolution (TRADE). After establishing proof-of-concept for this technology, we apply it in hopes of addressing the issue of efficiently transducing brain neurons following systemic vector administration in non-human primates (Chapter 3).

In summary, this dissertation describes the development of the TRADE system for directed evolution of the AAV capsid, its successful application to identify the HN1 variant, a novel AAV capsid with enhanced neuronal transduction efficiency and specificity following intravenous vector administration in both rodents and rhesus macaques. This work establishes a novel technology with the potential to facilitate the rapid development of customized AAV vectors targeting cell types of choice and has the potential to significantly impact a wide variety of problems in basic science and medicine where lack of an efficient vector for targeting a critical cell type poses a significant barrier. Thus, TRADE is poised to transform the field of AAV capsid directed evolution.

1   Chapter 1. Introduction

1.1   AAV-mediated gene therapy

1.1.1   *The adeno-associated virus*

Discovery and classification

Fifty-six years ago, a defective viral particle was discovered as a contaminant in a simian adenoviral preparation (Atchison, Casto, and Hammon 1965). The particles were termed adenovirus-associated defective virus particles in reference to their inability to replicate without the presence of adenovirus and, later, were referred to as the adeno-associated virus (AAV). While the name is historically fitting, it is important to recognize that AAV is classified as a parvovirus. Parvoviruses are small, ~26nm, non-enveloped, icosahedral viruses carrying a linear single-stranded DNA genome flanked by inverted terminal repeats (ITRs). Further taxonomic classification places AAV in the genus Dependoparvovirus, in reference to the dependence of its life cycle on coinfection with an adenovirus, herpesvirus, or vaccinia virus. Since the original discovery, hundreds of naturally-occurring AAV variants have been identified from a wide variety of species (Bossis and Chiorini 2003; Schmidt et al. 2004; Farkas et al. 2004; Y. Li et al. 2010), including the commonly referred to thirteen primate serotypes (AAV1-AAV13) (Gao et al. 2002, 2004). Many variants have been isolated from human tissues, demonstrating that humans are a natural host for AAV. AAV variants are now phylogenetically organized into Clades A-F such that members of a given clade are often functionally similar (Gao et al. 2004; Vandenberghe et al. 2009).

The AAV Genome

The prototypical AAV2 genome consists of a ~4.7 kb sequence consisting of homotelomeric ITRs flanking the Rep and Cap genes. Despite its small size, the AAV genome is able to encode at least nine unique protein products (see below). This is achieved by each gene encoding multiple distinct proteins using mechanisms such as a shared intron and polyadenylation signal, alternative splicing and use of non-canonical start codons, and overlapping open read frames (Figure 1.1).

The 145 nucleotide ITRs located at each of the 5' and 3' ends of the AAV genome include 125 nucleotides that facilitate formation of a hairpin structure and an additional 20 nucleotide non-palindromic region, termed the D-sequence. The ITRs play a vital role in replication by priming DNA replication, mediating replication resolution, and also contain the packaging signal that mediates genome encapsidation (Lusby, Fife, and Berns 1980; Xu Shan Wang, Ponnazhagan, and Srivastava 1995; X S Wang, Ponnazhagan, and Srivastava 1996).

The *rep* gene contains a single open reading frame encoding four non-structural proteins. The large Rep78 and Rep68 proteins are translated from the p5-driven transcript. Expression of the small Rep52 and Rep40 proteins is driven by the p19 promoter. Rep68 and Rep40 share an intron that results in the C-terminus overlapping with the *cap* gene. Together, the Rep proteins play critical roles in controlling activity of viral promoters (Pereira, McCarty, and Muzyczka 1997; Weger et al. 1997), genome replication and resolution (Im and Muzyczka 1990), and genome packaging (King, Dubielzig, and Grimm 2001).
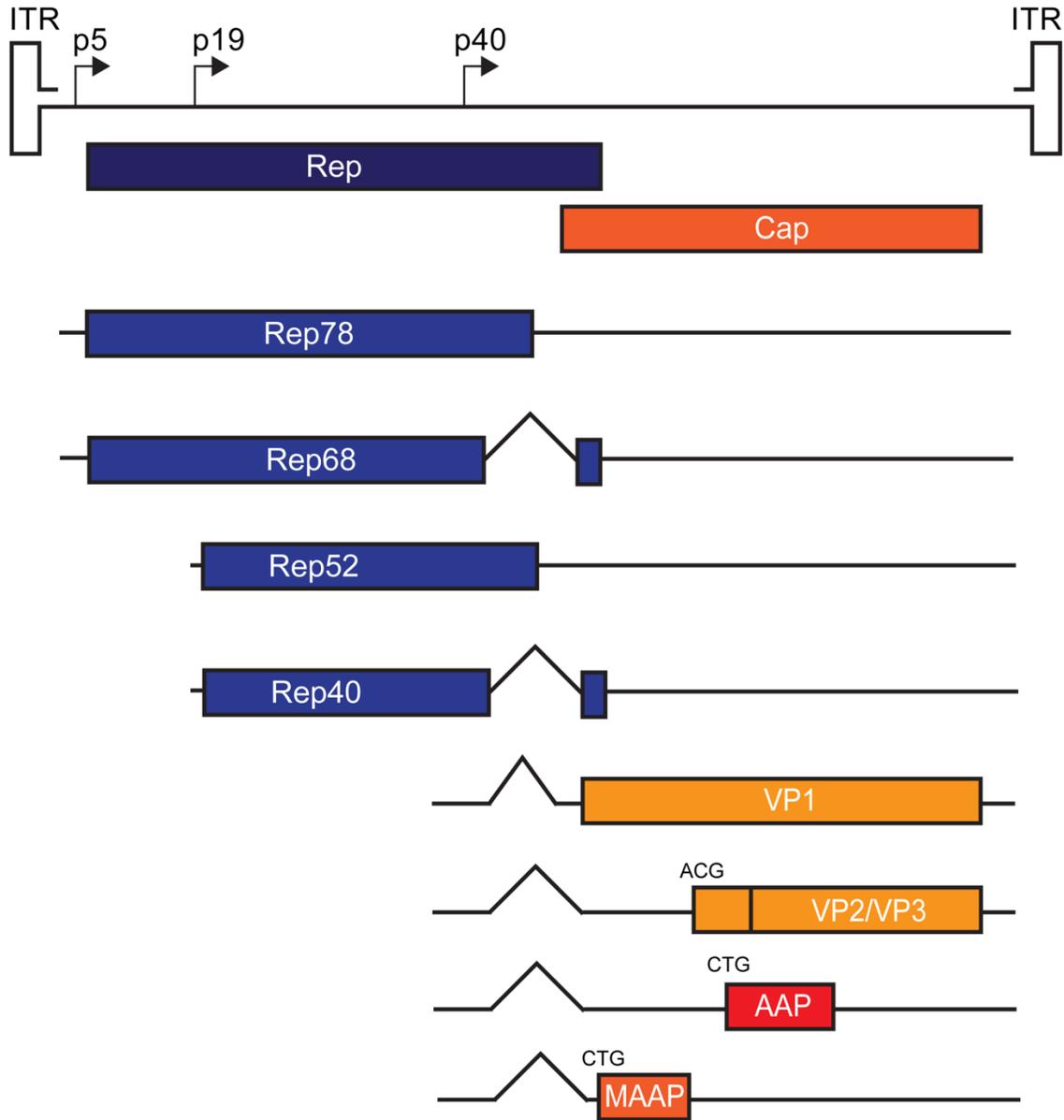
Figure 1.1 Schematic of the prototypical AAV2 genome
The AAV2 genome is flanked by inverted terminal repeats (ITRs) flanking the *Rep* and *Cap* genes. *Rep* encodes four non-structural proteins named Rep78, Rep68, Rep52, and Rep40 by using the p5 and p19 promoters and differential splicing. *Cap* encodes three capsid viral proteins (VP1, VP2, VP3) from p40-driven transcripts by utilizing differential splicing and a non-canonical ACG start codon. In addition, *Cap* encodes the assembly activating protein (AAP) and membrane-associated accessory protein (MAAP) from overlapping reading frames.

The *cap* gene contains three known open reading frames. The Cap open reading frame encodes the three viral proteins (VP1, VP2, and VP3) using two p40-driven mRNA transcripts generated by alternative splicing. All three VPs share a C-terminus, but VP1 and VP2 contain N-terminal extensions relative to VP3. The minor splice form encodes the full-length VP1 protein, whereas the major splice form encodes both VP2 and VP3 through the use of a non-canonical ACG start codon for VP2 and a downstream ATG start codon for VP3 (Figure 1.1). Ultimately, this results in a relative protein expression of approximately 1:1:10 for VP1:VP2:VP3, respectively.

Until as late as 2010, 45 years after the initial discovery of AAV, the genome was commonly reported to have only two open reading frames. At this time, a second open reading frame with a +1 frameshift relative to the Cap open reading frame was reported (Sonntag et al. 2011). This open reading frame, located between the VP2 and VP3 N-termini, utilizes yet another non-canonical CTG start codon and encodes the assembly-activating protein (AAP), so named to reflect its critical role in viral capsid assembly (Figure 1.1). Surprisingly, yet another open reading frame was recently discovered within the *cap* gene, overlapping the VP1 N-terminal unique region and encoding the so-called membrane-associated accessory protein (MAAP) (Ogden et al. 2019). The MAAP open reading frame utilizes a CTG start codon, similar to AAP. The precise functions of this protein remain to be determined, but interestingly, its positioning within the VP1 unique region is associated with a complete lack of ATG codons (Karlin 2020), perhaps suggesting an evolutionary pressure to avoid start codons that might interfere with MAAP expression.

AAV, like other parvoviruses, uses rolling hairpin replication that results in the generation of plus and minus AAV genomes (Figure 1.2). Formation of the hairpin loop at the 3' ITR primes second-strand synthesis (Lusby, Fife, and Berns 1980) that is then carried out by a replication complex consisting of the large Rep proteins and host proteins, including replication factor C, proliferating cell nuclear antigen, and polymerase δ (Nash, Chen, and Muzyczka 2008). Replication continues through the 5' ITR by strand displacement. This double-stranded intermediate still contains one hairpin loop, but is transcriptionally active. Subsequently, the large reps mediate nicking activity at the terminal resolution site (TRS) in the remaining hairpin loop, which allows the remaining portion of the genome to be replicated by strand displacement through the hairpin loop. At this point, replication is complete in that there are two full-length copies of the AAV genome. However, the "rolling" hairpin replication reinitiates when the ITRs reform their hairpin structure, thus allowing new strands to be synthesized and released by strand displacement. The replication process appears to be linked to packaging into pre-formed capsids by interactions of the large Rep proteins, that covalently attached to the AAV genome during terminal resolution, and the 5-fold pore (see below) of the AAV capsid (Prasad, Zhou, and Trempe 1997).

AAV utilizes homotelomeric ITR sequences for replication, resulting in the equimolar production of plus and minus strands during replication that are both efficiently packaged into capsids (Zhou et al. 2008). Intriguingly, to date, no coding sequences have been identified on the AAV minus strand. However, an effort to comprehensively characterize the AAV genome using RNA-seq, identified minus strand

transcripts upstream of the p5 promoter, though their function remains to be determined

(Stutika et al. 2016).
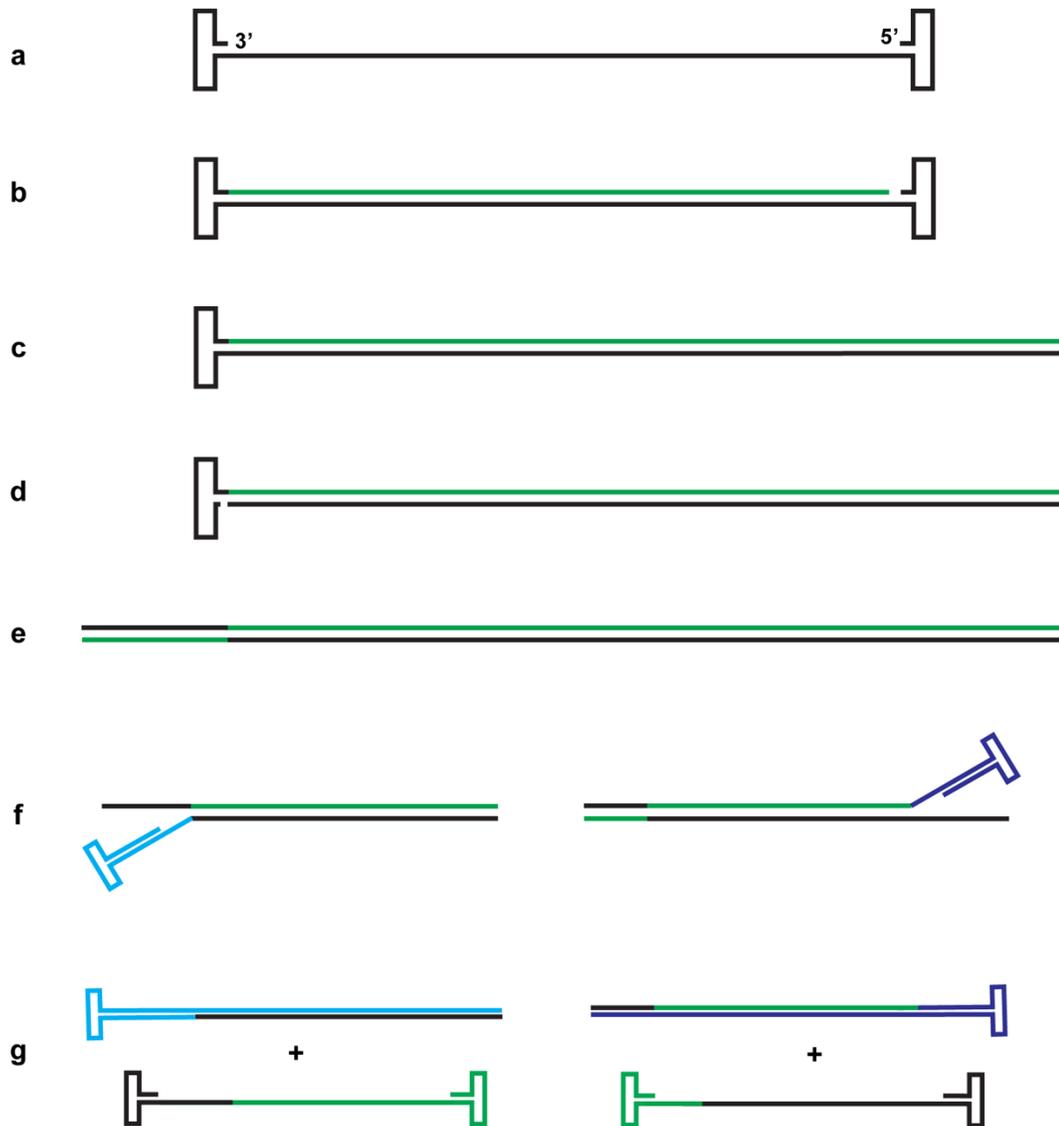
Figure 1.2 Rolling hairpin replication of the AAV genome
(a) The single-stranded AAV ITR sequences form hairpin structures. The 3' end serves as a primer for DNA replication. (b) The single-stranded AAV genome is converted to a double-stranded genome by second strand synthesis, mediated by a replication complex consisting of AAV Rep proteins and host replication machinery. (c) Replication continues through the 5' ITR hairpin structure. (d) The terminal resolution site is nicked on the intact ITR by the large Rep proteins. (e) The remaining ITR hairpin is replicated by extension of the newly liberated 3' end. (f) ITR hairpin structures reform and the 3' end again self-primes replication. (g) DNA synthesis then releases new single-stranded genomes by strand displacement and the new double double-stranded intermediate is recycled for further replication (return to (c)). Single-stranded progeny are generated in two polarities at an equimolar ratio.

The AAV Capsid

The AAV genome is surrounded by a 60mer T = 1 icosahedral capsid made up of VP1, VP2, and VP3 in approximate ratio of 1:1:10. The general structure of the icosahedral capsid can be organized into twelve pentamers such that VP monomers make 2-fold, 3-fold, and 5-fold interactions that define the surface topology (Xie et al. 2002; Govindasamy et al. 2006; Nam et al. 2007). The 5-fold axis contains a pore connecting to the internal surface of the capsid, the 3-fold axis consists of protrusions that extend above the surface of the capsid, and the 2-fold axis forms a depression between the 3-fold protrusions. Since the AAV virion consists of only the genome and the capsid, the capsid is responsible for mediating all of the functions necessary for viral infection, up to delivery of the viral genome. Since AAV capsids remain intact until reaching the nucleus (Sonntag et al. 2006), and perhaps remain intact for some time within the nucleus, (Johnson and Samulski 2009), it is unsurprising that differential capsid-host interactions will ultimately define the properties of the vector. In short, the AAV capsid determines its infective properties (Hauck and Xiao 2003).

As a case-in-point, the capsid's role in the initial step of infection, cellular attachment, clearly demonstrates that viral tropism is capsid-dependent. Early studies identified glycans as important host cellular determinants of AAV tropism. It is well-known that AAV2 binds heparan sulfate proteoglycan as its primary receptor (Summerford and Samulski 1998). AAV3, AAV6, and AAV13 can also bind heparan sulfate, but not as a primary receptor for AAV6 (Lerch and Chapman 2012; Halbert, Allen, and Miller 2001; Mietzsch et al. 2014). AAV1, AAV4, AAV5, and AAV6 are all able to bind sialic acid residues, while AAV9 has a notable terminal galactose-binding

footprint (Bell et al. 2012). Given this basic biological understanding, it is then unsurprising that differences in the ability to bind heparan sulfate correlate with different tropism profiles. This has been formally demonstrated in multiple studies. One such study determined that the differential ability of AAV1 and AAV6 to bind heparan sulfate is determined by a single amino acid (Wu et al. 2006). The AAV1-E531K and AAV6-K531E variants effectively reverse the heparan sulfate-binding capability of the parental serotypes. Furthermore, liver transduction was predicted by the heparan sulfate-binding capability of each AAV capsid (low efficiency exhibited by AAV1 and AAV6-K531E, but higher efficiency exhibited by AAV6 and AAV1-E531K). The importance of heparan sulfate-binding in determining vector tropism was subsequently extended to skeletal muscle and brain tissue. Heparin-binding variants demonstrated enhanced skeletal muscle transduction, but attenuated transduction in the brain following intraparenchymal injection into each tissue (Arnett et al. 2013), predictably demonstrating that modulating heparan sulfate-binding would not be a universal means of enhancing transduction. Indeed, another study showed that deletion of the heparan sulfate binding motif on the AAV2 capsid increased transduction of photoreceptors following subretinal delivery (Boye et al. 2016). Of historical interest, this basic biological understanding also explains, retrospectively, why the first AAV clinical trial attempting to treat cystic fibrosis by intranasal administration of AAV2 was not efficacious (Flotte et al. 2003). Namely, heparan sulfate is differentially expressed on the apical and basolateral membranes of the airway epithelium. Low levels of heparan sulfate on the apical surface, which is exposed to AAV vectors following intranasal administration, do not mediate efficient capsid attachment (Srivastava 2016).

Post-attachment mechanisms mediating viral entry are not precisely defined, but remain capsid-dependent (Figure 1.3). Different AAV serotypes recognize different protein co-receptors. AAV2 has been reported to recognize integrins αVβ5 and α5β1, as well as the fibroblast growth factor receptor 1 (FGFR1), and the hepatocyte growth factor receptor (Qing et al. 1999; Summerford, Artlett, and Samulski 1999; Asokan et al. 2006; Kashiwakura et al. 2005). AAV2 also utilizes the laminin receptor, as does AAV3, AAV8, and AAV9 (Akache et al. 2006). Subsequently, or perhaps coincident with co-receptor recognition, AAV is endocytosed. Serotypes as divergent as AAV2 and AAV5 are known to use a common mechanism of dynamin-regulated clathrin-coated pits (Bartlett, Wilcher, and Samulski 2000; Bantel-Schaal, Hub, and Kartenbeck 2002), though other mechanisms may also play a role. Endosomal escape is then achieved via acidification-catalyzed extrusion of the VP1 unique region through the 5-fold pore, resulting in the unveiling of a phospholipase A2 domain (Sonntag et al. 2006). Mutations to this domain were tolerated with respect to the production of AAV2 virions that were able to bind to and enter HeLa cells, but infectivity measured by gene expression was drastically reduced (Girod et al. 2002). A similar observation was made for mutations affecting the 5-fold pore that prevent the exposure of the VP1 N-terminus, that is no loss of vector production, but a clear attenuation of viral infectivity (Bleker, Sonntag, and Kleinschmidt 2005).

Following entry and intracellular trafficking (Figure 1.3), AAV then accumulates at the perinuclear space. Notably, an in vitro assay identified that AAV2 could bind to microtubule proteins, suggesting a direct interaction for intracellular trafficking (Kelkar et al. 2006). A later study used Cy3-labeled AAV particles to

determine that trafficking was indeed along the microtubule network, however, virions appeared to still be within endosomes based on electron microscopy (P.-J. Xiao and Samulski 2012). In the same study, the authors found that pharmacological disruption of the microtubule network delayed endosomal acidification. These observations are interesting in light of earlier observations that microinjections of AAV2 directly into the cytoplasm only led to detectable infection (i.e., Rep expression) in less than 1% of injected cells despite apparent perinuclear accumulation (Sonntag et al. 2006). Together, these data suggest that AAV capsid modifications within the endosome may be necessary for nuclear entry.

After nuclear entry, it appears that the capsid remains intact for some period of time and mediates trafficking between the nucleolus and nucleoplasm (Johnson and Samulski 2009) until uncoating of the genome. One might assume at this point that the AAV capsid has played its full role in AAV infection. However, there are a limited number of studies that suggest that the AAV capsid plays a role in transcription of the genome (Aydemir et al. 2016; Salganik et al. 2014). The authors reach this conclusion by comparing wild-type AAV2 to AAV2 variants with mutations in the 2-fold depression. Intriguingly, some of the mutants were able to achieve a similar number and proportion of uncoated genomes in the nucleus compared to AAV2 (measured as Benzonase-sensitive particles in the nuclear fraction), but with a far lower transduction efficiency (measured by detectable GFP expression) that correlated well with a decrease in mRNA expression (measured by qPCR). Very recent studies now suggest that context-specific capsid-promoter interactions dictate transcription in different cell types.

To summarize this section, AAV is a seemingly simple virus, but much of the basic biology requires further study. Relatively recent investigations have revealed entire open read frames in the *cap* gene that were previously unknown to the field (Sonntag et al. 2011; Ogden et al. 2019; Cao, You, and Hermonat 2014). Likewise, substantial effort has been invested in establishing capsid structure-function relationships and, despite exciting progress (Adachi et al. 2014), we are only beginning to understand the complex roles that individual amino acids can play in multiple different stages of the AAV lifecycle. Thus, basic AAV biology remains an area of intense investigation because developing a foundational knowledge of the factors influencing AAV infection will facilitate the manipulation of AAV vectors for better gene delivery technologies.

1.1.2 Recombinant AAV vectors

Development of the recombinant AAV vector system

A landmark advance for the budding field of AAV biology was the successful cloning of the AAV2 genome into a plasmid and subsequent demonstration that viral particles could be rescued from human cells transfected with the AAV plasmids followed by adenoviral infection (Samulski et al. 1982; Laughlin et al. 1983). Subsequently, a recombinant AAV vector system was generated that replaced the viral *rep* and *cap* genes flanked by the ITRs with a custom transgene cassette, in this case a SV40-driven neomycin, and provided the *rep* and *cap* functionalities in trans with another plasmid. Transfection of both plasmids into human cells followed by Adenoviral infection resulted in the recovery of vectors with an AAV capsid, but a synthetic and customized genome

flanked by AAV ITRs (McLaughlin et al. 1988). Incredibly, the pair of 145 nucleotide ITRs are the only cis-acting requirement for packaging of the AAV genome.

Another major step was the development of the adenovirus free triple transfection method that allowed production to occur in the absence of live Adenovirus infection by transferring the necessary helper functions to a third plasmid (X. Xiao, Li, and Samulski 1998; Matsushita et al. 1998). Although this method was reported more than two decades ago, it remains the standard in both academic labs and industry. Essentially, triple transfection requires delivery of three plasmids into a packaging cell line that supplies the adenoviral E1A gene. The three plasmids are (1) pVector or pITR, a plasmid containing an ITR-flanked AAV genome carrying the transgene cassette; (2) pRepCap, a plasmid carrying the desired Rep and Cap genes (note that AAV2 Rep can support many, but not all Cap sequences); and (3) pHelper is a plasmid containing the adenoviral E4orf6, E2A, and VA RNA. It is notable that this particular three plasmid configuration is convenient, but the system supports virtually any number of plasmids so long as all of the necessary components are provided in trans (Powers et al. 2018).

Subsequent work demonstrated that pseudotyped AAV vectors could be generated that consisted of unmatched AAV serotypes and genotypes. Namely, it appeared that any of the five AAV capsid serotypes available at the time could cross-package an AAV2 ITR-containing vector genome (Chao et al. 2000), and this finding has since been extended to virtually any capsid variant. Not only did this finding justify the tremendous convenience of using a single well-validated pVector, but the generation of recombinant vectors with different capsids but identical genomes led to the conclusion that many

vector phenotypes were primarily dependent on the capsid (Hauck and Xiao 2003; Grimm et al. 2006).

AAV-mediated transduction

Although the processes are highly related, we differentiate wild-type AAV infection from recombinant AAV vector-mediated transduction. Here, transduction is defined as the ability to successfully mediate all of the steps necessary to achieve transgene expression (Figure 1.3). The process essentially takes advantage of the natural AAV life cycle, which is arrested due to substitution of the viral *rep* and *cap* genes. Notably, the lack of a Rep transgene prevents the efficient integration of the AAV genome into the host (D. M. McCarty, Young, and Samulski 2004). Instead, the vector achieves a stable double-strand DNA form that is transcriptionally active and is known to predominantly persist as extrachromosomal concatemers (Nakai et al. 2001; Flotte, Afione, and Zeitlin 1994).
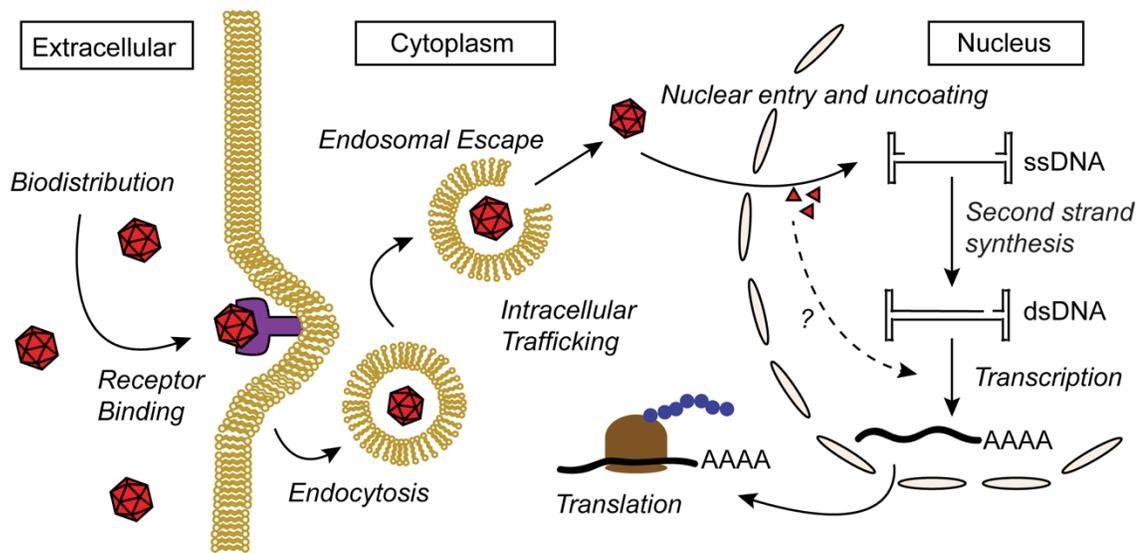
Figure 1.3 AAV transduction is a complex, multi-step process
Following administration, the AAV vector must biodistribute to the target tissue and cell type. Subsequently, the vector is internalized by receptor-mediated endocytosis. Acidification of the endosome mediates externalization of a phospholipase A2 domain that facilitates endosomal escape. After trafficking and entering the nucleus, the capsid must be shed and the single strand DNA genome is converted to a double stranded form that is transcriptionally active, allowing expression of transgene mRNA and protein.

1.1.3 A new kind of medicine

The ability to deliver genetic material into target cell types has wide applications for both basic science research and medicine, but the immediate implications of gene therapy for well-defined recessive diseases that currently lack any effective intervention are readily apparent. In the simplest case of AAV-mediated gene replacement therapy, a disease phenotype is caused by loss of function mutations to a single gene within a well-defined cell population. An AAV capsid is identified that efficiently transduces the target cell population and used to encapsidate an AAV genome harboring a transgene cassette that mediates expression of functional gene copies. The customized AAV vector is generated and delivered, thus curing the disease. Obviously, this description is incredibly oversimplified and does not take into account the critical nuances that must be considered when developing a gene therapy. Nevertheless, AAV-mediated gene therapies rapidly progressed to the clinic. As of 2019, AAV vectors were administered to over 3,000 patients across 149 clinical trials with no major safety issues reported (Kuzmin et al. 2021). Indeed, the rapid developments have been recognized by regulatory agencies and the US Food and Drug Administration has issued a statement that they are preparing to meet the demands of a predicted ~200 investigative new drug applications and 10-20 expected approvals per year by 2025 (Gottlieb 2019).

Clinical Successes

Hemophilia B is an X-linked bleeding disorder caused by loss of coagulation factor F.IX (Franchini and Mannucci 2012). Hemophilia B represents perhaps the most obvious use case for AAV gene therapy because it is caused by mutation to a single gene

that easily fits within AAV vectors. In addition, only small amounts of F.IX, on the order of 5% of normal serum levels, are sufficient to mediate a profound clinical effect and, in theory, a wide variety of cell targets can be transduced so long as they can mediate secretion of the factors into circulation (Mannucci and Tuddenham 2001). Thus, preclinical studies were carried out in a canine model of Hemophilia B, demonstrating that intramuscular administration of AAV2 vectors could mediate sustained therapeutic levels of Factor IX expression (Herzog et al. 1999). Subsequently, the first gene therapy clinical trial for Hemophilia B delivered an AAV2-cytomegalovirus intermediate early promoter (CMV)-F.IX at a dose of 2×10^{11} vg/kg intramuscularly and demonstrated safety, but limited efficacy that suggested dosing based on large animal studies did not translate to human patients (Kay et al. 2000). Further investigations found that systemically-administered AAV2 could also elicit a therapeutic response in hemophilic dogs (Mount et al. 2002). Therefore, another clinical trial was initiated, this time with a dose-escalation design up to 2×10^{12} vg/kg (Manno et al. 2006). No safety issues were reported and therapeutic levels of F.IX were achieved, but only for a limited time as patients developed a cell-mediated immune response against transduced hepatocytes. Years later, yet another clinical trial was carried out with six hemophilic patients, but using an optimized AAV vector with a self-complementary genome, liver-specific LP1 promoter-enhancer, AAV8 capsid, and codon optimized F.IX transgene (Nathwani et al. 2011). This time, therapeutic levels of F.IX were maintained for over a year, though patients in the high dose (2×10^{11} vg/kg) group required steroid treatment to control aminotransferase levels. To date, AAV-mediated gene therapy for Hemophilia B has not achieved regulatory approval, though evidence of clinical success continues to build.

Lipoprotein lipase deficiency (LPLD) is a rare autosomal recessive disease characterized by the loss of LPL and dysregulation of lipid metabolism leading to severely elevated levels of triglycerides and chylomicrons. Patients are at risk of acute pancreatitis and must adhere to strict low-fat diets. LPLD lent itself to a gene therapy approach because it was monogenic and the disease could easily be monitored by serum triglyceride levels that reach such high levels a visible lipemia can be observed. Pre-clinical proof-of-concept was established in a feline model of LPLD using an AAV1-CMV vector to express the LPL^{S447X} variant, which exhibits higher enzymatic activity, though long-term efficacy required an immunosuppressive regimen (Ross et al. 2006). The following clinical trial used a similar vector with a dose up to 1 x 10^{12} vg/kg delivered intramuscularly into 12 patients (Gaudet et al. 2013). No serious adverse events were reported, but the therapeutic effect on the defined primary outcome of plasmid triglyceride levels appeared to be transient, as was observed in the feline model. However, transgene expression was detected at 26 weeks post-vector administration in muscle biopsies taken from consenting patients. Additionally, patients self-reported improved quality of life. While tenuous given the subjective report in a small sample, investigators argued that plasma triglyceride levels were highly variable and not an appropriate predictor of efficacy. The European regulatory body agreed, leading to a follow-up trial demonstrating that the gene therapy mediated significant improvements in postprandial chylomicron metabolism, an important clinical outcome given the importance of postprandial chylomicron metabolism on acute pancreatitis risk (Carpentier et al. 2012). Ultimately, though by no means through a straightforward mechanism, Glybera (alipogene tiparvovec) received approval from the European

commission, thus establishing it as the first gene therapy to receive regulatory approval. Notably, the clinical efficacy of Glybera is not firmly established and would remain a matter of debate if Glybera had not been removed from the market due to its ultra-rare patient population and steep price tag that made it commercially non-viable. Regardless, Glybera can still be viewed as a clinical success in that it demonstrated sufficient safety data throughout the entire process, including vector manufacturing, to convince a regulatory body that AAV-mediated gene therapy is fit for clinical use.

Leber Congenital Amaurosis (Type 2) is one of a family of early-onset inherited retinal degenerations. The LCA2 subtype is defined as recessive loss of the retinal pigment epithelium-specific 65 kDa protein (*RPE65*) gene that encodes all-trans retinyl ester isomerase, a critical enzyme in the visual cycle, that leads to complete blindness and eventual photoreceptor degeneration (Pierce and Bennett 2015). Preservation of photoreceptors despite functional loss was an important finding for application of a gene therapy to this disease (Jacobson et al. 2005). Long-standing mechanistic studies coalesced into rapid pre-clinical development of an AAV-based gene therapy that was greatly facilitated by the discovery of spontaneously arising canine model of LCA2 (Veske et al. 1999; Bennicelli et al. 2008). Also critical to the success of retinal gene therapy was the non-trivial development and optimization of subretinal gene delivery (Xue et al. 2017; Vasconcelos et al. 2020). Subsequently, multiple groups initiated clinical trials utilizing subretinal delivery of AAV2 to mediate high efficiency transduction of *RPE65* into the retinal pigment epithelium (Bainbridge et al. 2008; Hauswirth et al. 2008; Maguire et al. 2008). Ultimately, this resulted in the first randomized Phase III clinical trial for an AAV-mediated gene therapy, utilizing bilateral

subretinal injection of 1.5×10^{11} vg of an AAV2-hRPE vector (chicken β-actin hybrid promoter with cytomegalovirus enhancer) into 20 patients (Russell et al. 2017). A functional assay of vision was developed called the multi-luminance mobility testing (MLMT), essentially providing a standardized means of scoring patient ability to navigate a maze at different luminance levels. Patients showed statistically significant improvement on the MLMT, corresponding to functional improvement navigating in low-moderate light conditions, as well as improvement in full-field light sensitivity thresholds. The data were compelling and led to the first AAV-mediated gene therapy to gain regulatory approval in the US, Luxturna (voretigene neparvovec-rzyl).

Spinal muscular atrophy (SMA) is an autosomal recessive neurodegenerative disease caused by mutation in the survival motor neuron 1 (SMN1) gene leading to degeneration of motor neurons and muscular atrophy, clinically presenting as progressive muscle weakness and failure to achieve milestones. Ultimately, SMA progresses to death by respiratory failure or placement of a permanent ventilator by the age of 20 months (Burghes and Beattie 2009). Pre-clinical development for an AAV-mediated gene replacement therapy began with observations that intravenously administered AAV9 vectors could cross the blood-brain-barrier and efficiently transduce lower motor neurons in the murine spinal cord (Foust et al. 2009; Duque et al. 2009). These observations were extended to large animals including cats and non-human primate (Duque et al. 2009; Foust et al. 2010; Gray et al. 2011). Neonatal administration of an AAV9 vector carrying a self-complementary chicken β-actin (CB)-driven SMN transgene cassette dramatically improved survival curves (from < 25 days to > 250 days) in an SMA mouse model compared to a control CB-GFP vector (Foust et al. 2010). The subsequent clinical trial

treated 15 patients with the severe form of SMA using systemic administration of the scAAV9-CB-SMN vector at a dose up to 2×10^{14} vg/kg (Mendell et al. 2017). Of 12 patients receiving the high dose, 100% were alive and did not require permanent ventilation at a median age of 30.8 months. In addition, all but one patient (treated at the latest timepoint of 7.9 months) achieved developmental motor milestones that were never observed in a historical cohort of SMA patients. The key safety concern was elevated serum aminotransferases suggesting hepatotoxicity. Therefore, serum aminotransferase levels were controlled by administration of steroids prior to AAV vector infusion. This study was a landmark in demonstrating the successful use of a single intravenous infusion of AAV vectors to treat a CNS disease and ultimately resulted in bringing Zolgensma (onasemnogene abeparvovec-xioi), the first systemically administered AAV-mediated gene therapy, to market. A benefit of gene therapy in this young population is that the total vector load (dosed on per weight basis) and likelihood of pre-existing anti-AAV immune response were both reduced.

These early clinical successes share several common features. First, investigators selected monogenic recessive diseases with an unmet clinical need and known causative genes that fit within the confines of the AAV packaging limit. Second, these diseases had well-described animal models to facilitate pre-clinical studies. Third, target cell types could be conveniently transduced by one of the naturally occurring AAV serotypes. While these features may seem obvious, it is important to note that many diseases do not meet these simple criteria and therefore face substantial barriers to development of an AAV gene therapy.

It is also important to note that the above-described clinical successes all describe a gene replacement strategy, the most direct use case for AAV-mediated gene therapy. This is reasonable since gene therapy remains in its infancy. However, other potential strategies that are being utilized or developed include delivery of small interfering RNAs to address dominant diseases, indirect gene therapy that does not target a specific causative gene (e.g., expression of an anti-inflammatory molecule), and gene editing techniques using CRISPR/Cas systems or base editing.

1.1.4   Challenges to AAV-mediated gene therapy

A common refrain in the field of AAV vector biology is that the AAV capsid did not evolve to be a gene therapy vector. Thus, despite the incredible and transformative successes that have been achieved in a relatively short period of time, it is important to note the specific shortcomings of AAV vectors.

First, AAV vectors have a highly constrained packaging limit of ~5 kb (Wu, Yang, and Colosi 2010). This is a significant barrier to application for diseases containing mutations to large genes that exceed this packaging limit, including Hemophilia A, Duchenne muscular dystrophy, cystic fibrosis, Usher Syndrome 1B, and Stargardt disease. Additionally, many applications take advantage of the transduction efficiency enhancement conferred by a self-complementary genome, which effectively halves the limit of the transgene cassette that can be included.

Second, pre-existing immune responses to the AAV capsid are prevalent (Calcedo et al. 2009; Boutin et al. 2010). Even minute levels of serum neutralizing antibodies can nullify transduction following systemic delivery of AAV vectors (Gray et al. 2011),

although other routes of administration may not be as susceptible to neutralization (Gray et al. 2013; Bennett et al. 2012).

Third, although a wide variety of AAV variants have been identified and characterized, a cursory look at current clinical applications will show a strong bias towards diseases of the liver, skeletal muscle, and CNS, including both the retina and spinal cord (Kuzmin et al. 2021). Chief among the reasons for this is that the AAV serotypes, while diverse, do not effectively transduce a variety of clinically-important tissues, including, for example, the kidney and lung. In addition, many specific cell types, even in tissues that are well-transduced, remain refractory to AAV-mediated transduction.

Fourth, AAV is dogmatically touted as being non-pathogenic. This is reasonable because no disease association has been observed despite high frequency natural AAV infections. In the recombinant context, the majority of recombinant AAV genomes are maintained as non-integrating episomal concatamers, which should limit the risk of insertional mutagenesis. Furthermore, the field had completed countless pre-clinical studies and thousands of patients were administered AAV vectors, again with no observable disease association (Chandler et al. 2016). However, evidence of hepatocellular carcinoma (HCC) following AAV administration in neonatal mice was observed as early as 2007 (Donsante et al. 2007). Another study determined that development of HCC was dose-dependent and closely associated with the *Rian* locus (Chandler et al. 2015). A very recent study further confirmed the development of HCC following administration of AAV vectors into neonatal, but not adult mice (Dalwadi et al. 2021). Intriguingly, however, the same study demonstrated that HCC was induced in

adult mice with liver injuries, such as fatty liver disease, again with an association with AAV integration into the *Rian* locus. Although the *Rian* locus has no human ortholog, insertional mutagenesis due to natural AAV2 infection was discovered in human HCC samples (Nault et al. 2015). In addition, a clinical trial patient was recently diagnosed with hepatocellular carcinoma following systemic delivery of an AAV5 vector encoding F.IX (UniQure 2021). This report should be viewed in light of the patient's pre-existing genetic predisposition to developing HCC, including a TP53 mutation, and failure to clearly demonstrate a link to AAV genome integrations. A causal relationship between AAV gene therapy and HCC has not yet been established, but the field is rapidly investigating this possibility. In the meantime, caution and close monitoring are warranted, particularly for high dose systemic studies.

Fifth, AAV is often described as exhibiting low immunogenicity. It is important to consider that "low" is a relative term that was probably used in reference to adenoviral vectors (Flotte 2020), the previous standard for viral gene therapy that ultimately elicited a severe immune response leading to the death of Jesse Gelsinger (Sibbald 2001). Indeed, decades of research evidenced that immune response against AAV administration were largely limited to humoral development of neutralizing antibodies and cell-mediated responses against transduced cells that were largely controllable by judicious use of steroids. However, the pursuit of greater therapeutic efficacy through administration of higher doses has corresponded to a recent upsurge in safety reports. Most notably, and unfortunately, three patients in the ASPIRO clinical trial of high-dose (3×10^{14} vg/kg) systemically-administered AAV8-mediated gene therapy for X-linked myotubular myopathy have died (Audentes 2020). Details are lacking, but one report (Wilson and

Flotte 2020) speculates that the toxicity resembles previous observations in non-human primates (Hinderer et al. 2018). In that study, elevated transaminase levels were noted, as well as acute liver failure and shock four days after vector injection (compared to 3-4 weeks in the clinical trial). Adding to this are reports that the clinically-approved drug Zolgensma elicits subacute liver failure (Feldman et al. 2020). Furthermore, Hinderer and colleagues observed inflammation of the dorsal root ganglia (DRG) and degenerating sensory neurons. A follow up study determined that DRG pathology was nearly universal across 33 pre-clinical AAV studies in non-human primate (Hordeaux et al. 2020).

Taken together, the latest pre-clinical and clinical observations clearly proclaim serious concerns regarding the safety of AAV-mediated gene therapy. Critically, however, toxicities observed to date are all dose-dependent, including the above-described genotoxicity. Descriptively, 3/17 patients in the high dose of the ASPIRO trial died, but no serious adverse events have been reported for 6 patients in the low dose group (a three-fold reduction to 1×10^{14} vg/kg). Indeed, a prominent AAV biologist has recently noted that adults are easily able to tolerate a dose of AAV particles equal to the number of cells in the body (Srivastava 2020). This conclusion was based on a ~70kg subject and dose of 4×10^{11} vg/kg that elicited no observed immune response (Manno et al. 2006). The same study showed that a 5-fold increase in dose resulted in a cell mediated immune response against transduced cells. It is, therefore, unsurprising that doses of AAV vector increased by nearly 1000-fold, are eliciting immune responses of concern.

In summary, the therapeutic index of AAV may be prohibitively small for certain disease applications. In order to achieve efficacy despite relatively low vector

efficiencies, the field has been aggressively pursuing the high doses of AAV vector described above. Although this has resulted in transformative therapy for at least one disease, it is apparent that we have approached the maximal dose that can be safely delivered. Rather than pursuing these high dose strategies further, an alternative approach is to generate enhanced gene delivery vectors that can achieve clinical efficacy at lower doses.

1.2   Engineering AAV vectors for next generation gene therapy vectors

Among the many features of recombinant AAV vectors that make them ideal for gene therapy, its ability to be easily manipulated stands out. Largely due to the foundation built by the above-described efforts to generate an AAV vector plasmid system, the sequence of virtually any part of the AAV genome can be manipulated and interrogated. Although early work utilized this system to develop capsid mutants for the purpose of studying basic AAV biology, recent work has focused on optimizing and engineering AAV vectors for gene therapy applications.

A seminal example came from efforts to engineer the AAV ITRs that resulted in the development of the self-complementary AAV vector (D. McCarty et al. 2003). Knowledge of AAV replication led to the hypothesis that one of the two terminal resolution sites of the AAV genome could be deleted. Doing so prevents DNA nicking and resolution of replication such that DNA synthesis continues all the way through one of the hairpin loop ITRs. The result is a genome containing inverted terminal repeats on each end as well as an inverted repeat of the transgene cassette separated by the mutated

ITR in the middle. After successful delivery to a cell nucleus, this configuration of vector genome can rapidly self-anneal and serve as substrate for transcription. This effectively bypasses the need for conversion to a double-stranded genome by either second strand synthesis or annealing of plus and minus strand, the rate limiting step of AAV-mediated transduction. While self-complementary vectors are a landmark advance in the field that played a pivotal role in achieving sufficient levels of transduction for clinical efficacy, they do suffer from two critical weaknesses: (1) the need to limit transgene cassettes to ~2.3 kb (half of the already limited capacity of AAV), and (2) an apparent potentiation of host innate immune response (Martino et al. 2011). Thus, engineering of optimal AAV vectors is a never-ending process.

In addition to the many methods of optimizing the genetic payload, it is important to recall that many in vivo properties are primarily dependent on the AAV capsid (Grimm et al. 2006) and that the use of different AAV capsid serotypes provides access to different in vivo properties (Hauck and Xiao 2003). Additionally, AAV vectors are no longer limited to the naturally occurring AAV capsid isolates. Indeed, the natural diversity of AAV capsids strongly suggested that the AAV capsid would be highly tolerant to sequence variation and, despite the complexity of overlapping reading frame within the *cap* gene, it has now been demonstrated that the capsid is highly amenable to capsid engineering strategies. Engineering of the AAV capsid can largely be organized into two approaches: rational design and directed evolution. The rational design approach utilizes a fundamental understanding of capsid structure-function relationships that allow specific phenotypes to be programmed into novel vectors. In contrast, the directed evolution approach involves screening of genetically diverse libraries of AAV variants.

1.2.1 Rational Design

One approach to capsid engineering is rational design, whereby knowledge of AAV vector-host interactions and structure-function relationships are used to predict modifications that will enhance performance. In the quintessential example, observations that tyrosine residues on the AAV capsid are phosphorylated and marked for proteasomal degradation by ubiquitination led to the prediction that removal of phosphorylatable hydroxyl groups by site-directed mutagenesis of tyrosine residues to phenylalanine (so called Y-F mutations) would prevent proteasomal degradation and enhance transduction (Zhong et al. 2008). As a result, the authors demonstrated up to 30-fold greater transduction of hepatocytes in mice, in addition to a 10-fold lower dose requirement for achieving therapeutic levels of F.IX.

However, it is important to recall that AAV-mediated transduction is a complex and multi-step process (Figure 1.3), and that AAV entry to the cell is necessary, but not sufficient, for functional transduction. Case-in-point, a recent report suggests that failure of the naturally-occurring AAV8 capsid to transduce microglia, a long-sought after goal, is due to a post-entry mechanism that results in loss of AAV genomes before they can enter the nucleus and become transcriptionally active (S. K. Wang et al. 2020). Interesting observations like this can potentially lead to hypotheses and a rationally designed capsid to overcome the barrier. However, reasonable hypotheses depend on pre-existing knowledge of basic AAV biology in the host cell and also capsid-structure relationships. Despite regular advances, the field still lacks fundamental understanding on both fronts (Dudek et al. 2019; Adachi et al. 2014). Indeed, a rational design approach is

challenging for addressing many desirable vector phenotypes. Therefore, an alternative approach that is not predicated on pre-existing knowledge is required.

1.2.2 Directed evolution

Directed evolution is one such agnostic approach, invented by Frances Arnold and Willem Stemmer. Arnold shared the Nobel Prize in Chemistry in 2018, while Stemmer passed in 2013. Directed evolution is essentially the same process used by artificial selection strategies that breed dogs small enough to fit into tea cups, but does so at the molecular level, which allows for drastically accelerated timelines. In brief, modern molecular tools allow us to easily introduce genetic diversity into a protein of interest. Proteins are screened for some desirable trait. These selected proteins are then used as a parental population to generate a new pool of genetic diversity. The cycle is repeated until the desired result is achieved (Arnold 2018). A related approach is biopanning of phage display libraries, which involves the initial generation of a library that undergoes multiple rounds of screening for the ability to bind some substrate. Notably, as will be discussed, efficient screening of AAV libraries involves more than identifying capsids that bind to a target since functional transduction requires many post-attachment steps (Figure 1.3). Therefore, screening AAV libraries is not strictly defined as biopanning, which focuses on identifying binding partners. Furthermore, recent advances in screening AAV libraries have clearly demonstrated that diversification does not need to be introduced after each round of selection (discussed below). Therefore, screening AAV libraries is not strictly defined as directed evolution, which iteratively generates diversity and screens the novel variants. Thus, the field of AAV capsid engineering often uses

these terms interchangeably. Here, we similarly apply both terms to broadly encompass any library-based screen of AAV variants and organize the process into three steps: (1) library generation, (2) library selection, and (3) validation.

1.2.2.1 Library generation

DNA library

Genetic diversity must first be introduced into the *cap* open reading frame encoding the VP proteins. This is greatly facilitated by modern tools in molecular biology and the availability to manipulate the AAV sequence in plasmids. Perhaps the simplest means of introducing such diversity is to amplify the *cap* gene with an error-prone DNA polymerase. Doing so will introduce point mutations into the capsid sequence. Importantly, the rate of mutation can be controlled (Perabo et al. 2006; Maheshri et al. 2006). Too many mutations may increase the number of non-viable variants. Too few mutations, decreases the likelihood of generating a successful candidate. Although single amino acid mutations to the capsid that mediate drastic phenotypic differences have been identified (Wu et al. 2006; S. J. Huang et al. 2017), the likelihood of discovering one through random mutagenesis via an error-prone PCR method is incredibly low given the sequence search space that is possible. The probability can be drastically improved by applying elements of rational design to target specific regions of the AAV capsid. For example, focusing mutations on the surface-exposed three-fold protrusions led to the identification of liver-detargeted AAV9 variants that exhibit attenuated biodistribution to the liver (Pulicherla et al. 2011).

Despite its simplicity, error-prone PCR was not the first strategy used to generate a library for directed evolution of a viral vector. In 2000, the Stemmer group attempted "molecular breeding" of murine leukemia viruses to redirect in vitro tropism (Soong et al. 2000). In brief, their method pooled 6 different murine leukemia virus genomes, applied a DNase I treatment to fragment the genomes, then reassembled fragments using PCR. Amazingly, despite all 6 parents failing to infect CHOK1 cells, the study successfully identified chimeric variants with the ability to infect CHOK1. In contrast to generating point mutations that may have limited function, DNA shuffling potentially facilitates recombination of entire functional domains. Thus, in 2008, three independent groups applied the DNA family shuffling strategy to AAV *cap* (Grimm et al. 2008; Koerber, Jang, and Schaffer 2008; W. Li et al. 2008)*.* All three studies demonstrated the effectiveness for generating AAV capsid libraries. Ultimately, one effort led to the identification of AAV-DJ, which demonstrated excellent liver transduction efficiency and the ability to evade neutralizing antibodies (Grimm et al. 2008)

One limitation of the randomized approaches described above is that a large fraction (majority) of variants is non-viable. Thus, a third approach to generating AAV capsid libraries is in silico design. Although this can take a variety of forms, one example that addresses the issue of DNA shuffling generating non-viable variants is the use of structure-guided SCHEMA to design a high diversity AAV library with high predicted viability of assembled capsids (Ojala et al. 2018). This approach identified the SCH9 variant that is capable of transducing neural stem cells in the subventricular zone.

Another successful approach for in silico design is the generation of libraries of putative AAV ancestors using ancestral sequence reconstruction methods. Ancestral

AAVs were proposed as a potential starting point for directed evolution given that all of the divergent phenotypes observed in modern AAVs should be derived from ancestors (Santiago-Ortiz et al. 2015). A noteworthy success was achieved by the Vandenberghe group, who reconstructed AAV ancestors based on 75 modern AAV variants that were able to infect primates and selected a distal node between AAV2 and AAV8, called Anc80. The authors then generated a library of 2,048 variants based on 11 low confidence positions in the Anc80 sequence that each had two possibilities. Screening of the library identified Anc80L65, which exhibited enhanced transduction of skeletal muscle following intramuscular injection, enhanced transduction of the outer retina following subretinal injection, and some evidence of enhanced liver transduction in non-human primate following systemic injection (Zinn et al. 2015). A follow-up study determined that Anc80L65 was also able to efficiently cross the blood-brain-barrier and transduce the murine brain (Hudry et al. 2018).

Finally, a commonly used method of capsid library generation inserts a short peptide sequence on the capsid surface, the so-called AAV peptide display. The key advantage of this approach is its simplicity. Library generation is simple, as is selection because the peptide is at a pre-defined location and is of limited length, thus avoiding the need to sequence the entire *cap* sequence. A further advantage of this method is that common insertion sites for the peptide do not disrupt the overlapping AAP or MAAP reading frames, whereas the previously discussed library generation methods may benefit from providing these functions (in particular AAP) in trans during production. The earliest demonstrations of this approach focused on redirecting the in vitro tropism of the AAV2 capsid, which is highly dependent on its heparan sulfate proteoglycan-recognizing

domain (Kern et al. 2003). Thus, 7mer peptide insertions were introduced adjacent to the AAV2 R588 residue that disrupted the strong heparan sulfate interaction. Screening of these libraries successfully identified mutants with drastically modified tropisms (Perabo et al. 2003; Müller et al. 2003). While it was previously held that peptide insertion displays were most likely to only alter tropism and not drastically effect transduction (Grimm and Zolotukhin 2015), recent successes clearly show that this is not the case. In a seminal study, a 7mer peptide display library was generated on the AAV9 capsid and the library was screened to identify variants that could cross the blood-brain-barrier. Strikingly, the AAV-PHP.B variant was identified that mediated a dramatic 40-fold greater transduction of the murine brain than the parental AAV9, mediated only by a TLAVPFK insertion on the capsid surface (Deverman et al. 2016).

Importantly, generation of the AAV plasmid library is a limiting step with respect to the diversity of AAV variants that can be sampled. Although the diversity of a simple library, such as one containing a randomized 8mer is theoretically on the order of 10^{10}, the maximum diversity of generated plasmid libraries appears to be on the order of 10^{7}-10^8 (Deverman et al. 2016; Nonnenmacher et al. 2020).

AAV Library

Conversion of the plasmid library to an AAV capsid library is a critical but poorly understood step. Subsequent screening methods rely on a high genotype-phenotype correlation, that is, the AAV capsid should package its own genome. Loss of the genotype-phenotype relationship would render any screening effort meaningless since recovered sequences would not correspond to desirable variants. To date this is most

commonly achieved by providing limiting amounts of the *cap* library during vector production by plasmid transfection. The goal of this approach is to avoid delivering multiple copies of *cap* into a single cell as doing so can result in the generation of AAV capsids that package the wrong genome (cross-packaging) or heterogenous capsids made up of different VP monomers (mosaicism).

Intriguingly, early studies demonstrated an apparent disregard for this potential pitfall, yet achieved successful selection outcomes (Perabo et al. 2003; Grimm et al. 2008). A key study was carried out to investigate how this could be the case. Clever experimental design allowed the investigators to identify a propensity for capsids to package their own genomes, so long as *cap* was in cis to the ITRs (as it is in directed evolution experiments). Clear evidence of capsid mosaicism was observed if *cap* was provided in trans to the ITRs during vector production (as it is in most recombinant production systems) (Nonnenmacher et al. 2015). Although the rates of both cross-packaging and mosaicism were dramatically less than what might be expected due to chance, they were not absent. The authors went on to define that a transfection utilizing ~5,000 plasmid copies per cell would be the optimal tradeoff for maximizing titer while maintaining a high genotype-phenotype correlation. Recently, these findings were largely corroborated by another group (Schmit et al. 2020). While their findings were largely consistent, including the conclusion that limiting amounts of plasmid could effectively increase library genotype-phenotype correlation, they also found that rates of mosaicism were much higher than previously reported. However, mosaic capsids had a high likelihood of packaging the genome that encoded the majority of its capsomers, which might still allow for successful phenotypic selection.

1.2.2.2 Selection

After the AAV capsid variant library has been generated, it is subjected to selective pressure and successful variants must be recovered. This pressure can be as simple as an affinity column to identify AAV variants that bind a certain molecule, such as heparin (Maheshri et al. 2006). Alternatively, the library can be applied in vitro to select for cells targeting a specific cell type (Michelfelder et al. 2009). Depending on the goals of the study, this may be suitable. However, it is important to note that due to the complexity of AAV transduction, the identification of AAV variants with enhanced in vivo properties should be selected in that context and we focus on this goal.

Approaches to recovery rely on amplification techniques (Figure 1.4). Originally, AAV variants carrying a wild-type genome configuration could be screened for successful transduction by superinfection with adenovirus to elicit replication of infectious particles that could then be directly harvested (Figure 1.4a). An inherent limitation of this approach is that it requires that adenovirus be able to transduce the target cell type, ideally with high efficiency so that there is a high probability that most cells containing AAV variants will be superinfected. While this approach is often viable in vitro, it becomes much more challenging in vivo. Nevertheless, the first report of in vivo directed evolution was carried out using this approach to screen an AAV-DJ peptide display library in murine lung (Grimm et al. 2008). A second successful application of this approach came from the same group, applying an AAV variant library to chimeric human-mouse livers (FRG mice engrafted with human hepatocytes) and rescuing successful variants by adenovirus superinfection, ultimately leading to the identification

of AAV-LK03 (Lisowski et al. 2014). It is important to note that the successful application of this approach in vivo is limited to tissues that are well-transduced by adenovirus. On the other hand, it is notable that the adenovirus rescue method for recovering sequences offers a stringent pressure and selects for AAV variants that were able to achieve all of the steps of the AAV life cycle.

Subsequently, the field overcame the issue of limited adenovirus tropism by rescuing capsid variant sequences using PCR-based amplification (Figure 1.4b). For example, adenovirus is not efficient for transducing photoreceptors (Sweigard, Cashman, and Kumar-Singh 2010), yet the PCR-based amplification approach opened the door to successful application of a directed evolution approach to identify AAV-7m8, which is capable of transducing the outer retina following intravitreal administration in mice (Dalkara et al. 2013). The move to PCR-based amplification, however, was not without its disadvantages. Most notably, the highly sensitive recovery method facilitates recovery of any AAV genomic DNA extracted from the tissue. This includes ineffective capsid variants that mediate some, but not all, of the steps of transduction. Recovery of such variants increases the level of noise in screening technologies and makes it difficult to identify the variants that are truly enhanced (Körbelin and Trepel 2017).

Thus, the next step was to develop a directed evolution platform that would effectively limit the amount of background sequences by stringently selecting only those variants that were able to achieve a functional level of transduction. Functional transduction is here defined as AAV-mediated expression of the transgene. In a landmark study, the AAV genome was reconfigured to contain a double-floxed primer binding site (Figure 1.4c). An AAV library generated with this vector genome configuration could

then be administered to transgenic rodents that expressed Cre-recombinase (Cre) within a promoter-enhancer-specified population of cells. AAV capsid variants that successfully transduced cells expressing Cre would undergo irreversible inversion of the double-floxed primer binding site. Subsequent PCR-based recovery of capsid sequences could selectively amplify sequences that had undergone inversion of the primer binding sequence. The design effectively removes non-functional AAV variants because Cre-recombinase is only able to catalyze inversion of double-stranded DNA; therefore, only functional AAV variants that were able to overcome the rate-limiting step of achieving the double-stranded DNA form would be recovered.

This clever system was termed CREATE (Cre-recombination based AAV Targeted Evolution) and led to the successful identification of AAV-PHP.B, as described above (Deverman et al. 2016). Further evolution of AAV-PHP.B using CREATE yielded an enhanced version, AAV-PHP.eB, which is ~100-fold more efficient than AAV9 in the C57BL/6 central nervous system (Chan et al. 2017).

It should also be emphasized that the incorporation of a cell type-specific mechanism of selection is a major technological advance for performing selection in heterogeneous tissues. For example, the aforementioned selection of AAV-7m8 required the additional use of transgenic animals expressing GFP in photoreceptors and subsequent fluorescence-activated cell sorting to enrich for their target cell population. In contrast, CREATE was able to perform their screen using bulk tissue, though still requiring a transgenic animal. Inclusion of a cell type specific mechanism of enrichment in the technology is a significant advance that should not be overlooked and is likely to facilitate the selection of AAV variants targeting rare cell populations in the future.

Notably, the system was adapted by another group and again successfully applied to identify the SCH9 variant, described above (Ojala et al. 2018). Thus, a new generation of Cre-dependent selection strategies was born that demonstrated remarkable potential compared to the previous technologies that indiscriminately recovered all viral genomic DNA. Nevertheless, Cre-dependent selection was not without its limitations.

The first limitation is the reliance on Cre-driver lines to efficiently express Cre-recombinase in specific cell populations. A wide variety of Cre-driver lines are now available that should facilitate Cre-dependent selection schemes for a wide variety of specific cell types. However, if the Cre-driver line does not already exist, investigators must generate it themselves, spending considerable effort, resources, and time. Alternatively, and in theory, Cre-recombinase can be co-delivered as a separate vector. The limitations of this approach are multitude including: (1) need for a vector that can already transduce the target cell type, (2) potential issues in achieving highly efficient and cell type-specific expression of Cre-recombinase, and (3) exponentially lower hits since both vectors must transduce the same cell.

Second, the reliance on Cre-driver lines effectively limits the application of Cre-dependent selection to rodents. While less of a concern for the development of research tools, this serves as a potential barrier to the development of AAV vectors suitable for clinical gene therapy applications. This is because directed evolution approaches are context-specific. That is, assuming that the selection system correctly identifies the best variant within a library, "best" is defined only with respect to the precise conditions under which the variant was selected. Directed evolution itself provides no insight into the mechanisms that underlie the selected variant's enhancement, which may require the

highly specific conditions of selection, or, by luck, may generalize more broadly. Of note, the current literature offers only a few reports AAV variants that demonstrated enhancement in rodents and were subsequently evaluated in non-human primate, which likely suggests a publication bias. Among these, there is a notable failure of enhanced properties to translate, suggesting that species-specific differences can often serve as a major barrier to the identification of clinically-relevant AAV variants using rodent models (Dalkara et al. 2013; Hordeaux et al. 2018).

This has recently been highlighted by reports that AAV-PHP.B, which was evolved in C57BL/6 mice, does not mediate enhanced transduction efficiency in non-human primates, or even BALB/c mice (Matsuzaki et al. 2018; Hordeaux et al. 2018; Matsuzaki et al. 2019). Multiple groups took advantage of the differential performance of AAV-PHP.B in C57BL/6 and BALB/c mouse strains to determine that enhanced performance was dependent on lymphocyte antigen 6 complex (LY6A) expression on brain microvasculature endothelial cells (Hordeaux et al. 2019; Q. Huang et al. 2019; Batista et al. 2020). In addition, LY6A has no primate homolog, explaining the failure to translate. Thus, the highly successful directed evolution approach identified an optimal solution for the system that the experiments were carried out in (i.e., C57BL/6 mice). In this case, however, serendipity did not favor a translational outcome.

Third, CREATE selects for AAV variants that successfully achieve conversion to a double-stranded DNA form. This is highly effective because it overcomes the rate limiting step in AAV-mediated transduction (D. McCarty et al. 2003). However, it falls short of selecting at the level of transduction (i.e., transgene expression). Given recent evidence that the capsid plays a role transgene expression (Salganik et al. 2014; Aydemir

et al. 2016; Powell, Samulski, and McCown 2020; Bohlen et al. 2020), there may be an advantage to developing expression-based selection platforms.

To address this limitation, investigators designed a novel directed evolution platform, iTransduce, that offered selection at the level of vector-mediated protein expression (Hanlon et al. 2019). In essence, the authors designed the AAV library vector to include a Cre-expressing transgene cassette (Figure 1.4d). Administration of the library into Ai9 mice that harbor a double-floxed stop upstream of a tdTomato gene would lead to reporter expression only in cells that were successfully transduced by an AAV variant. Subsequent isolation by fluorescence-activated cell sorting allowed for recovery of enhanced variants. This approach resulted in the identification of AAV-F, which performed similarly to AAV-PHP.B in C57BL/6 mice and also demonstrated a similar level of enhancement in BALB/c mice.

This approach was successful and demonstrated the ability to screen AAV capsid variant libraries at the level of protein expression. Interestingly, the authors used an identical library design (7mer insertion at AAV9 VP1 588) to that used in the CREATE report, and report that overall performance of AAV-PHP.B and AAV-F exhibit somewhat comparable performance. This perhaps intimates a lack of clear benefit to the expression-dependent iTransduce system over the CREATE system, though the theoretical advantage remains.

In addition, it is important to recognize that the iTransduce system is still dependent on transgenic animals, and the corollaries of that are described above. Additionally, the iTransduce system may suffer from two new limitations. First, the use of fluorescence-activated cell sorting may serve as a bottleneck for library recovery,

whereas the CREATE system could utilize bulk tissue analysis. Second, there is a

theoretical concern that multiple AAV variants may enter a single cell. Even if the vast

majority of variants fail to mediate expression, one successful variant expressing Cre in

the cell will lead to recovery of all the defunct capsid sequences.
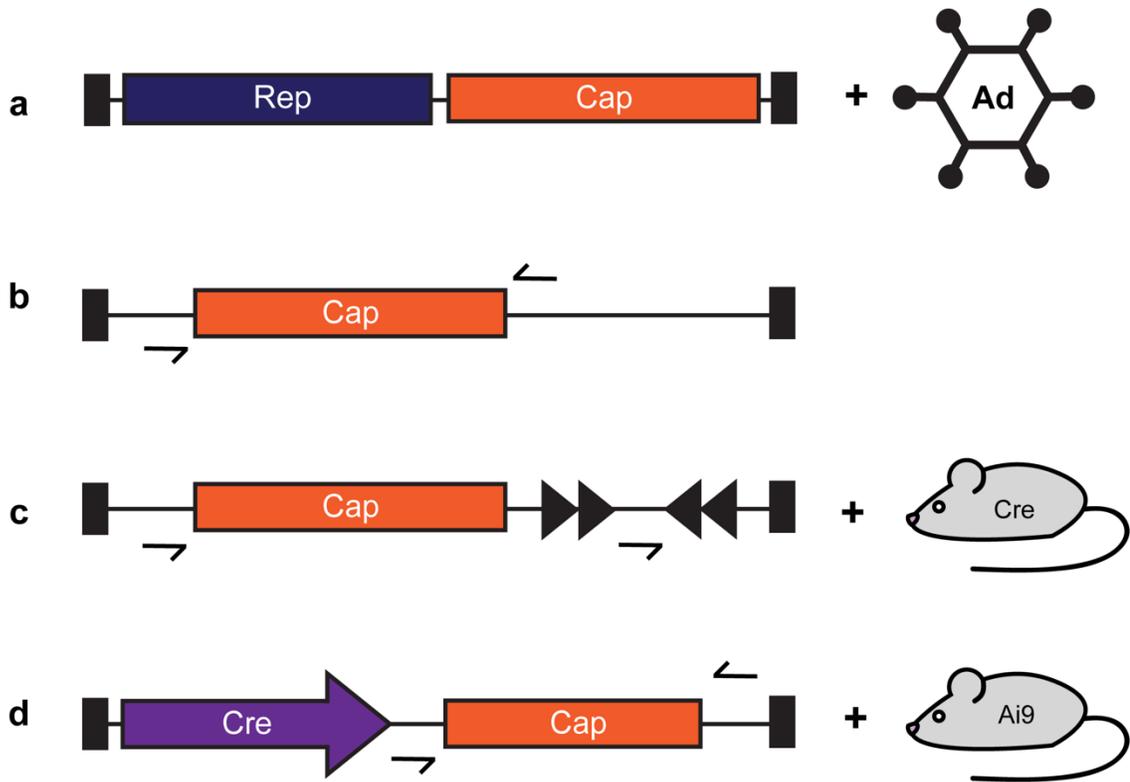
Figure 1.4 Selection strategies utilized for AAV capsid directed evolution
(a) A wild-type AAV configuration facilitates rescue of successful variants by Adenoviral superinfection. (b) The development of AAV Cap recovery by PCR rescue allowed for directed evolution without the need for Adenovirus, but with far less selection stringency. (c) The inclusion of a double-floxed primer binding site coupled with the use of Cre-driver animals allowed for PCR rescue of AAV Cap while stringently selecting for AAV variants that successfully mediated functional transduction, defined as achieving a double-stranded genome that could undergo Cre-mediated inversion of the primer binding site. (d) Selection for variants at the level of AAV vector-mediated protein expression was accomplished by inclusion of a Cre-recombinase transgene cassette in the AAV genome coupled with Ai9 reporter mice that express TdTomato in cells that are successfully transduced.

1.2.2.3 Validation and characterization

Identified variants must undergo a final validation due to the inherent noise of directed evolution approaches that may lead to false positives. For example, AAV-F was actually the second most frequently recovered capsid sequence reported in the iTransduce study. The most frequently recovered sequence, named AAV-S, actually demonstrated no enhancement compared to the AAV9 parent (Hanlon et al. 2019).

The classic approach to evaluating an AAV variant involves the generation of individual vector preparations packaged with identical reporter genomes. Most commonly, fluorescent reporters are used, though the Luciferase system can be useful in some contexts. The vectors are then administered into individual mice, each mouse is evaluated separately, and statistics are applied determine the effect. As the number of variants to be evaluated increases, the approach becomes incredibly time-consuming and resource-intense (Lochrie et al. 2006). In addition, achieving sufficient biological replicates is incredibly challenging in large animal models due to individual-to-individual variation that can preclude meaningful comparisons of vector performance (Gray et al. 2011).

Although these classic methods remain standard, the field is increasingly employing high throughput phenotyping approaches, such as AAV barcoding (Adachi et al. 2014; Marsic, Méndez-Gómez, and Zolotukhin 2015; Pekrun et al. 2019; Weinmann et al. 2020). The AAV DNA Barcode-Seq system was pioneered by the Nakai lab and allows for simultaneous characterization of hundreds of AAV variants in a small number of animals (Adachi et al. 2014). In essence, hundreds of AAV clones comprising a unique combination of AAV capsid and barcoded genome can be pooled and administered into a

small number of animals. Subsequent analysis can utilize next-generation sequencing to measure barcode reads as a surrogate for AAV capsid performance. Importantly, the analysis focuses on quantifying phenotypes relative to a reference control included in the library (e.g., AAV9). This allows for powerful within-subjects comparisons that can then be shown to generalize across biological replicates. Furthermore, transcription of the barcode allows for analysis to be carried out at the level of transgene expression, providing a measure of relative transduction efficiencies (AAV RNA BC-Seq, Adachi & Nakai, unpublished) that can be compared to relative biodistributions (AAV DNA BC-Seq). Such approaches allow for high confidence conclusions regarding the relative performance of AAV variants.

Admittedly, a non-trivial effort is required to establish AAV barcoding within a lab, including the generation and authentication of reagents and bioinformatic pipelines. Indeed, AAV Barcode-Seq method relies on manual association of known barcode sequences with specified capsid sequences through independent production of each vector, though purification is a pooled process (Adachi et al. 2014). Recently, the BRAVE (Barcoded Rational AAV Vector Evolution) system was reported (Davidsson et al. 2018). Though described as a system for directed evolution, it is better described as an AAV barcoding system for high-throughput phenotyping in the context of the framework presented here. This is because throughput is limited to hundreds of variants, whereas the directed evolution selection strategies above typically screen a diversity on the order of 10^6-10^7 (Nonnenmacher et al. 2020; Deverman et al. 2016). Nevertheless, BRAVE presents a methodology for multiplexing the generation of barcoded AAV plasmid libraries, resulting in ~50 unique barcodes associated with each capsid variant following

a single cloning reaction. BRAVE plasmid libraries can be directly used for vector production and the capsid-barcode relationship is established post-hoc. It remains to be seen if this technique will be embraced by the field.

A major limitation of high throughput AAV barcoding approaches is that they are currently limited to bulk tissue analysis. The development of a single cell AAV barcoding strategy has not yet come to fruition, but remains a possibility in the future. As such, use of a ubiquitous promoter to drive barcode expression, as originally described in the AAV Barcode-Seq system, cannot distinguish different cell types. If there is only one cell type of interest, then a cell type specific promoter can be utilized. If many cell types need to be characterized, then it is feasible to generate multiple libraries, each using a cell type specific promoter-enhancer, and multiplex the experiment, but this has not yet been reported. Relatedly, bulk analysis cannot distinguish between high levels of expression in a few cells, vs low level expression in many cells. Although some applications may benefit from high expression in a limited number of cells, most require a meaningful level of transduction in many cells.

Therefore, the best approach for directed evolution of the AAV capsid may require both modalities. Following, library generation and selection, many candidates can be selected for high-throughput characterization using AAV barcoding. Subsequently, a select few can be more thoroughly characterized using classic approaches in single animals.

1.2.3 Conclusion

To summarize, engineering the AAV capsid is a promising approach for developing next-generation gene delivery vehicles. In particular, directed evolution is a powerful technique for developing enhanced AAV vectors despite lacking a comprehensive understanding of structure-function relationships. However, the black box process is still being optimized. Libraries must be generated that contain a solution and maintain genotype-phenotype relationships. Selection strategies are prone to identifying false positives or successful candidates with properties that do not generalize outside the precise specifications of the screen. Validation techniques are rapidly improving, but multiple techniques are still required. Regardless of these limitations, the potential of this strategy for identifying enhanced vectors has clearly been demonstrated and it is only a matter of time before translational milestones are achieved.

2 Chapter 2. Development of a transcription-dependent directed evolution system

2.1 Abstract

Adeno-associated virus (AAV) vectors are utilized for safe and efficacious in vivo gene therapy for a wide variety of diseases. However, broader application will require development of next-generation AAV capsids that demonstrate enhanced properties compared to currently available vectors based on naturally-occurring isolates. High-throughput AAV capsid evolution techniques are highly successful for the identification of improved capsid variants in rodents, but robust translation to a clinically-relevant model is lacking. In order to maximize the probability of developing a clinically-relevant vector, evolution should be carried out in a clinically-relevant context, but current state-of-the-art technologies cannot accomplish this. Therefore, we developed the transcription-dependent directed evolution (TRADE) system that mediates the recovery of the AAV vector *cap* sequence from cell type-specific vector genome-mediated mRNA expression. This system should facilitate stringent recovery of highly functional AAV *cap* variants without any restrictions on the experimental host.

2.2   Introduction

Recombinant adeno-associated virus (AAV) is now an established vector for safe and efficacious in vivo gene therapy (Nathwani et al. 2014; Russell et al. 2017; Mendell et al. 2017). Three in vivo therapies have now achieved regulatory approval: Glybera for lipoprotein lipase deficiency (Gaudet et al. 2013), Luxturna for Leber congenital amaurosis (Russell et al. 2017), and Zolgensma for spinal muscular atrophy (Mendell et al. 2017). All approved AAV therapies, and the vast majority of clinical trials, take advantage of naturally-occurring AAV serotypes that serendipitously mediate efficient gene transfer to target cell types. However, for many gene therapy applications, gene delivery remains a key barrier and enhanced vectors are needed.

Despite recent advances moving the field toward a comprehensive understanding of AAV capsid-phenotype relationships (Adachi et al. 2014), we remain unable to rationally design AAV capsids that meet the multifactorial criteria for clinical utility. Instead, the field turned to high throughput screening of AAV capsid variant libraries to identify novel AAV capsids with enhanced gene therapy properties, so-called directed evolution of the AAV capsid.

The earliest attempts at *in vivo* library selection relied on recovery of vector genome DNA from crudely dissected tissue. Theoretically, this strategy results in recovery of both effective AAV variants, as well as AAV variants that mediate some, but not all of the steps required for vector-mediated transgene expression, potentially leading to a high background recovery of AAV variants that are completely ineffective gene therapy vectors. Furthermore, targeting a specific cell type required further processing, such as fluorescence-activated cell sorting or laser capture microdissection. Nonetheless,

there have been several reports of successfully employing this technology (Grimm et al. 2008; Excoffon et al. 2009; Dalkara et al. 2013; Lisowski et al. 2014).

Efforts to engineer the AAV9 capsid for enhanced brain transduction culminated with the landmark development of Cre-dependent selection methods. Cre-recombination based AAV targeted evolution (CREATE) was successfully applied to identify the AAV-PHP.B capsid, which mediates 40-50 folder greater gene delivery than AAV9 in mice (Deverman et al. 2016). CREATE highlights that even simple libraries consisting of a peptide insertion on the AAV capsid surface can yield high-performing variants if stringent selection methods are applied to accelerate the directed evolution process. Although the field rapidly attempted to show clinical relevance of the AAV-PHP.B capsid, it soon became apparent that AAV-PHP.B was no better than the parental AAV9 serotype in the non-human primate context (Matsuzaki et al. 2018; Hordeaux et al. 2018). Further studies found that AAV-PHP.B utilized a mechanism of enhancement that was specific not only to rodents, but even particular mouse strains (Hordeaux et al. 2019; Q. Huang et al. 2019; Batista et al. 2020), thus highlighting a key limitation of AAV directed evolution technologies – namely, the potential to select for non-translatable mechanisms of enhancement in rodents.

In order to address this capsid evolution was pursued directly in clinically-relevant non-human primates (Byrne et al. 2020). However, these attempts were met with limited success, requiring six rounds of in vivo selection, due to the non-stringent recovery of AAV genomic DNA. Indeed, use of the more stringent Cre-dependent selection was intractable in non-human primates due to the requirement for transgenic animals that are only readily available in rodents. Thus, there remained a critical need in

the field of gene therapy – the ability to perform directed evolution experiments in large non-human primate models while enacting stringent/functional levels as well cell type specificity. Here, we describe one such system to accomplish this task based on transcription-dependent directed evolution of the AAV capsid (TRADE).

2.3   Methods and Materials

Plasmids

We first constructed the pAAV9-N272A-2BsaI-hSYN1-TRADE construct that contains two expression cassettes oriented antisense to each other and flanked by the AAV2 inverted terminal repeats (ITRs). On one side of the genome, the AAV2 p40 promoter and intron is followed by the AAV9 *cap* gene and AAV2 poly-adenylation sequence. The AAV9 cap sequence was modified by (1) an N272A mutation that confers liver detargeting (S. J. Huang et al. 2017), (2) silent mutations to facilitate cloning as previously described (Adachi et al. 2014)(Appendix 6.1) and (3) a 42 bp substitution at position Q588 of the AAV9 VP1 ORF that encodes for 3 frameshifted stop codons flanked by BsaI restriction sites and followed by a GGGGS linker sequence. On the other side of the genome and antisense to the p40-driven cap ORF, we designed an overlapping transgene cassette consisting of a 471 bp human synapsin I (hSYN1) promoter-enhancer sequence, a 92 bp minute virus of mice (MVM) intron, the non-coding antisense sequence, and an SV40 poly-adenylation sequence placed in the AAV intron such that it avoided the predicted splicing branch point as determined using the Human Splicing Finder (Desmet et al. 2009). The pAAV9-N272A-2BsaI-hSYN1-GFP-TRADE construct is similar, but additionally contains a PCR-generated, codon-modified enhanced GFP

(GFP) sequence in the hSYN1-driven transgene cassette, anti-sense to the *cap* open reading frame (Figure 2.1).

For the proof-of-concept study, we generated the pAAV9-N272A-PHP.B-hSYN1-GFP-TRADE plasmid by synthesizing the AAV9-PHP.B cap sequence (Genscript) and using simple restriction enzyme cloning to insert the TLAVPFK sequence into the pAAV9-N272A-2BsaI-hSYN1-GFP-TRADE backbone.

We generated the non-splicing (NS) constructs by using a human codon optimization of the Cap open reading frame (sense) at the regions corresponding to the identified splice acceptor and splice donor (antisense). Relative to the pAAV9-N272A-PHP.B-hSYN1-GFP-TRADE Cap open reading frame, the NS1 construct contain 13 mutations covering the antisense splice acceptor, the NS2 construct contain 11 mutations covering the antisense splice donor, and the NS3 construct contain both sets of mutations.

The pHLP-Rep plasmid was derived from a non-ITR-containing plasmid expressing the AAV2-derived Rep and Cap genes with modifications to the p5 promoter to limit expression of the large Rep proteins (Grimm et al. 2003). A large deletion was made to the Cap gene so that only a small fragment of the VP1 protein is expressed. This plasmid was used to supplement Rep2 functions during production of TRADE vectors and has been previously described (Powers et al. 2018). pRepCap9 is a non-ITR-containing plasmid that expresses AAV2 Rep and AAV9 Cap. The plasmid contains a number of silent mutations to the Cap open reading frame.

Cell culture

Human embryonic kidney (HEK) 293 cells (Stratagene) were maintained in
Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine
serum (Sigma-Aldrich), 1% L-Glutamine (Lonza), and 1% Penicillin-Streptomycin
(Lonza). Neuro-2a (N2a) cells (ATCC) were maintained in Eagle's Minimum Essential
supplemented with 10% fetal bovine serum (Sigma-Aldrich), 1% L-Glutamine (Lonza),
and 1% Penicillin-Streptomycin (Lonza). Transient transfection was carried out using
polyethyleneimine (PEI) at a DNA:PEI mass ratio of 1:2 and 2 µg of total DNA per well
in a 6-well format.

Images of GFP expression were acquired two days post-transfection using a
Nikon Eclipse Ts2 microscope equipped with CoolSNAP DYNO camera (Photometrics)
and FITC filter set. RNA was harvested directly by adding 1 mL TRIZOL to each well.

AAV Vectors

We produced AAV vectors by adenovirus-free triple transfection of HEK293 cells
in T225 flasks using PEI. Transfection was carried out using a DNA:PEI mass ratio of
1:2, a total of 45 µg DNA per flask, and 1:1:1 mass ratio of pITR, pRepCap, and
adenoviral pHelper (Agilent) plasmids. Cells and media were harvested five days post-
transfection and freeze-thawed. Virus was precipitated from the cleared lysate by adding
NaCl to a final concentration of 500 mM and polyethylene glycol (PEG) to a final
concentration of 8%. PEG pellets were routinely resuspended in 50 mM HEPES, 150
mM NaCl, 1% Sarkosyl, pH 8.0. For certain vectors, the PEG pellet was resuspended in
500 mM NaCl, 40 mM Tris, 10 mM $MgCl_2$, pH 9.5. Resuspended vector was purified by
two rounds of cesium chloride ultracentrifugation in a 50.2Ti rotor (Beckman) spun at

45,000rpm x >21 hr and 65,000 rpm x >5 hr. Fractions were collected and combined on the basis of refractive index and dot blot titers. Pooled fractions were then dialyzed against PBS supplemented with 0.001% Pluronic-F68 using a Slide-a-lyzer 10K MWCO (Thermo Fisher). We performed three buffer changes, each with 400x excess of buffer relative to the dialyzed volume. Finally, the buffer was exchanged with a 5% sorbitol in PBS supplemented with 0.001% Pluronic-F68. Vector recovered from the dialysis cassette was filter sterilized using a 0.22 μm filter, aliquoted, and stored at -80C. We used 25 flasks for vectors to be used in mouse studies and up to 400 flasks for vectors to be used in non-human primate studies. Vectors were quantified by measuring Benzonase-resistant AAV genomes using a radioactive dot blot assay (Powers et al. 2018). Vectors used in a single experiment were titered side-by-side with a probe targeting a common region of the genome, either the hSYN1 promoter-enhancer region or GFP. Vector production was scaled down to a 6-well format for evaluating relative production efficiency.

Animals

Animals were maintained by the OHSU Department of Comparative Medicine husbandry and veterinary staff. Eight week-old C57BL/6J mice were purchased from the Jackson Laboratory. AAV vectors were administered via lateral tail vein in a total volume of 300uL by diluting vectors with a 5% sorbitol solution made in PBS.

Mouse tissues for molecular studies involving DNA/RNA recovery were harvested as quickly as possible and frozen on dry ice. For histology, animals were transcardially perfused with heparinized 0.01 M PBS followed by 4% paraformaldehyde

(PFA) in 0.1 M phosphate buffer (PB). The brain was post-fixed in 4% PFA at 4ºC overnight, subsequently equilibrated with 30% sucrose in PBS, embedded in OCT compound (Sakura), flash frozen in dry ice-equilibrated isopentane and stored at -80ºC.

Recovery of anti-sense transcripts

For the proof-of-concept study, total RNA was isolated from brain tissue using TRIzol reagent (Invitrogen) per the manufacturer protocol and DNase-treated with the DNA-free DNA Removal Kit (Ambion). 2µg of DNase-treated RNA was reverse transcribed using oligo-dT and the RETROscript First Strand Synthesis Kit (Ambion). MMLV-RT was substituted with water for the RT-negative conditions. The entire cap open reading frame was then recovered as a 2403 bp amplicon by PCR using Platinum Pfx DNA Polymerase (Invitrogen) and the AAV9capF and AAV9capR primers. For PCR we used an initial denaturing step at 95ºC x 2min followed by 35 cycles of 95ºC x 15 seconds, 55ºC x 30 seconds, 68ºC x 3 minutes, and a final extension of 68ºC x 5 minutes. Subsequently, the region of the MVM intron was PCR amplified using the MVMF and MVMR primers and similar PCR conditions, adjusting the annealing temperature to 60ºC and the extension time to 75 seconds. Expected amplicon sizes were evaluated by gel electrophoresis. PCR products were then purified using QIAQuick Gel Extraction Kit (QIAGEN). The truncated antisense amplicon was cloned into a TOPO vector using the Zero Blunt TOPO PCR cloning kit (Thermo Fisher). Sanger sequencing was outsourced to Elim Biopharmaceuticals Inc. (Hayward, CA). The cap ORF was sequenced using the AAV9PepF1 primer, the MVM intron was sequenced using the MVMseq primer, and TOPO clones were sequenced with the universal M13F primer.

Table 2.1 Primer sequences

Primer ID	Sequence (5' → 3')
AAV9pepF1	AAAGAAGGAGAGGACCGTTTC
AAV9pepR1	GAAAGTTGCCGTCCGTGTGA
AAV9pepF2	GATGAATCCTGGACCTGCTATG
AAV9pepR2	GGTGCTTCATTCCAAACCCT
AAV9pepF3	AGAGTGCTGGAGGTGGTTCT
AAV9pepR3	TGTGCAGATCCGCCGCCTCC
AAV9RT1	GGGCTCTCAATGGACGTAATAG
AAV9capR	GCCGGAAGCTCATCTAGAAA
AAV9capF	ATCTGGTCAATGTGGATTTGGATGAC
MVMF	GCAGGACCATGTGATCGCGC
MVMR	GAGGAGTCGTGTCGTGCCTGAGAG
MVMseq	TACCTGCTGAAGCATTGCAC
VBC1F	ACCTACGTACTTCCGCTCAT
VBC1R	TCCCGACATCGTATTTCCGT
VBC2F	ACGGAAATACGATGTCGGGA
VBC2R	CTTCTCGTTGGGGTCTTTGC

<u>Immunofluorescence</u>

Mouse brains were cryosectioned in the sagittal plane at a thickness of 10 μm and directly mounted on glass slides. Sections were washed in PBS then blocked in 5% donkey serum (Sigma) + 0.5% Triton-X100 in PBS for one hour at room temperature. Primary antibodies (Table 2.2) were diluted in blocking buffer and applied overnight at 4ºC. After washing, sections were incubated with secondary antibodies (Table 2.2) diluted in blocking buffer at room temperature for two hours. Samples were washed with PBS, counterstained with Hoechst 33342 (MolecularProbes, 1:10,000), and coverslipped using Fluoromount-G (SouthernBiotech). Images were acquired using a Keyence BZ-X710 equipped with CFI Plan Apo lambda 20X objective and structured illumination module to achieve optical sectioning.

Table 2.2 List of antibodies utilized for immunofluorescent studies

Host	Target (Conjugate)	Vendor	Product	Working Dilution
chicken	anti-GFP	Abcam	ab13970	1:1000
rabbit	anti-NeuN	Millipore	ABN78	1:1000
goat	anti-chicken (AF647)	LifeTechnologies	A21449	1:500
donkey	anti-rabbit (Cy3)	JacksonImmunoResearch	111-165-144	1:500

Figures

Plots were generated using the ggplot2 package in R(3.6.0). Sequence logos were generated on WebLogo (https://weblogo.berkeley.edu/logo.cgi). Statistics were carried out using the rstatix wrapper of base stats package in R. Figures were formatted using Adobe Illustrator.

2.4 Results

Design of an AAV vector configuration that facilitates cell type-specific, transcription-dependent recovery of the AAV cap sequence

We developed an AAV vector genome configuration (i.e., the TRADE configuration) that enables expression of the AAV *cap* sequence under the control of a cell type-specific promoter-enhancer sequence of choice. In this configuration, successful recovery of the capsid sequence from transduced cells by reverse-transcription presupposes functional transduction at the level of vector-mediated transgene expression within a cell population defined by the cell type-specific promoter. The TRADE configuration consists of a bicistronic AAV2 ITR vector genome with overlapping antisense expression cassettes. The first expression cassette consists of the AAV2 p40 promoter, an AAV *cap* gene, and the AAV2 polyadenylation sign (Figure 2.1a, purple).

56

We reasoned that retention of the endogenous AAV *cap* expression cassette would facilitate in vitro vector production by more or less preserving viral capsid protein expression stoichiometry. However, the majority of the AAV2 *rep* gene that lies upstream of the p40 promoter was deleted and provided in trans during vector production (see Methods). The second expression cassette is placed antisense to the first cassette and consists of a cell type-specific promoter of choice, the minute virus of mice (MVM) intron, an optional reporter gene, the entire antisense AAV *cap* sequence, and the SV40 polyadenylation signal, which has been embedded in the AAV intron (Figure 2.1a, green). For rare cell populations, the reporter gene may be utilized for fluorescence cell sorting in order to enrich for the target population. Alternatively, the reporter gene can be removed in order to avoid expression of a foreign protein in the host, or to make space for a larger promoter-enhancer sequence.

In theory, the TRADE configuration minimizes co-expression of the p40 and cell type-specific promoter-driven transcripts (Figure 2.1b). Upon successful transduction of the TRADE vector into target cells, the cell type-specific promoter is (ideally) highly active, facilitating recovery of the *cap* sequence. However, the p40 promoter is largely suppressed in the absence of helper functions (e.g. adenoviral superinfection) (Mouw and Pintel 2000), which limits both the potential for recovering p40-driven *cap* sequences and the potential to elicit a cell-mediated immune response against expressed viral capsid proteins. Conversely, during virus production in HEK293 cells supplemented with adenoviral helper plasmids, the p40 promoter is highly active while the cell type-specific promoter will drive a variable level of expression.
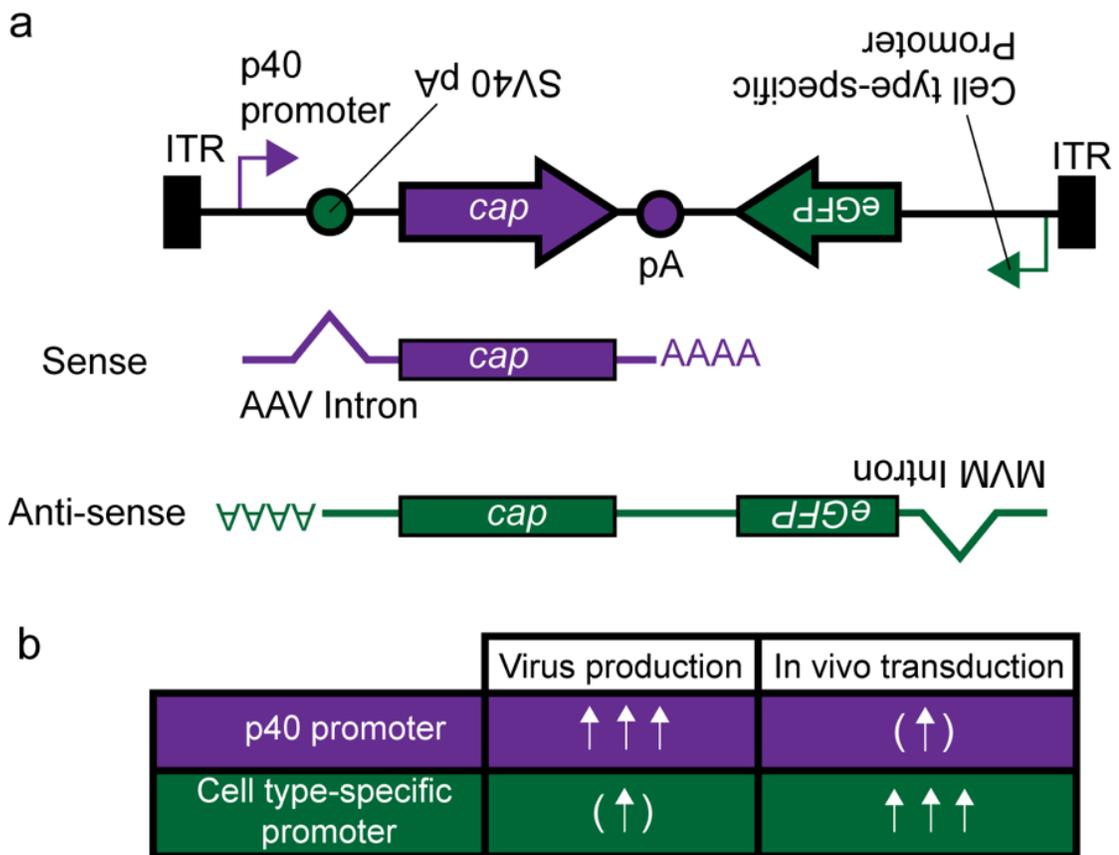
Figure 2.1 The TRADE vector configuration

The <u>TRA</u>nscription-dependent <u>D</u>irected <u>E</u>volution (TRADE) vector configuration enables identification of AAV capsids that mediate cell type-specific transcription while limiting the expression of capsid protein in vivo. Depending on the chosen promoter, expression of anti-sense *cap* transcripts may be limited in vitro during vector production. (a) The TRADE configuration consists of two expression cassettes in a head-to-head orientation. (b) During production, the p40 promoter drives expression of the *cap* open reading frames (purple). Following successful transduction of human synapsin I (hSNY1)-expressing neurons, a transcript containing an optional GFP reporter and the antisense cap sequence is expressed and the minute virus of mice (MVM) intron is spliced out (green).

In vitro evaluation of the TRADE construct

We thus set out to test the bicistronic TRADE design in vitro. For this proof-of-concept study, we chose to target neurons and generated a TRADE vector using a human synapsin 1 promoter-enhancer (hSYN1) sequence that expresses specifically in neurons (Schoch, Cibelli, and Thiel 1996; Hioki et al. 2007) and the AAV9-N272A-PHP.B capsid, which is known to efficiently transduce neurons in the mouse brain following intravenous administration (Adachi, unpublished)(Deverman et al. 2016; S. J. Huang et al. 2017). Cloning of these elements into the TRADE configuration generated the pAAV9-N272A-PHP.B-hSYN1-GFP-TRADE plasmid.

We then tested expression of the antisense transcript in vitro using PEI-mediated plasmid transfection. We observed robust GFP expression in the murine neuroblastoma Neuro-2a (N2a) cell line, supporting the expected strong expression of the hSYN1 promoter-enhancer in neuronal cells (Figure 2.2a). However, we also observed low levels of GFP expression in HEK293 cells, which suggested a low level of non-specific expression in the cell line used for virus production.

Concerned that co-expression of p40-driven sense and hSYN1-driven antisense transcripts during production may adversely affect vector yields, we sought to benchmark production of TRADE vectors to standard recombinant AAV vectors (Figure 2.2b). We therefore produced AAV9-hSYN1-TRADE vectors by triple transfection of an ITR-cap-TRADE plasmid (pAAV9-hSYN1-TRADE) with trans supplementation of an AAV2 Rep plasmid (pHLP-Rep) and an adenoviral helper plasmid, and AAV9-hSYN1-GFP vectors using the standard triple transfection of an ITR-containing hSYN1-GFP AAV genome with an AAV2 Rep / AAV9 Cap helper plasmid and adenoviral helper plasmid. As a

negative control, we performed a mock production using the ITR-containing hSYN1-GFP AAV genome and pHLP-Rep (No Cap). We found that TRADE vectors were produced with approximately 2/3 the efficiency of the standard recombinant AAV production system, but the decrease did not reach statistical significance (Figure 2.2b). Thus, low level expression of antisense transcripts did not significantly impact our ability to produce high viral titers. It is also worth noting that our laboratory has successfully generated CAG-TRADE vectors despite high levels of antisense transcript expression during production (Takahama, unpublished; Furosho, unpublished).
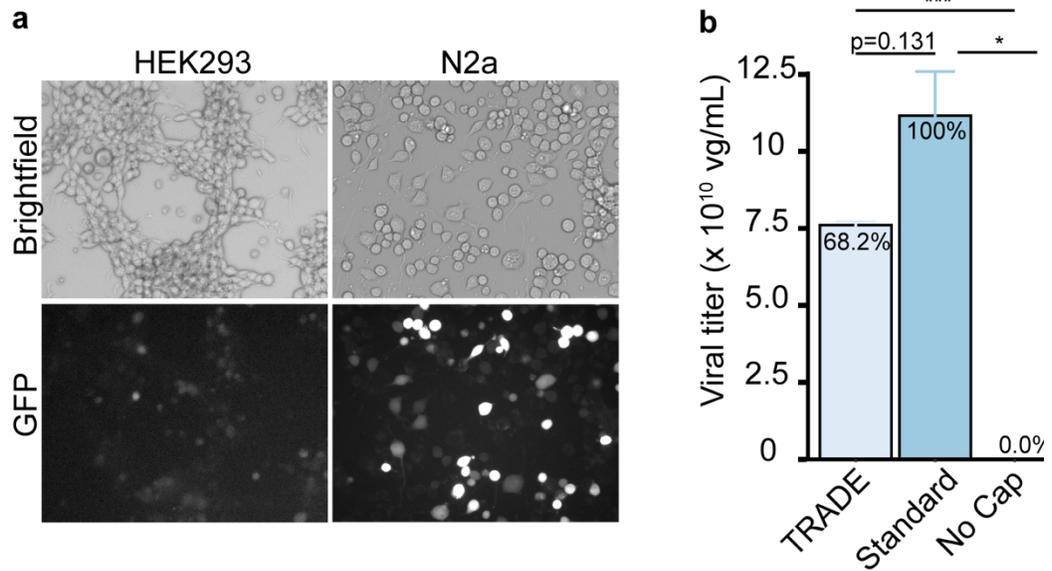
Figure 2.2 Validation of the bicistronic TRADE construct design in vitro

(a) Expression of the cell type-specific antisense transcript was confirmed by visualizing native GFP expression in transfected cell lines. Cell lines were PEI-transfected with 2 µg of the pAAV9-N272A-PHP.B-hSYN1-GFP-TRADE plasmid. Two days later, weak GFP expression was observed in the HEK293 packaging cell line, while stronger expression was observed in the mouse neuroblastoma Neuro-2a (N2a) cell line. (b) AAV9 vectors were produced using either a TRADE vector supplemented with a *rep* helper (AAV9-hSYN1-TRADE), or a hSYN1-GFP vector plasmid supplemented with a *rep-cap* helper (AAV9-hSNY1-GFP). As a negative control, we performed a mock production using the hSYN1-GFP vector and provided *rep* helper, but no *cap* helper. Benzonase-resistant titers were quantified by radioactive dot blot assay that recognized a shared hSYN1 sequence and are plotted as mean ± SEM (n = 3 independent transfection / condition). Uncorrected Welch's t-tests, *p < 0.05, ***p < 0.001.

In vivo, transcription-dependent capsid sequence recovery

We then moved to an in vivo model and tested the ability of the TRADE platform

to facilitate recovery of the capsid sequence from cell type-specific-driven antisense

transcripts (Figure 2.3). AAV9-N272A-PHP.B-hSYN1-TRADE vectors with (GFP+) and

without (GFP-) a GFP reporter in the cell type-specific transgene cassette were produced

and intravenously administered into 8-week-old C57BL/6J mice via tail vein at a dose of

3×10^{11} vector genomes (vg) per mouse. Twelve days post-injection, brains were either

perfused with 4% PFA and harvested for immunohistology, or flash frozen for

subsequent RNA extraction, RT-PCR, and sequencing. In the animal that was

administered the GFP+ vector, we visualized strong GFP expression across the mouse

brain, consistent with our use of the AAV9-N272A-PHP.B capsid. Counterstaining with

the neuronal marker NeuN confirmed that the hSYN1 promoter/enhancer restricted GFP

expression to neurons (Figure 2.3a) despite our observation of non-specific expression in

vitro (Figure 2.2a). For both vectors, we were able to successfully recover the full-length

cap gene by RT-PCR of DNase-treated total RNA extracted from brain (Figure 2.3b).

PCR amplicons were predominantly of the expected size (~2.3 kb). Importantly, we

included a negative control that omitted reverse transcriptase in order to demonstrate that

there was no significant contribution of DNA contamination to recovery of this amplicon.

Subsequent sequencing of the recovered amplicon resulted in the expected AAV-PHP.B

cap sequence consisting of the 21 nucleotide peptide insertion in the AAV9-N272A

backbone (Figure 2.3c). A separate PCR reaction, targeting the MVM exon-exon

junction, resulted in dominant amplicons of the expected size (305 bp for the –GFP

construct and 893 bp for the +GFP construct) assuming splicing of the 82 bp MVM

intron (Figure 2.3d). Sequencing confirmed the expected splicing event (Figure 2.3e) and

provided further evidence that the PCR template was derived from mature antisense

hSYN1-driven transcripts, and not contaminating vector genomic DNA.
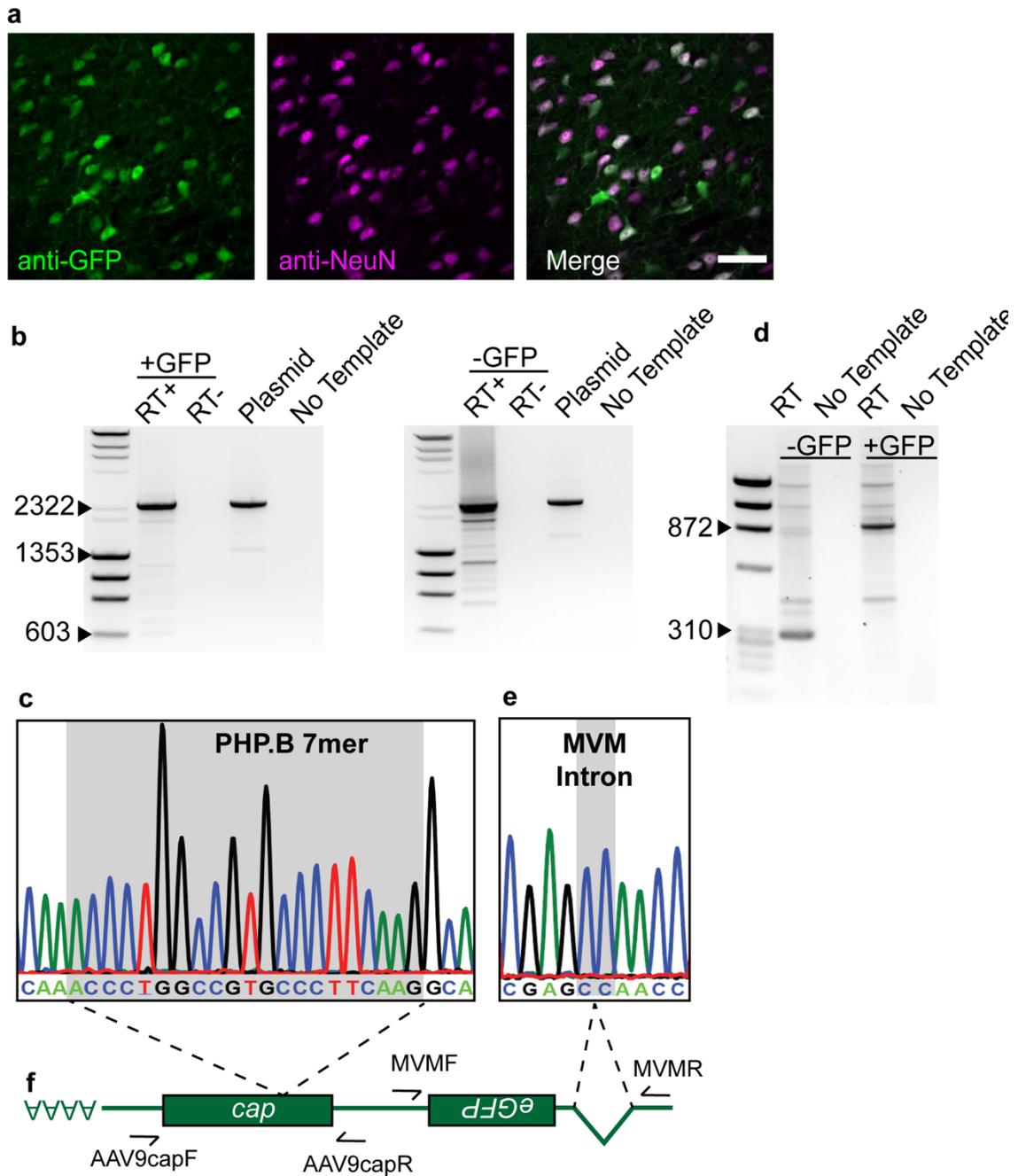
Figure 2.3 Transcription-dependent recovery of AAV *cap* from a cell type-specific transcript

AAV9-N272A-PHP.B-hSYN1-GFP-TRADE (+GFP) and AAV9-N272A-PHP.B-hSYN1-GFP-TRADE (-GFP) vectors were administered intravenously into C57BL/6J mice. (a) GFP expression in thalamus following administration of the +GFP vector co-localized with the neuronal NeuN marker. Scale bar = 50 µm. (b) RT-PCR generated an amplicon consistent with recovery of the full-length *cap* sequence for animals receiving either vector. (c) Electropherogram demonstrating that

Sanger sequencing of the amplicon correctly identifies the PHP.B peptide (highlighted in gray). (d) RT-PCR spanning the exon-exon junction recovered amplicons of the expected size and (e) sequencing confirmed splicing of the MVM intron. (f) Map of the hSYN1 transcript showing approximate PCR primer binding sites.

Identification of an unexpected splicing event on the antisense transcript

In the course of the successful proof-of-concept study described above, we also discovered a potential limitation of the TRADE platform. We attempted transcription-dependent recovery of the anti-sense cap sequence in vitro using the samples seen in Figure 2.2a. Given the clear GFP expression, we expected to easily recover the expected amplicon of ~2.4 kb, but were surprised to generate only ~0.7 kb amplicons across all conditions tested (Figure 2.4). Given that PCR controls using a plasmid template or no template gave the expected results, the observed amplicons were unlikely to be artifactual. Therefore, we decided to investigate this phenomenon and sequenced the truncated species. We discovered that the short length was due to a single ~1.7kb deletion (Figure 2.5a) that conspicuously began and ended with sequences matching highly conserved splice donor (GT) and splice acceptor (AG) dinucleotides (Zhang 1998). Additionally, a poly-pyrimidine tract immediately upstream of the splice acceptor was readily apparent such that the 9nt immediately upstream of the AG consensus sequence were pyrimidines (Figure 2.5b). Multiple sequence alignment of 122 naturally-occurring AAV variants suggest that the consensus splicing motifs may be evolutionarily conserved (Figure 2.6, AAV serotypes 1-13 are shown).

Because this unexpected splicing could interfere with our ability to recover capsid sequencing information using the TRADE platform, we developed modified constructs that ablated splicing by disrupting the splice donor and acceptor motifs. We accomplished this by mutagenizing each splicing element, such that the mutations were synonymous in the *cap* sense orientation Figure 2.7a. We developed three constructs, no splice acceptor (NS1), no splice donor (NS2), and neither splice acceptor nor donor

(NS3). All three constructs successfully suppressed splicing following transfection in N2a cells (Figure 2.7b).

Notably, mutations in the antisense splice acceptor region disrupt the AAP open reading frame and are expected to interfere with vector production. We therefore assayed the ability of the NS constructs to mediate vector production in the presence and absence of supplemented AAP9 (Figure 2.7c). As expected, constructs with mutations in the antisense splice acceptor (NS1 and NS3) failed to produce viral particles in the absence of AAP9. However, all constructs mediated effective vector production in the presence of AAP9.
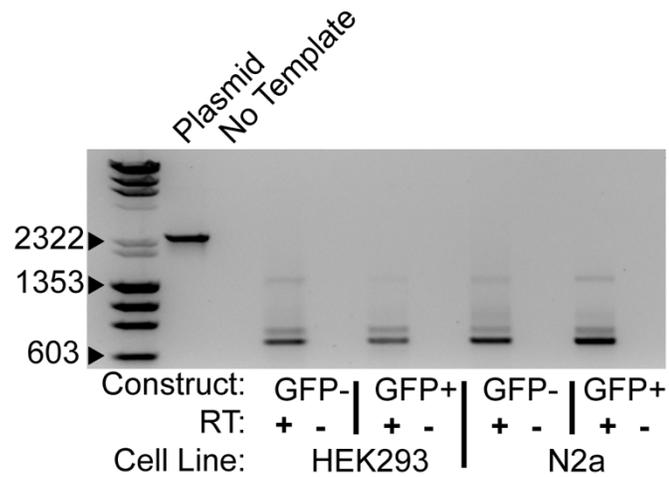
Figure 2.4 In vitro expression of the cell type-specific TRADE transcript following PEI transfection

RNA was extracted from HEK293 and N2a cell transfected with the pAAV9-N272A-PHP.B-hSYN1-GFP-TRADE (GFP+) or pAAV9-N272A-PHP.B-hSYN1-TRADE (GFP-) constructs three days post-transfection. Samples are the same as seen in Figure 2.2a. RT-PCR attempted to recover the ~2.4 kb full length AAV9-N272A-PHP.B *cap* sequence. Unexpectedly, we consistently recovered a truncated species of ~0.7 kb. Construct indicates which plasmid was transfected. RT indicates whether or not reverse transcriptase was included in the reaction.

a

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTTAGTGAAGGAATTCGCGAGTGGTGGGCTTTGAA
ACCTGGAGCCCCTCAACCCAAGGCAAATCAACAACATCAAGACAACGCTCGAGGTCTTGTGCTTCCGGGTTACAAAT
ACCTTGGACCCGGCAACGGACTCGACAAGGGGGAGCCGGTCAACGCAGCAGACGCGGCGGCCCTCGAGCACGACAAG
GCCTACGACCAGCAGCTCAAGGCCGGAGACAACCCGTACCTCAAGTACAACCACGCCGACGCCGAGTTCCAGGAGCG
GCTCAAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAGGCCAAAAAGAGGCTTCTTGAACCTC
TTGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCTGGAAAGAAGAGGCCTGTAGAGCAGTCTCCTCAGGAACCGGAC
TCCTCCGCGGGTATTGGCAAATCGGGTGCACAGCCCGctaaaaagagactcaatttcggtcagactggcgacacaga
gtcagtcccagaccctcaaccaatcggagaacctcccgcagcccctcaggtgtgggatctcttacaatggcttcag
gtggtggcgcaccagtggcagacaataacgaaggtgccgatggagtgggtagttcctcgggaaattggcattgcgat
tcccaatggctgggggacagagtcatcaccaccagcacccgaacctgggccctgcccacctacaacaatcacctcta
caagcaaatctccaacagcacatctggaggatcttcaaatgacaacgcctacttcggctacagcacccctgggggt
attttgacttcaacagattccactgccacttctcaccacgtgactggcagcgactcatcaacaacaactggggattc
cggcctaagcgactcaacttcaagctcttcaacattcaggtcaaagaggttacggacaacaatggagtcaagaccat
cgccaataaccttaccagcacggtccaggtcttcacggactcagactatcagctcccgtacgtgctcgggtcggctc
acgagggctgcctcccgccgttcccagcggacgttttcatgattcctcagtacgggtatctgacgcttaatgatgga
agccaggccgtgggtcgttcgtccttttactgcctggaatatttcccgtcgcaaatgctaagaacgggtaacaactt
ccagttcagctacgagtttgagaacgtacctttccatagcagctacgctcacagccaaagcctggaccgactaatga
atccactcatcgaccaatacttgtactatctctcaaagactattaacggttctggacagaatcaacaaacgctaaaa
ttcagtgtggccggacccagcaacatggctgtccaggggaagaaactacatacctggacccagctaccgacaacaacg
tgtctcaaccactgtgactcaaaacaacaacagcgaatttgcttggcctggagcttcttcttgggctctcaatggac
gtaatagcttgatgaatcctggacctgctatggccagccacaaagaaggagaggaccgtttctttcctttgtctgga
tctttaatttttggcaaacaaggaactggaagagacaacgtggatgcggacaaagtcatgataaccaacgaagaaga
aattaaaactactaacccggtagcaacggagtcctatggacaagtggccacaaaccaccagagtgcccaagcacagg
cgcagaccggctgggttcaaaaccaaggaatacttccgggtatggtttggcaggacagagatgtgtacctgcaagga
cccatttgggccaaaattcctcacacggacggcaactttcacccttctccgctgatgggagggtttggaatgaagca
cccgcctcctcagatcctcatcaaaaacacacctgtacctgcggatcctccaacggccttcaacaaggacaagctga
actctttcatcacccagtattctactggccaagtcagcgtggagatcgagtgggagctgcagaaggaaaacagcaag
cgctggaacccggagatccagtacacttccaactattacaagtctaataatgttgaatttgctgttaatactgaagg
tgtatatagtgaaccccgcccccattggcaccagatacCTGACTCGTAATCTGTAA

b
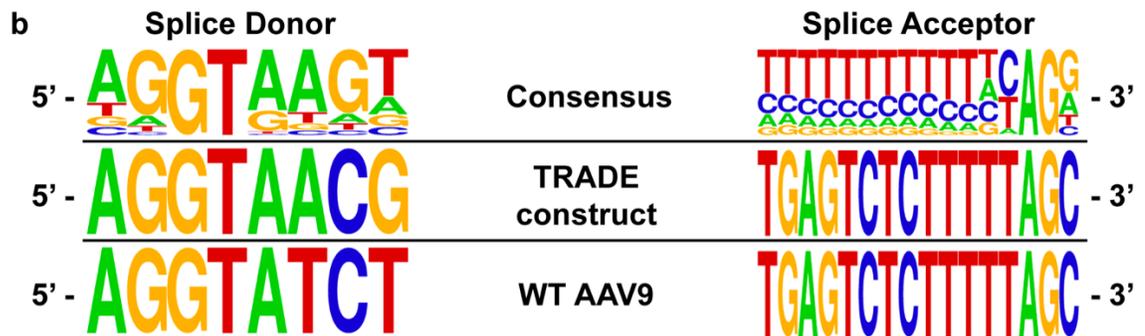
Figure 2.5 Truncation of the antisense transcript is consistent with statistical features of a splicing event

(a) The AAV9-N272A-PHP.B VP1 sequence. The portion of the sequence in lower case and red was not recovered by RT-PCR. (b) The missing sequence is consistent with a splicing event in the antisense orientation and contains the consensus splice donor and acceptor sequences, as well as a pyrimidine-rich region upstream of the splice acceptor. Aligned sequence logos are shown for the human splicing consensus sequences, the observed putative splicing event in the TRADE construct, and wild-type AAV9.
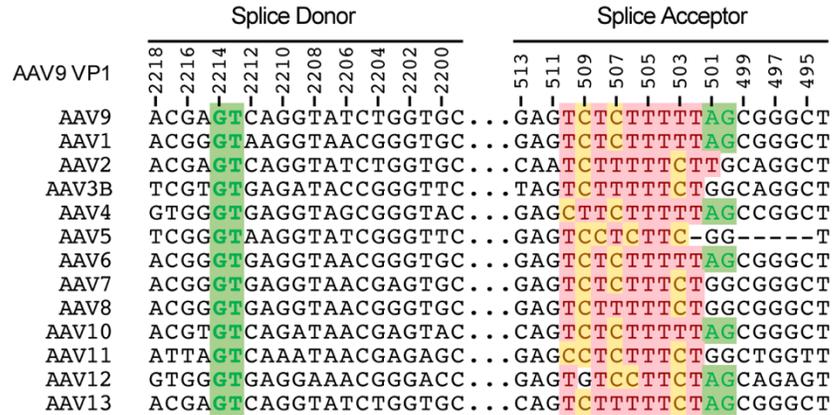
Figure 2.6 Multiple sequence alignment of AAV serotypes surrounding the identified antisense splice donor and acceptor
AAV serotype Cap sequences were aligned and are presented in reverse complement. Position numbers are provided with respect to the AAV9 VP1 sequence. Green text identifies a perfect match and apparent conservation of the consensus dinucleotide splice donor (GT) or splice acceptor (AG) identified in Figure 2.5. Red and yellow text highlight the poly-pyrimidine tract immediately upstream of the splice acceptor.

a

```
          500
           |
Native AGCCCGCTAAAAAGAGACTCAATTTCGGTCAGACTGGCGACACAGAGTCAGTT
  NS1  ....T..C..G...C.C..G..C..T..C.....C............AGC...
```

```
                        2214
                          |
Native TGAACCCCGCCCCATTGGCACGCGTTACCTGACTCGTAATCTGTAATTGCTTG
  NS2  .............A..C..G..AA.A..T.....AA.A..C.............
```

b

Construct: NS1 | NS2 | NS3
RT: + - | + - | + -
Cell Line: N2a

c

Standard

1 7
2 8
3 NoCap
4 NS1
5 NS2
6 NS3
   WT AAV9
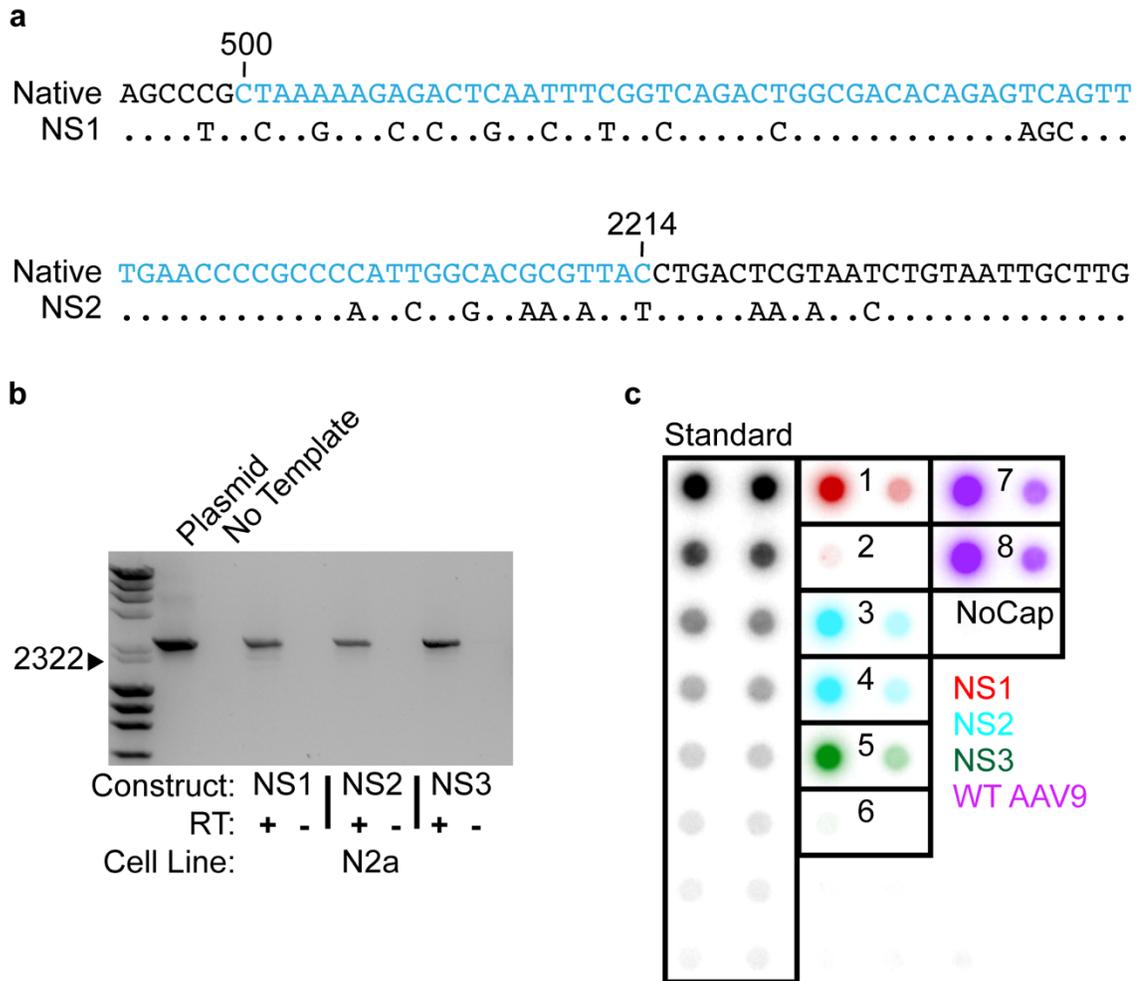
Figure 2.7 Antisense splicing can be suppressed without disrupting vector production

We generated three non-splicing (NS) constructs to disrupt antisense splicing. Constructs were based on pAAV9-N272A-PHP.B-hSYN1-GFP-TRADE with mutations surrounding the conserved splice acceptor (NS1), splice donor (NS2), or both the splice acceptor and donor (NS3). (a) Mutagenesis of the splice acceptor and splice donor. Bottom sequence indicates changes from the native sequence, with identity indicated as dots. Blue color highlights the intron. Position numbers are with respect to AAV-PHP.B VP1 *cap*. (b) Recovery of the entire AAV9-N272A-PHP.B *cap* amplicon by RT-PCR of total RNA isolated from transfection of Neuro-2a (N2a) cells. (c) Production of AAV TRADE vectors using NS constructs in the presence or absence of supplemented AAP9. A recombinant AAV9-hSYN1-GFP vector was produced using a standard triple transfection method as a positive control. The negative control (NoCap) was the same as the positive control, but substituted the AAV9 RepCap plasmid with the pRep plasmid that does not express Cap. Boxes

71

with odd numbers (1, 3, 5, 7) represent vector production in the presence of a supplementary plasmid providing AAP9. Boxes with even numbers (2, 4, 6, and 8) represent production in the absence of supplementary AAP9. For each condition, the left dot represents a 10x higher concentration than the right dot.

2.5   Discussion

We describe the development of a transcription-dependent directed evolution platform, termed TRADE. This platform facilitates *in vivo* biopanning of AAV capsid libraries by selectively recovering capsid sequences from vector genome-derived mRNA. As such, we apply a selective pressure for functional AAV variants that are able to complete all of the complex steps involved in AAV-mediated transduction, up to the point of transgene expression. Critically, the TRADE system is highly flexible and biopanning requires only standard molecular biology tools. It is amenable to virtually any promoter-enhancer sequence of choice (up to 1.8kb if the GFP transgene is excluded). Because transgenic animals expressing Cre-recombinase are not required to achieve stringent selection of functionally viable variants (Deverman et al. 2016; Chan et al. 2017; Ojala et al. 2018), the TRADE system is readily applied to more clinically-relevant models such as non-human primates, human-derived xenograft models (Lisowski et al. 2014; Keswani et al. 2012; GuhaSarkar et al. 2017) or human organoids (Garita-Hernandez et al. 2020; Depla et al. 2020; Kim, Koo, and Knoblich 2020).

The key limitation to employing TRADE is the requirement to identify an appropriate promoter-enhancer sequence. For relatively homogeneous tissues with a large population of the target cell, a strong ubiquitous promoter is likely to suffice. However, for more rare populations of difficult to transduce cell types or heterogeneous tissues, a cell type-specific promoter is likely required for efficient biopanning. Notably, there have been significant advances with respect to developing small, yet strongly and specifically expressing promoter-enhancer sequences (Portales-Casamar et al. 2010; Jüttner et al. 2019).

During the course of this work, a similarly motivated transcription-dependent AAV directed evolution platform termed Tropism Redirection of AAV by Cell-type-specific Expression of RNA (TRACER) was recently reported by Voyager Therapeutics (Nonnenmacher et al. 2020). In contrast to the TRADE configuration that divorces the p40-driven capsid expression that is required for vector production from cell type-specific promoter-driven expression of the capsid sequence that is required for capsid sequence recovery, the TRACER platform utilizes overlapping transcripts by placing the cell type-specific promoter-enhancer upstream of the AAV2 p40 promoter. While the platforms are otherwise conceptually similar, we note that the seemingly simple change in vector configuration results in several important differences.

(1) The TRADE system minimizes expression of capsid proteins in transduced host cells because the AAV2 p40 promoter is repressed in the absence of superinfection by a helper virus (Mouw and Pintel 2000; Stutika et al. 2016). The cell type-specific promoter drives expression of a non-coding antisense Cap sequence. In contrast, the TRACER system will drive high levels of potentially immunogenic AAV capsid proteins in host cells since the cell type-specific promoter transcript is in the sense orientation. The host immune response against the AAV capsid and foreign transgenes expressed following AAV transduction are well-reported (Samaranch et al. 2014; Manno et al. 2006; Mingozzi et al. 2007). Thus, it is highly likely that expression of foreign capsid proteins in the host will also elicit a robust immune response, and such an immune response may negatively impact biopanning efforts.

(2) It is well-established that the naturally-occurring p40 promoter achieves a stoichiometry of VP1:VP2:VP3 transcripts that is critical for efficient production

(Becerra et al. 1988; Farris and Pintel 2008). In the TRADE system, this stoichiometry is likely undisrupted during production because we retain the p40 promoter sequence. In contrast, it is possible that surplus expression of transcripts driven by the cell type-specific promoter (e.g., SYN1) during production with the TRACER system interferes with this delicate stoichiometry, leading to the reported reduction in viral titer to roughly 1/3 of a standard triple transfection method. While it is difficult to directly compare the results of our experiments given differences in plasmid constructs and titering methods, we observed viral titers that were roughly 2/3 of a standard triple transfection method when using the TRADE system, though the difference was not statistically significant (Figure 2.2). Therefore, TRADE may afford greater production efficiency than TRACER by approximately 2-fold.

(3) The TRADE system allows for relative independence of promoter sequences by placing them on opposite ends of the vector genome. Although AAV construct designs ultimately require empiric validation, we expect that exchanging promoter sequences will be a straightforward process (and this has been true to date with ongoing experiments). In contrast, we posit that the proximity of cell type-specific and p40 promoters in the TRACER platform may require additional optimization steps for each iteration.

(4) We developed TRADE constructs with and without a GFP reporter. The reporter can be used to enrich for a target population using fluorescence-activated cell sorting. In addition, expression of a vector-mediated transgene reporter can theoretically facilitate biopanning at the level of protein expression, similar to the previously reported iTransduce system (Hanlon et al. 2019). Inclusion of the GFP reporter in the antisense transcript is unlikely to affect the sense transcript, which is responsible for vector

production. It is not obvious how the reporter gene would be optimally introduced into the TRACER design since it must be co-expressed with AAV cap and doing so may affect vector production.

(5) In addition, the TRADE configuration enables the selective recovery of p40-driven sense transcripts or cell type-specific promoter-driven antisense transcripts if a strand-specific primer is utilized for reverse transcription. In contrast, the overlapping transcripts generated by the TRACER genome cannot be easily separated by RT-PCR.

(6) Finally, a potential obstacle to application of the TRADE system, not applicable to the TRACER system, was the unexpected observance of an antisense splicing event that may require modification of the TRADE configuration if the entire cap sequence is to be recovered. Importantly, however, we did not observe any clear evidence of antisense cap splicing in vivo following AAV vector administration and transduction. Thus, it remains possible the robust splicing in vitro may be artifactual.

To elaborate on this point, it is important to note that these data were generated in an artificial vector genome (i.e., the TRADE configuration) that harbored silent mutations (previously introduced to facilitate cloning) surrounding the putative splice donor site (see Figure 2.5b). In addition, we used a transient transfection method rather than viral infection. Furthermore, although we compared the observed splicing sequences to the statistical features of splicing motifs in humans (Zhang 1998), the well-described AAV intron on the plus strand utilizes a non-consensus splice donor sequence (Farris and Pintel 2008). Therefore, future studies will be required to study if this splicing event is observed in the wild-type AAV genome context. Of note, an RNA-seq experiment was carried out to comprehensively characterize the transcriptome of AAV2 (Stutika et al. 2016). The

authors discovered new splicing sequences and also a new transcript on the AAV minus strand; however, they did not report the detection of a transcript corresponding to what we observed in our experiments under any condition, including adenoviral superinfection. On the other hand, the robust splicing we observed in vitro was consistent across two cell lines and the splice acceptor and donor sequences that we identified appeared to be evolutionarily conserved in naturally-occurring AAV variants. It is important to recognize that the significance of this discovery to basic AAV biology remains questionable. Nevertheless, we were able to overcome the potential issue this posed for the TRADE vector design by targeted mutagenesis to ablate the anti-sense splicing phenomenon.

To reiterate, the TRADE configuration is designed to minimize co-expression of the p40-driven sense and cell type-specific promoter-driven antisense transcripts during both vector production and following in vivo transduction. This affords us maximal flexibility with respect to manipulating the expression of the vector genome in these two different contexts. In conclusion, this work establishes the TRADE platform for transcription-dependent directed evolution of the AAV capsid, adding to recent technological advances that bring the field ever closer to engineering designer gene delivery vehicles for clinical applications. In addition, we describe the discovery of a splicing event within the antisense *cap* open reading frame. The broader implications of this finding, if there are any, will require further investigation.

3 Chapter 3. Identification of an efficient and highly neurotropic AAV capsid variant for gene therapy in the CNS

3.1 Abstract

Adeno-associated virus (AAV) vectors can cross the blood-brain-barrier and mediate long-lasting expression of therapeutic payloads in terminally differentiated cells. Recent clinical success demonstrated a highly efficacious, and theoretically curative, intervention for infants with spinal muscular atrophy following a one-time administration of AAV9 vectors (i.e., Zolgensma). Nevertheless, broader applications of AAV-mediated gene therapy in the central nervous system (CNS), are hindered by shortcomings of the AAV9 vector, low transduction efficiency of brain neurons following intravenous administration and safety concerns at high dose. Here, we apply the transcription-dependent directed evolution system to successfully identify novel AAV variants with enhanced brain neuronal transduction efficiency and reduced biodistribution to the liver following systemic delivery in both mice and non-human primate.

3.2   Introduction

Adeno-associated virus (AAV) vectors are ideal gene delivery vehicles for treating diseases of the central nervous system (CNS). The ability of AAV vectors to traverse the blood-brain-barrier and mediate long-lasting transgene expression in terminally differentiated target cells suggests a future in which CNS diseases are cured following a one-time intravenous administration. Major milestones toward this future include the approval of subretinally-administered AAV2-RPE65 (Luxturna) for Leber congenital amaurosis and systemically-administered AAV9-SMN (Zolgensma) for spinal muscular atrophy (Russell et al. 2017; Mendell et al. 2017). In addition, to retina and spinal cord, AAV is also capable of transducing the brain via three routes of administration: intraparenchymal, intra-cerebrospinal fluid (CSF), and intravenous (systemic).

Clinical trials have demonstrated the efficacy of intraparenchymal AAV administration into the brain. For example, incredible outcomes were achieved following bilateral intraputaminal administration of AAV2-hAADC for aromatic l-amino acid decarboxylase (AADC) deficiency (Kojima et al. 2019), and a similar vector mediated safe and long-lasting expression in patients with Parkinson's disease (Mittermeyer et al. 2011). However, intraparenchymal injection mediates only limited coverage of the brain, which may be insufficient for many disease applications that require widespread gene transfer.

The blood-brain-barrier can also be bypassed via intra-CSF administration routes that include intrathecal delivery via lumbar puncture, intracerebroventricular injection by open surgery, or intra-cisterna magna injection. Technological advances, have improved

the less invasive lumbar puncture route by allowing a catheter to be advanced through the subarachnoid space (Bey et al. 2020). This technique also benefits from physically crossing the blood-brain-barrier and allows a lower overall dose (per kg body weight) to be administered. However, transduction appears to biased to periventricular and surface structures (Hinderer et al. 2014; Gray et al. 2013). Additionally, dorsal root ganglion pathology is common following AAV administration into the CSF (Hordeaux et al. 2020).

Systemically delivered CNS transduction, if it can be achieved safely, remains an optimal method since it is far less invasive and achieves broad distribution in the CNS. The AAV9 capsid serotype utilized by Zolgensma is able to cross the blood-brain-barrier of primates and transduce motor neurons in the spinal cord and remains the gold standard for systemic gene therapy targeting the CNS. However, high doses (~1-2 x 10^{14} vg/kg) are required and transduction of the brain is nominal (Duque et al. 2009; Gray et al. 2011). In addition, the tropism of AAV9 appears to be significantly biased toward glial transduction following intravenous administration (up to 30 astrocytes transduced per neuron), which may prevent treatment of cell-autonomous diseases affecting neurons (Gray et al. 2011). Critically, significant safety concerns have been raised regarding the use of such high doses following recent deaths in a clinical trial (Wilson and Flotte 2020). Evidence to date suggests that such high doses of AAV9 vectors are hepatotoxic and also carry the risk of genotoxicity (Hinderer et al. 2018; Feldman et al. 2020; Chandler et al. 2015). There is therefore a strong impetus to develop enhanced AAV vectors that can broaden and improve current gene therapy applications.

Finally, it is important to recognize that hereditary diseases involve mutations to every cell, even though pathology may be apparently restricted to specific tissues and cell types. Intriguingly, intra-CSF administration of an AAV9-SMN (survival motor neuron) vector with neuron-specific expression was less efficacious in rescuing a mouse model of spinal muscular atrophy than was intravenous administration of a similar vector with a ubiquitous promoter, despite achieving similar levels of SMN expression in motor neurons (Besse et al. 2020). This suggests that SMN may be important outside of the primary target motor neuron cell population.

To this end, engineering of AAV capsid, and more specifically directed evolution of the AAV capsid, has yielded novel variants that more efficiently cross the blood-brain-barrier and transduce the CNS (Gray et al. 2009; Zinn et al. 2015; Deverman et al. 2016). One such study utilized an improved capsid evolution platform (CREATE, Cre-Recombination-based AAV Targeted Evolution) to isolate the AAV-PHP.B capsid from C57BL/6 brain tissue. AAV-PHP.B mediates 40-50x greater brain transduction than the parental AAV9 serotype following intravenous administration in C57BL/6 mice. Various groups enthusiastically pursued the potential for translation and evaluated the AAV-PHP.B capsid in non-human primates, leading to the resounding conclusion that there was no observable enhancement (Matsuzaki et al. 2018; Hordeaux et al. 2018; Liguore et al. 2019). Intriguingly, enhancement was also lacking in BALB/c mice (Hordeaux et al. 2018). Subsequent analyses determined that the directed evolution platform had selected a mechanism that was dependent on the LY6A protein expressed on the brain microvasculature endothelium of C57BL/6 mice, but not BALB/c mice nor primate (Hordeaux et al. 2019; Q. Huang et al. 2019; Batista et al. 2020). Application of the

CREATE system to the non-human primate context was precluded by the strict requirement for Cre-driver animals.

Given the translational potential of developing a vector that could efficiently cross the blood-brain-barrier and transduce the CNS, and the clear demonstration that directed evolution of the AAV capsid could accomplish this in rodents, we chose to address this problem by leveraging the advantages of our transcription-dependent directed evolution (TRADE) system combined with a rational design approach to reduce vector biodistribution to the liver. Here, we demonstrate successful application of TRADE in mice to identify novel AAV variants with enhanced transduction efficiency and specificity for brain neurons following intravenous administration. Furthermore, we characterize the lead HN1 candidate in both mouse and non-human primate.

3.3   Methods and Materials

<u>Plasmids</u>

We first constructed the pAAV9-N272A-2BsaI-hSYN1-TRADE construct that contains two expression cassettes oriented antisense to each other and flanked by the AAV2 inverted terminal repeats (ITRs). On one side of the genome, the AAV2 p40 promoter and intron is followed by the AAV9 *cap* gene and AAV2 poly-adenylation sequence. The AAV9 cap sequence was modified by (1) an N272A mutation that confers liver detargeting (S. J. Huang et al. 2017), (2) silent mutations to facilitate cloning as previously described (Adachi et al. 2014)(Appendix 6.1) and (3) a 42 bp substitution at position Q588 of the AAV9 VP1 ORF that encodes for 3 frameshifted stop codons flanked by BsaI restriction sites and followed by a GGGGS linker sequence. On the other

side of the genome and antisense to the p40-driven cap ORF, we designed an overlapping transgene cassette consisting of a 471 bp human synapsin I (hSYN1) promoter-enhancer sequence, a 92 bp minute virus of mice (MVM) intron, the non-coding antisense sequence, and an SV40 poly-adenylation sequence placed in the AAV intron such that it avoided the predicted splicing branch point as determined using the Human Splicing Finder (Desmet et al. 2009). The pAAV9-N272A-2BsaI-hSYN1-GFP-TRADE construct is similar, but additionally contains a PCR-generated, codon-modified enhanced GFP (GFP) sequence in the hSYN1-driven transgene cassette, anti-sense to the *cap* open reading frame.

For the proof-of-concept study, we generated the pAAV9-N272A-PHP.B-hSYN1-GFP-TRADE plasmid by synthesizing the AAV9-PHP.B cap sequence (Genscript) and using simple restriction enzyme cloning to insert the TLAVPFK sequence into the pAAV9-N272A-2BsaI-hSYN1-TRADE backbone.

The initial TRADE plasmid peptide display library (pAAV9-N272A-hSYN1-GFP-TRADE-Lib0) was constructed by PCR amplification of an AAV9 cap sequence harboring a GGGS[NNK]$_8$GGGGS substitution at position Q588 of the AAV9 VP1. The 355 bp amplicon and vector plasmid were digested with AflII and NdeI. Library construction was carried out using a 1:20 molar ratio of vector to insert, 500 ng vector, and T4 DNA Ligase (NEB) at 16ºC overnight. The optimal molar ratio was determined empirically on a small scale prior to the large-scale reaction. The reaction was purified using phenol-chloroform and isopropanol precipitation containing 0.25 M potassium acetate. DNA was resuspended in 10 µL water and transformed into ElectroMAX DH10B cells (Invitrogen). E.coli were allowed to recover in SOC media while

shaking at 37ºC for one hour, and then spread over 20 LB agar plates containing Ampicillin (50ug/mL). We additionally spread four titering plates in order to calculate the total colony number and theoretical diversity of the library. The next day, colonies were collected by resuspending colonies with Luria Broth. E.coli were allowed to grow shaking at 37ºC until reaching an OD600 > 2.3, then plasmid was prepared by maxiprep (QIAGEN).

After the first round of selection in mouse brain, we determined that we could recover sufficient material without the need for enriching for GFP+ cells. We reasoned that inclusion of this transgene would limit the size of future tissue-specific promoters and also incorporated an unnecessary risk of GFP overexpression-induced inflammation (Samaranch et al. 2014; Ciesielska et al. 2013) and potential liver toxicity (Gray et al. 2011). We therefore, utilized the pAAV9-N272A-2BsaI-hSYN1-TRADE plasmid that does not contain a GFP transgene as the vector for generation of subsequent libraries. Furthermore, we opted to use DNA assembly by NEBuilder HiFi DNA Assembly Master Mix (NEB) to generate subsequent libraries because it demonstrated ~5-fold higher efficiency than the T4 DNA ligation in our hands. The NEBuilder reaction was carried out per the manufacturer protocol using an empirically-optimized molar ratio of vector to insert and the product was purified and transformed as described above.

RepCap helper plasmids were generated for 21 novel variants using a cloning service (Genscript) and the pAAV9-N272A-SBBANXB RepCap plasmid as a backbone for insertion of the peptide sequences.

For characterization of neuronal transduction, we used constructs containing a hSYN1 promoter-enhancer. The pdsAAV-hSYN1-GFP plasmid is a self-complementary

AAV construct containing a hSYN1 promoter-enhancer, GFP transgene, and SV40 polyadenylation signal. We also generated a set of hSYN1-GFP-BC plasmids (pdsAAV-hSYN1-GFP-BCx) by assembling PCR-amplified viral barcode cassettes into the 3' untranslated region of pdsAAV-hSYN1-GFP using NEBuilder. Barcode cassettes contain two unique 12 nucleotide barcodes flanked by 20 bp primer binding sites and were previously validated in the original report of AAV Barcode-Seq (Adachi et al. 2014). Individual clones were independently maxiprepped (QIAGEN) then characterized by both sequencing and restriction enzyme digestion, including BglI digestion of the ITRs.

For further characterization, we utilized ubiquitously expressing CAG constructs. The single strand pAAV-CAG-FLAGnlsGFP and pAAV-CAG-HAnlsGFP plasmids containing a CAG promoter (CMV enhancer, chicken beta-actin promoter, and rabbit beta-globin intron), N-terminally epitope-tagged and nuclear localized GFP transgene, WPRE sequence, and human growth hormone polyadenylation signal were derived from CAG-NLS-GFP, a kind gift from Viviana Gradinaru (Addgene #104061), by adding a FLAG (DYKDDDDK) or HA (YPYDVPDYA) epitope to the N-terminus of the nlsGFP open reading frame. Notably, the NLS-GFP open reading frame in the Addgene #104061 construct contains two potential start codons, one introduced with the nuclear localization sequence, and one remaining from the GFP open reading frame. When adding FLAG/HA tag sequences we replaced the start codon at the 5' end of the NLS, but we neglected to remove the start codon of GFP. Thus, cytoplasmic expression of GFP may occur. Each plasmid was then used to generate individual barcode clones (pAAV-CAG-FLAGnlsGFP-BC$_x$ and pAAV-CAG-HAnlsGFP-BC$_x$) as described above for the pdsAAV-hSYN1-GFP-BCx constructs.

<u>AAV Vectors</u>

We produced AAV by adenovirus-free triple transfection of HEK293 cells in
T225 flasks using polyethyleneimine (PEI). Transfection was carried out using a
DNA:PEI mass ratio of 1:2, a total of 45 µg DNA, and 1:1:1 mass ratio of pITR,
pRepCap, and adenoviral pHelper (Stratagene) plasmids. Cells and media were harvested
five days post-transfection and freeze-thawed. Virus was precipitated from the cleared
lysate by adding NaCl to a final concentration of 500 mM and polyethylene glycol (PEG)
to a final concentration of 8%. PEG pellets were routinely resuspended in 50 mM
HEPES, 150mM NaCl, 1% Sarkosyl, pH 8.0. For one experiment, the PEG pellet was
resuspended in 500 mM NaCl, 40 mM Tris, 10 mM $MgCl_2$, pH 9.5. Resuspended vector
was purified by two rounds of cesium chloride ultracentrifugation in a 50.2Ti rotor
(Beckman) spun at 45,000rpm x >21 hr and 65,000rpm x > 5 hr. Fractions were collected
and combined on the basis of refractive index and dot blot titers. Pooled fractions were
then dialyzed against PBS supplemented with 0.001% Pluronic-F68 using a Slide-a-lyzer
10K MWCO (Thermo Fisher). We performed three buffer changes, each with 400x
excess of buffer relative to the dialyzed volume. Finally, the buffer was exchanged with a
5% sorbitol in PBS supplemented with 0.001% Pluronic-F68. Vector recovered from the
dialysis cassette was filter sterilized using a 0.22 µm filter, aliquoted, and stored at -80ºC.
We used 25 flasks for vectors to be used in mouse studies and up to 400 flasks for vectors
to be used in non-human primate studies. Vectors were quantified by measuring
Benzonase-resistant AAV genomes using a radioactive dot blot assay (Powers et al.
2018). Vectors used in a single experiment were titered side-by-side with a probe

targeting a common region of the genome, either the hSYN1 promoter-enhancer region or GFP.

We modified the protocol for production of AAV peptide display libraries. The initial AAV library (AAV9-N272A-hSYN1-TRADE-Lib0) used a 0.01:1:1 plasmid mass ratio of TRADE vector:pHLP-Rep (AAV2 RepCap plasmid with large deletion of the Cap gene):pHelper, respectively. This transfection ratio was empirically determined to generate sufficient vector for our studies while limiting the ratio of TRADE plasmids per cell to 1,000, which substantially decreases the risk of cross-packaging and mosaicism (Nonnenmacher et al. 2015; Schmit et al. 2020). Subsequent libraries were produced using a plasmid mass ratio of 0.1:1:1. Vectors were purified and titrated as described above.

Barcoded AAV libraries were generated by individually producing viral barcode constructs into capsid variants with a pre-established 1-to-1 relationship between the barcode sequence and the capsid variant. DNase-resistant vector genomes in crude lysate were measured by radioactive dot blot. Crude lysate was then pooled such that each viral clone was represented at a roughly equimolar ratio. The pooled lysate was then purified as described above. A precise equimolar mixture is not critical for the AAV Barcode-Seq analysis pipeline.

For the CAG-nlsGFP-BC libraries, we independently produced nine different barcoded AAV vectors for each of the four capsid-epitope pairs (i.e., AAV9-FLAG, AAV9-HA, HN1-FLAG, HN1-HA). We mixed crude lysate and independently purified four barcode libraries without mixing clones with different capsid-epitope relationships. We then performed three independent side-by-side radioactive dot blot assays to

determine the concentration of DNase-resistant genomes and mixed two equimolar libraries. Thus, each CAG-BC library contained 18 different barcode clones, nine representing one capsid-epitope relationship and nine representing the reciprocal capsid-epitope relationship.

<u>Animals</u>

Mice: Animals were maintained by the OHSU Department of Comparative Medicine husbandry and veterinary staff. Eight-week old C57BL/6J and BALB/cJ male mice were purchased from the Jackson Laboratory. Mice with humanized livers were generated as previously described (Azuma et al. 2007). AAV vectors were administered via lateral tail vein in a total volume of 300 µL by diluting vectors with a 5% sorbitol solution made in PBS. Mouse studies were blinded when possible using two researchers. Researcher1 prepared de-identified vector solutions and Researcher2 administered vector solutions, recording a syringe ID and the animal ID. Analysis was carried out in a blinded fashion, taking advantage of researchers completely naïve to the goals of the study to analyze subjective data whenever possible.

Mouse tissues for molecular studies involving DNA/RNA recovery were harvested as quickly as possible and frozen on dry ice. For histology, animals were transcardially perfused with heparinized 0.01 M PBS followed by 4% paraformaldehyde (PFA) in 0.1 M phosphate buffer (PB). The brain was post-fixed in 4% PFA at 4ºC overnight, subsequently equilibrated with 30% sucrose in PBS, embedded in OCT compound (Sakura), flash frozen in dry ice-equilibrated isopentane and stored at -80ºC.

Non-human primate: A list of non-human primate experiments, including sex, age, vector, dose, and post-injection interval can be found in Table 3.1. Indian rhesus macaques were bred and housed at the Oregon National Primate Research Center (Beaverton, OR). AAV2 and AAV9 neutralizing antibody assays and ELISAs to detect AAV-binding antibodies were used to screen all animals and only animals that were negative at a 1:5 serum dilution in both assays were utilized. Neutralizing antibody assays were carried out by the OHSU Molecular Virology Core.

Vector administration was performed under anesthesia by trained ONPRC veterinary staff via bilateral cephalic catheters. Infusion began at a rate of 0.6 mL/min/kg and was incrementally increased while monitoring animal health until a maximal rate of 4.8 mL/minute/kg was achieved. Vector solutions were diluted to a total volume of 10 mL/kg using 5% sorbitol solution in PBS.

At necropsy, animals were sedated with ketamine followed by a lethal injection of sodium pentobarbital. Animals were perfused with 2-3 L ice-cold saline. The brain was removed from the skull and 4 mm coronal slabs were made using a brain matrix. For molecular studies involving DNA/RNA extraction, tissue was dissected from regions of interest and frozen on dry ice. For histology, slabs were placed in 4% PFA for 48 hours and subsequently equilibrated with 30% sucrose solution. The spinal cord was removed from the vertebral column and carefully segmented between exiting dorsal root ganglia. Segments from the cervical, thoracic, and lumbar regions were divided and either fixed in 4% PFA for histology or frozen on dry ice for molecular studies. Various peripheral tissues including, but not limited to dorsal root ganglia, liver, heart, kidney, lungs, skeletal muscle, pancreas, and spleen were also collected.

All experiments were carried out in accordance with the OHSU and ONPRC Institutional Animal Care and Use Committee.

Table 3.1 List of non-human primate experiments

Experiment	Age (days)	Sex	Vector	Dose (vg/kg)	Interval
Directed evolution	296	M	ssAAV9-N272A-hSYN1-TRADE-PepLib0	2.8×10^{12}	2 weeks
Validation	478	M	scAAV-hSYN1-GFP-BCLib	2×10^{13}	2 weeks
Characterization	848	F	ssAAV9-CAG-FLAGnlsGFP-BCLib	1.5×10^{13}	3 weeks
			ssAAV9-N272A-HN1-CAG-HAnlsGFP-BCLib	1.5×10^{13}	
Characterization	310	M	ssAAV9-CAG-HAnlsGFP-BCLib	1.5×10^{13}	3 weeks
			ssAAV9-N272A-HN1-CAG-FLAGnlsGFP-BCLib	1.5×10^{13}	

Recovery of anti-sense transcripts

For the proof-of-concept study, total RNA was isolated from brain tissue using TRIzol reagent (Invitrogen) per the manufacturer protocol and DNase-treated with the DNA-free DNA Removal Kit (Ambion). 2 µg of DNase-treated RNA was reverse transcribed using oligo-dT and the RETROscript First Strand Synthesis Kit (Ambion). MMLV-RT was substituted with water for the RT-negative conditions. The entire cap open reading frame was then recovered as a 2403 bp amplicon by PCR using Platinum Pfx DNA Polymerase (Invitrogen) and the AAV9capF and AAV9capR primers. For PCR we used an initial denaturing step at 95ºC x 2min followed by 35 cycles of 95ºC x 15 seconds, 55ºC x 30 seconds, 68ºC x 3 minutes, and a final extension of 68ºC x 5 minutes. Subsequently, the region of the MVM intron was PCR amplified using the MVMF and MVMR primers and similar PCR conditions, adjusting the annealing temperature to 60ºC and the extension time to 75 seconds. Expected amplicon sizes were evaluated by gel electrophoresis. PCR products were then purified using QIAQuick Gel Extraction Kit (QIAGEN) and Sanger sequencing was outsourced to Elim Biopharmaceuticals Inc. (Hayward, CA). The cap ORF was sequenced using the AAV9PepF1 primer and the MVM intron was sequenced using the MVMseq primer.

For transcription-dependent directed evolution, total RNA was isolated using TRIzol and treated with the TURBO DNase Kit (Ambion). 2 µg of DNase-treated RNA was reverse transcribed using the AAV9RT1 primer which recognizes the hSYN1-driven antisense AAV9 cap sequence ~200 bp distal to the peptide region, or oligo-dT. A 355 bp amplicon spanning the peptide insertion region was recovered using SuperFi DNA Polymerase and the AAV9pepF1 and AAv9pepR1 primers. PCR conditions were 98ºC x 30s to denature followed by 35 cycles of 98ºC x 10 seconds, 55ºC x 15 seconds, 72ºC x 30 seconds, and a final extension of 72ºC x 5 minutes. For non-human primate samples, Reverse transcription was carried out as above, but recovery required a nested PCR strategy wherein the product of a first PCR using the AAV9capF1 x AAV9capR1 primers was used as the template for second PCR using the AAV9capF2 x AAV9capR2 primers.

Over the course of these studies, the RETROscript kit was discontinued and we switched to the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). The Platinum Pfx DNA Polymerase was also replaced by the Platinum SuperFi DNA Polymerase and we switched to the TURBO DNase Kit (Ambion) due to its increased efficiency.

In vivo biopanning of AAV libraries

In vivo selection was carried out in 8-week-old C57BL/6J male mice. The AAV9-N272A-hSYN1-TRADE-Lib0 vector was administered into two mice and brain tissue was harvested 12 days post-injection. In this first round, we chose to use a brain region that was relatively enriched for neurons and thus extracted RNA from crudely dissected frontal cortex using TRIzol reagent. Total RNA was isolated from each animal

independently, mixed in a 1:1 ratio, and DNase-treated. A region containing the peptide insertion was recovered by RT-PCR and assembled in a TRADE vector backbone to generate a plasmid library representing capsid sequences that had undergone one round of selection (i.e., pAAV9-N272A-hSYN1-TRADE-Lib1). This plasmid library was used to generate the corresponding AAV library and the process was repeated for 3 total rounds of selection in mice. Rounds 2 and 3 were carried out in a similar manner to the first round with a few modifications. In the latter rounds, we administered doses ranging from 1×10^{12} to 1×10^9 vg per mouse in 10-fold dilutions (n = 2 animals per dose), but only samples from the animals administered 1×10^{11} vg dose were carried forward. In addition, we isolated total RNA from entire hemispheres of mouse brain, reasoning that regional neuronal enrichment was unlikely to be necessary after one round of selection.

In parallel, a single round of selection was performed using a male rhesus macaque. The AAV9-N272A-hSYN1-TRADE-Lib0 vector was administered intravenously and the animal was euthanized 14 days post-injection. Tissues were collected and frozen on dry ice.

RT-PCR amplicons of *cap* containing the peptide fragment were sub-cloned into a plasmid backbone, transformed into DH10B cells, and individual clones were mini-prepped and sent for Sanger sequencing (ElimBio). We selected the top five most common variants following three rounds of selection in C57BL/6J mice, as well as 21 variants that were recovered from various regions of the rhesus brain after a single round of selection, for further screening.

AAV DNA/RNA Barcode-Seq Analysis

Viral barcode clones were individually packaged into novel, TRADE-identified, capsids using the triple transfection method described above. In addition, we generated barcoded vectors in key reference capsids including AAV9, AAV9-N272A (parental), and AAV-PHP.B. The number of clones for each variant is given in Table 3.3. Crude lysate titers of DNase-resistant vector genomes were determined by radioactive dot blot, probing against the hSYN1 promoter. Each barcoded viral clone was then mixed at an equimolar ratio and the pooled crude lysate was purified as a library as described above to generate the scAAV-hSYN1-GFP-BCLib. Although we have utilized this method in the past, this method assumes that the purification process does not differentially impact different AAV variants. In the course of these studies, we found that this assumption may not always be valid.

Processing of samples and AAV Barcode-Seq analysis were performed as previously described (Adachi et al. 2014). In brief, DNA was extracted from animal tissues using the KingFisher Flex Purification System (Thermo), and from blood using the Extract-N-Amp Blood PCR kit (Sigma). RNA extraction from animal tissues and reverse-transcription were performed as described above. AAV vector genomes from the purified AAV vector stock were extracted using Proteinase K (Invitrogen) and the DNA Extractor Kit (Wako). In addition, AAV genomes were extracted from the mixed crude lysate prior to purification, but treated with Benzonase (0.2 U/mL) overnight at 37C prior to Proteinase K treatment and DNA extraction. Human hepatocytes were harvested from highly engrafted xenograft FRG chimeric mice (serum human albumin = 3.4-6.7 mg/mL) by collagenase perfusion followed by staining and fluorescence activated cell sorting, as

previously described (Azuma et al. 2007). DNA was extracted from human hepatocytes using the MasterPure Complete DNA and RNA Purification Kit (Lucigen).

For each sample, both viral barcodes (VBC1 and VBC2) were amplified by independent indexing PCR reactions, using primers (VBC1F x VBC1R and VBC2F x VBC2R) containing unique sample barcodes (SBCs). We then generated a pool of viral barcode amplicons by mixing each PCR reaction at an equimolar ratio based on ethidium bromide densitometry. The pooled amplicons were purified and adapted for Next-Generation Sequencing (NGS) using an Illumina HiSeq 2000 as previously described (Adachi et al. 2014). Raw sequence quality was assessed using FASTQC.

We utilized custom programs to analyze NGS data. Reads were binned based on SBCs and VBC counts were summed within each SBC. VBC counts were normalized to read counts from the AAV vector input library and subsequently underwent global normalization to the mean number of reads within each SBC to calculate global normalized ratios (GNR). In these studies, a single GNR value was calculated for each VBC clone by averaging the independently calculated VBC1 and VBC2 GNR values. Therefore, the number of GNR values representing each AAV variant in the library is equal to the number of VBC clones in the library representing that particular variant multiplied by the number of animals. We report phenotypic difference (PD) values, which are equal to the mean of GNR values that have been normalized to the mean of AAV9 GNR values within each SBC to calculate phenotypic difference (PD) values. Therefore, the PD of AAV9 is always 1.0 and PD values calculated from DNA and RNA barcode reads measure biodistribution and transduction efficiency relative to AAV9, respectively.

In the original AAV BC-seq report, we stated that global normalized ratio values may not be normally distributed and therefore utilized a non-parametric Mann-Whitney U-test. In the current study, we analyzed all AAV9 GNR distributions for normality using the Shapiro-Wilks test implemented in R using the rstatix package. Our rationale for excluding non-AAV9 variants is that the low sample number would be underpowered to detect deviation from normality and we would, therefore, underestimate the frequency of non-normal distributions. Of 430 samples, 31 rejected the null hypothesis of normality with alpha set at 0.05. Thus, the observed rejection rate of 0.072 is close to the expected rate. Put another way, we expected that if the null hypothesis of normality was true, then the distribution of p-values would be roughly uniform and this is indeed what we observed (Figure 3.1). Based on this analysis, we determined that parametric tests could be used and implemented Welch's t-tests with false discovery rate correction for multiple pairwise comparisons between AAV variant PD scores.
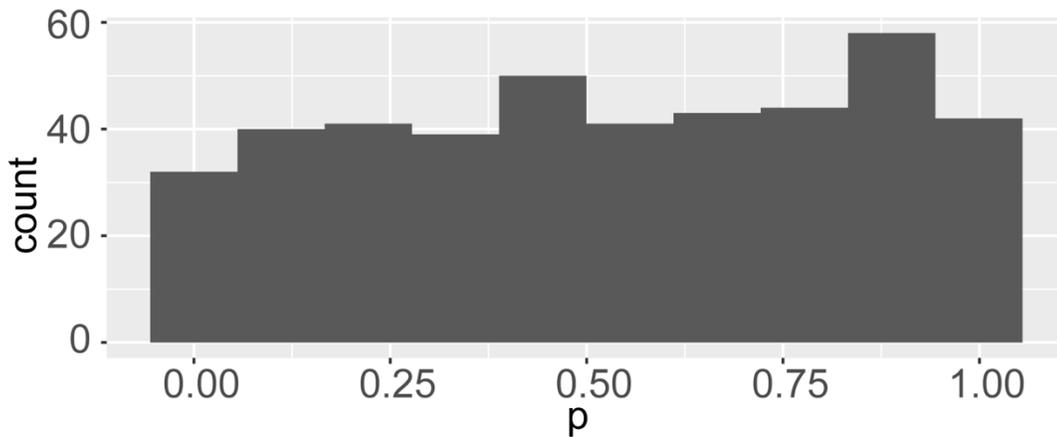
Figure 3.1 Histogram of p-values following Shapiro-Wilks tests of normality for 430 distribution of AAV9 global normalized ratios
We analyzed data from 430 samples associated with the scAAV-hSYN1-GFP-BCLib. GNR values were calculated for VBCs associated with AAV9 and the distributions were tested for normality using a Shapiro-Wilks test with alpha set at 0.05. Of the 430 distributions, 31 rejected the null hypothesis of normality, a rate of 0.072.

Total RNA was extracted from one hemisphere of each mouse brain using

TRIzol. Replicate animals from a round of selection were pooled. mRNA was purified

using the OligoTex mRNA mini kit (QIAGEN), digested with TURBO DNase (Thermo),

and reverse transcribed using the High-Capacity cDNA Reverse Transcription kit

(Applied Biosciences). Indexing PCR was used to amplify peptide sequences using the

AAV9pepF3 and AAV9pepR3 primers containing SBCs. Such amplicons were named

RNALib1, RNALib2, and RNALib3 where the number refers to the number of in vivo

selection rounds that have been applied to that library of sequences. Additionally, vector

genomes were extracted from the original AAV stock library (named AAVLib0) and

used as a template for indexing PCR. Amplicons were pooled in an equimolar ratio,

purified, and prepared for NGS as described above. Quality was assessed by FastQC.

We first generated a custom script to extract peptide sequences from raw reads

and group them based on SBC. Next, we tallied read counts for each unique peptide

sequence and generated a summary table for each SBC. Subsequently, we attempted to

correct for Illumina error by applying a bottom-up sequence collapsing algorithm to each

summary table. The final summary table was filtered for peptide sequences of the

expected 8mer length.

Enrichment scores for each peptide sequence were calculated by dividing the

proportion of reads in each RNALib by the proportion of reads in AAVLib0. If a peptide

sequence was detected in RNA, but not the original AAVLib0, we imputed a pseudo-read

count of 1. We then generated a list of all peptide sequences in RNALib2 and RNALib3

with enrichment scores above 5.8 and 6.4, respectively. These values correspond to the

enrichment score when the raw read count in the RNALib and AAVLib0 with both 1. Different values are due to different sequencing depths. We then performed a top-down (starting with the most common peptides sequence) grouping analysis to identify peptide sequences within a Hamming distance of 1, selecting the most common sequence as the exemplar of each group.

The resulting list of 416 peptides (105 from RNALib2 only, 175 from RNALib3 only, and 136 from both libraries) was combined with a random sample of 80 peptide sequences from AAVLib0. The total list of 496 peptide sequences were then backtranslated to generate five different DNA sequences per 8 amino acid peptide sequence. In order to optimize mismatches between synonymous sequences, we generated a program to randomly sample five sequences from a list of all possible backtranslation sequences. For each sample, all 10 pairwise Hamming distances were calculated and multiplied to create an aggregate score of mismatches. We repeated this 10,000 times for each peptide sequence and selected the combination with the highest aggregate Hamming distance score. Combinations were subsequently screened for the following criteria: (1) a minimum pairwise Hamming distance of at least 3, (2) lack of homology with the 10 nucleotides at the 3' end of any primer set to be used for cloning or sequencing (allowing for 1 mismatch), and (3) no consensus poly-adenylation signal (defined as AATAAA) in the forward or reverse directions.

DNA sequences were synthesized as an oligo pool (Genscript) containing homology arms to facilitate DNA assembly into the pAAV9-N272A-2BsaI-hSYN1-TRADE plasmid. We generated the plasmid library and AAV library as described above. The amount of ITR-containing TRADE vector plasmid was reduced to 1/30th the level of

the Rep and Adenoviral helpers in order to minimize mosaicism and cross-packaging. The AAV library was administered intravenously into C57BL/6J and BALB/cJ mice (n = 4 / strain) at a dose of 1×10^{11} vg. Brain tissue was harvested two weeks post-injection and RT-PCR was used to generate RNA libraries for enrichment analysis as described above.

We carried out Shapiro-Wilks tests for normality and found that ~81% of enrichment score distributions rejected the null hypothesis of normality. Therefore, we turned to a non-parametric Wilcoxon test. In the absence of any reference capsid variants in the library, we tested the null hypothesis that the enrichment scores were less than or equal to 1 using the rstatix package implement in R. Rejection of the null hypothesis would then indicate that the peptide made up a greater proportion of RNA reads following transcription-dependent selection than DNA reads in the AAV input stock library.

Histological evaluation of transduction efficiency and tropism in mouse

Mouse brains were cryosectioned in the sagittal plane at a thickness of 16-20 μm and directly mounted on glass slides, or as free-floating 40 μm sections placed in PBS. Sections were washed in PBS then blocked in 5% donkey serum (Sigma) + 0.5% Triton-X100 in PBS for one hour at room temperature. Primary antibodies (Table 3.2) were diluted in blocking buffer and applied overnight at 4ºC. After washing, sections were incubated with secondary antibodies (Table 3.2) diluted in blocking buffer at room temperature for two hours. Samples were washed with PBS, counterstained with Hoechst

33342 (MolecularProbes, 1:10,000), and coverslipped using Fluoromount-G (SouthernBiotech).

For quantification of neuronal transduction using the self-complementary hSYN1-driven GFP promoter, slides containing stained brain sections were tilescanned at the OHSU Advanced Light Microscopy Core using a Zeiss AxioScan.Z1 and 20X objective. Image analysis was carried out by an observer blinded to experimental groups as well as the overall goals of the study. The blinded observer used Zen Blue 2.6 software to perform semi-automated counts of GFP+, NeuN+, and Hoechst+ loci using the dynamic thresholding function for segmentation and counting. Smoothing and thresholding values were manually adjusted by the blinded observer to optimize the accuracy of the counts. We quantified transduction in 1.5 mm x 1.5 mm areas. Six brain sections were stained and imaged per animal and technical replicates were pooled to obtain a single neuronal transduction efficiency for each region of each animal. The hSYN1 promoter is highly specific for neurons and we did not observe any instances of GFP+/NeuN- cells. Therefore, we calculated neuronal transduction efficiency by dividing the number of GFP+ cells by the number of NeuN+ cells. Neuronal transduction efficiencies were compared across AAV capsid variants, but within mouse strains and regions using one-way ANOVA and Tukey's Honestly Significant Difference post-hoc comparisons with alpha = 0.05.

For quantification of cell type tropism and transduction efficiency using the CAG-nlsGFP vectors, slides containing stained brain sections were prepared as described above and imaged using a Keyence BZ-X710 equipped with CFI Plan Apo lambda 20X objective and structured illumination module to achieve optical sectioning. Manual

counts of transduction and colocalization were carried out by an observer blinded to experimental groups. Counts were made in three brain sections per animal. Images were first viewed with the GFP and Hoechst channels. Double positive loci were then overlaid with a cell type specific marker channel. We utilized NeuN as a marker of neurons, S100 as a marker of glial cells that are predominantly astrocytes, Olig2 as a marker of cells in the oligodendrocyte lineage, and Iba1 as a marker of microglia. Each cell type was assayed independently. To measure tropism, we calculated the percentage of transduced cells (GFP+) that were positive and negative for each cell type-specific marker. To measure relative transduction efficiency, we used the same dataset, but focused on the average number of cells that were double positive for GFP and each cell type marker within a 20X field. We compared tropism for neurons and glia for each variant in each brain region, as well as transduction efficiencies between capsid variants for neurons and glia within each brain regions, using student's t-tests with Holm-Bonferroni correction for multiple comparisons.

Histological evaluation of transduction efficiency and tropism in NHP

Histological evaluation in large animal studies is often challenging due to the recruitment of relatively small samples consisting of heterogeneous subjects with respect to age, sex, weight, environmental exposures, genetic background, and other factors. Therefore, we developed a strategy to allow for a head-to-head comparison of AAV-mediated transgene protein expression within a single test subject. We generated single strand AAV genomes that expressed an epitope-tagged and nuclear-localized GFP transgene under the control of a ubiquitously expressing CAG promoter-enhancer

sequence (see above). In addition, these constructs harbored AAV DNA/RNA Barcode-Seq cassettes in the 3'untranslated region. The idea is that transduction by either vector could be determined by detection of GFP signal and the capsid identity could then be determined based on detection of the epitope tag. In order to control for differences in antibody sensitivity we generated vectors with all four capsid-epitope relationships (i.e., AAV9-FLAG, AAV9-HA, HN1-FLAG, HN1-HA). We then mixed reciprocal pairs to generate two vector libraries and injected each library into one animal. The female animal (Animal1) received a vector mixture such that detection of HA represented HN1 transduction, while detection of FLAG represented AAV9 transduction. The male animal (Animal2) received the reciprocal mixture such that detection of HA represented AAV9 transduction, while detection of FLAG represented HN1 transduction.

Rhesus brain was sectioned in the coronal plane at a thickness of 40 μm using a freezing stage microtome and sections were stored free-floating in a cryoprotectant solution (30% sucrose + 30% ethylene glycol in 0.5X PBS) at -20ºC. Staining was carried out as described above.

For quantification of relative transduction efficiency and cell type tropism in rhesus, we first acquired whole section tilescanned images using a Zeiss AxioScan.Z1 and 10X objective housed at the OHSU Advanced Light Microscopy Core. Subsequently, we generated large optically-sectioned tilescan images of specific brain regions using the Keyence BZX-710 microscope with structured illumination module. Due to the 40um thickness, optical sectioning was necessary for confident determination of colocalization. We manually identified transduced cells (GFP+) in each brain region and determined the cell to be either positive or negative for staining against HA and the cell type specific

marker (either NeuN or S100). Using this information, we calculated the proportion of transduced neurons (GFP+/NeuN+) and glia (GFP+/S100+) that could be attributed to HN1 or AAV9. Subsequently, we measured tropism by calculating the proportion of transduced (GFP+) cells that were neuronal (NeuN+) or non-neuronal (NeuN-), and glial (S100+) or non-glial (S100-). The observed proportion of transduced cells attributable to HN1 or AAV9 was compared to a null hypothesis assuming no difference in transduction efficiency using a one-proportion z-test (mu = 0.5). To compare AAV9 and HN1 tropism, we performed a two-proportion z-test comparing AAV9 to HN1 based on the proportion of GFP+ cells that were positive for NeuN or S100. Given the sample size of 1, we corrected p-values with the stringent Bonferroni method.

Table 3.2 List of antibodies utilized for immunofluorescent studies

Host	Target (Conjugate)	Vendor	Product	Working Dilution
chicken	anti-GFP	Abcam	ab13970	1:1000
mouse	anti-FLAG	Sigma	F1804-50UG	1:1000
rat	anti-HA	Sigma	11867423001	1:500
rabbit	anti-NeuN	Millipore	ABN78	1:1000
goat	anti-ChAT	EMDMillipore	AB144P	1:100
rabbit	anti-Olig2	Millipore	AB9610	1:1000
rabbit	anti-Iba1	Wako	019-19741	1:500
rabbit	anti-S100	Dako	GA504	1:2000
rabbit	anti-PGP9.5	MilliporeSigma	AB1761-I	1:500
donkey	anti-chicken (AF488)	JacksonImmunoResearch	703-545-155	1:500
donkey	anti-rat (Cy3)	JacksonImmunoResearch	712-165-153	1:500
goat	anti-chicken (AF647)	LifeTechnologies	A21449	1:500
donkey	anti-goat IgG (AF647)	JacksonImmunoResearch	705-605-147	1:500
donkey	anti-rabbit (AF647)	ThermoFisher	A-31573	1:500
goat	anti-mouse (AF647)	JacksonImmunoResearch	115-605-166	1:500

Quantification of GFP in brain lysates

Mouse cortex was microdissected at the time of harvest and frozen on dry ice. We quantified GFP protein expression in these samples using the GFP CatchPointSimpleStep ELISA Kit (Abcam, ab229403). In brief, brain tissue was homogenized in buffer PTR using a 28G insulin syringe to apply shear force. Total protein concentration in cleared supernatants was quantified using the BCA assay (Thermo). We loaded 300 or 30 ng total protein per well and carried out the ELISA in duplicate per the manufacturer protocol. Fluorescence was measured using a FLUOstar OMEGA (BMG Labtech). MARS Data Analysis Software (BMG Labtech) was used to model a 4-parametric logarithmic fit (r^2 = 0.998) and calculate concentrations of GFP in each sample. Because groups did not exhibit equivalent variance, GFP concentrations were compared across AAV capsid variants and within mouse strains using Welch's one-way ANOVA with post-hoc Welch's t-tests corrected with the Holm-Bonferroni method.

Hydrophobicity analysis

We determined the relative abundance of AAV variants in the scAAV9-hSYN1-GFP-BCLib before and after purification using the AAV DNA Barcode-Seq pipeline and then calculated the relative recovery by dividing the relative abundance post-purification to the relative abundance pre-purification. We observed the formation of two groups that could easily be separated by drawing a threshold relative recovery score of 0.5. Nevertheless, we used a more objective approach to categorizing the variants by standard clustering algorithms. We determined the optimal number of clusters using the Hartigan-Wong elbow method as implemented by the R stats package. Subsequently, we

performed k-means clustering (k = 2) to formally categorize the variants into high and

low recovery groups.

We calculated the proportion of all amino acids that were hydrophobic (here

defined as ALIMFWV) within each group. The proportion of amino acids within the high

and low recovery groups were compared using a two-proportion z-test. We also

calculated the proportion of hydrophobic amino acids at each position, excluding the

AAV-PHP.B variant from this analysis due its lack of glycine-serine linkers and 7mer

size that may obviate any comparison. For the positional analysis, no statistical test was

attempted given the relatively small sample size.

Plots were generated using the ggplot2 package in R(3.6.0). Statistics were carried

out using the rstatix wrapper of base stats package in R. Figures were formatted using

Adobe Illustrator. Tables were formatted in Microsoft Excel.

3.4   Results

In vivo selection of AAV variants that efficiently transduce brain neurons following systemic administration

In order to apply the TRADE system toward identification of an AAV vector that

efficiently transduces brain neurons following systemic administration, we generated a

TRADE construct using a hSYN1 promoter-enhancer sequence that expresses

specifically in neurons. Many labs have now utilized the AAV9-7mer peptide display

library that places 7 random NNK-encoded amino acids on the three-fold axis (AAV9

VP1 position 588) (Perabo et al. 2003; Müller et al. 2003; Deverman et al. 2016; Hanlon

et al. 2019; Nonnenmacher et al. 2020). Rather than resampling the same protein landscape, we sought to explore a similar, but distinct library design. Thus, we generated a library with the following modifications: (1) we incorporated the N272A mutation into AAV9 cap, which results in a strong liver detargeting phenotype without compromising biodistribution to the brain (S. J. Huang et al. 2017), (2) we utilized an NNK-encoded 8mer flanked by glycine-serine linkers (GGGS[NNK]$_8$GGGGS), and (3) we performed a substitution at AAV9 VP1 Q588, rather than an insertion (Figure 3.2a).

Iterative in vivo biopanning using the TRADE platform is outlined in Figure 3.2b. We generated the pAAV9-N272A-hSYN1-GFP-TRADE-Lib0 plasmid library ("0" indicates the number of *in vivo* selection rounds that the library has undergone) using standard cloning techniques and achieved a total colony number (i.e., estimated diversity) of ~1.6 x 10^7. We then utilized the plasmid library to produce an AAV library, correspondingly named AAV9-N272A-hSYN1-GFP-TRADE-Lib0. Next-generation sequencing of DNase-resistant vector genomes completed at the end of this study measured an 8mer diversity of ~3.8 x 10^6 at the amino acid level, but with a depth of only 9.8 x 10^6 reads. The library was administered intravenously into 8 week-old C57BL/6J mice (n = 2, dose = 3 x 10^{11} vg per mouse or ~1 x 10^{13} vg/kg) and one male rhesus macaque (n = 1, dose = 2.7 x 10^{12} vg/kg). Tissues were harvested 12-14 days after injection.

In the first round of selection, we reasoned that we could enrich for a neuronal cell population by using cortical tissue. Therefore, total RNA was extracted from grossly dissected mouse frontal cortex and subsequently used to generate a 355 bp amplicon containing the peptide insertion by RT-PCR (gel in Figure 3.2b). This fragment was

cloned into the pAAV9-N272A-2BsaI-hSYN1-TRADE backbone to generate the

subsequent library, pAAV9-N272A-hSYN1-TRADE-Lib1. The cycle was repeated for a

total of three rounds of selection in mice. However, in subsequent rounds we extracted

RNA from an entire brain hemisphere and we also reduced the dose to 1 x 10^{11} vg (~3.3 x

10^{12} vg/kg). We then sub-cloned the RT-PCR product and Sanger sequenced 58 clones to

identify the most common 8mer peptide insertions, ultimately selecting the five most

common sequences for further evaluation (Table 3.3).

We also isolated total RNA from various rhesus brain regions and attempted to

recover variant sequences. Unfortunately, we failed to recover the 355 bp amplicon using

the same RT-PCR conditions utilized in the mouse study. We thus utilized a nested PCR

strategy (see Methods), which allowed for successful amplification of the peptide region,

but we were only able to recover 1-2 unique peptide sequences per RT-PCR reaction.

Nevertheless, we selected 21 rhesus-derived variants for further study (Table 3.3).
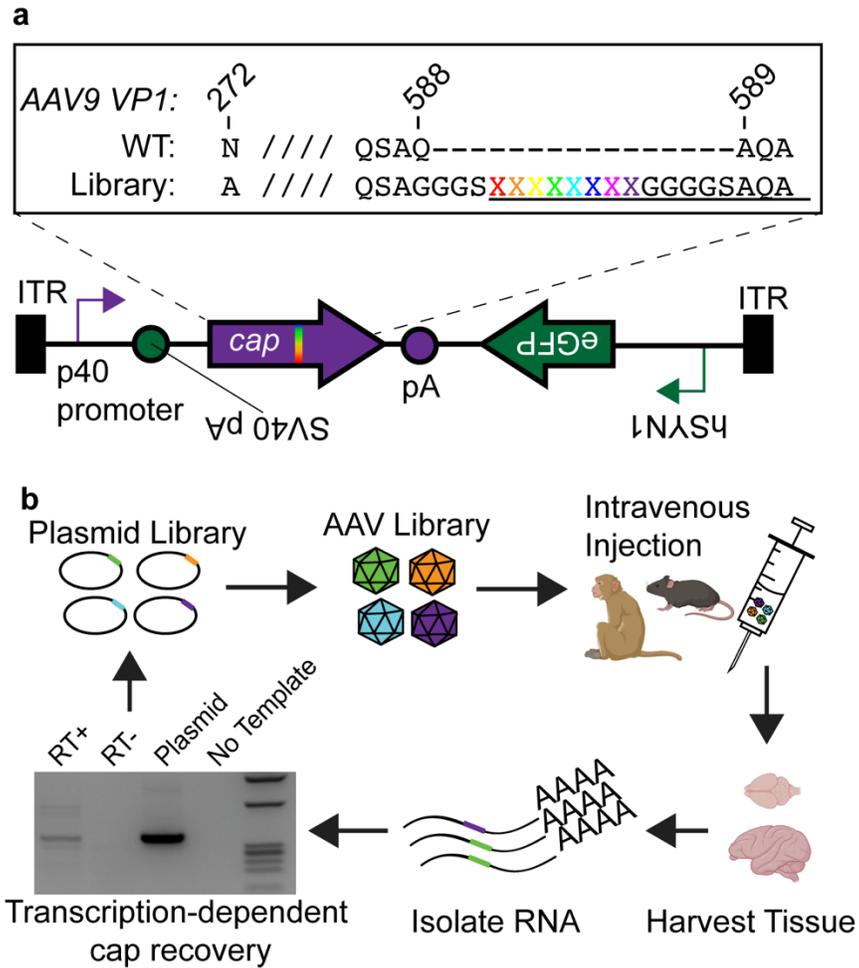
Figure 3.2 Experimental design for in vivo selection of AAV variants with enhanced neuronal transduction using TRADE

(a) A plasmid peptide display library was generated using NNK motifs to encode random 8mer peptides that were flanked by glycine-serine linkers and substituted at Q588 of the AAV9-N272A VP1 sequence. (b) The plasmid library was used to generate a corresponding AAV library that was biopanned in C57BL/6J mice and rhesus macaque using the TRADE system.

TRADE identifies AAV variants with enhanced brain neuronal transduction following intravenous administration

We next sought to validate and benchmark the TRADE-identified variants using our high-throughput AAV Barcode-Seq technology. We adapted the original AAV DNA Barcode-Seq pipeline (Adachi et al. 2014) to allow evaluation of vector-mediated RNA expression (i.e., AAV RNA Barcode-Seq, Adachi & Nakai, unpublished) and generated novel constructs, pdsAAV-hSYN1-GFP-BC, to express unique hSYN1-driven mRNA barcode sequences downstream of a GFP reporter (Figure 3.3a). Viral barcode (VBC) sequences were derived from a pool of previously validated 2 x 12 nucleotide VBC cassettes. Importantly, the use of the hSYN1 promoter-enhancer allowed us to focus on the identification of variants with enhanced neuronal transduction efficiency, whereas use of a ubiquitous promoter might confound our validation if glial transduction, but not neuronal transduction, was enhanced.

We individually packaged the pdsAAV-hSYN1-GFP-BC vector genomes into the selected AAV variant capsids to generate the scAAV-hSYN1-GFP-BCLib. Each capsid variant was used to package two unique VBC clones that were used as an internal control in downstream analysis. In addition, the barcoded library also included AAV9 (n = 15 VBC clones), the parental AAV9-N272A (n = 5 VBC clones), and AAV-PHP.B (n = 2 VBC clones) as reference controls.

This vector was intravenously administered to C57BL/6J and BALB/cJ mice at a dose of 5 x 10^{11} vg (n = 3 mice / strain, ~1.7 x 10^{13} vg/kg) via tail vein. Tissues were harvested two weeks post-injection. Brain tissue was grossly dissected to separate the forebrain from the brainstem and cerebellum. Total RNA was extracted, reverse-

transcribed, and indexed barcode amplicons were generated by PCR and submitted for next-generation sequencing. AAV RNA Barcode-seq analysis revealed that, as expected, AAV-PHP.B demonstrated enhanced transduction of brain neurons compared to AAV9 in C57BL/6J mice, but not BALB/cJ mice (Hordeaux et al. 2018, 2019; Q. Huang et al. 2019), though the degree of enhancement was less than previously reported based on other methods (Deverman et al. 2016) and variation between animals was high (Figure 3.3b).

Among the mouse-derived variants, two out of five, HN1 and HN2, showed significantly enhanced transduction of neurons in all three C57BL/6J brain tissue samples. Notably, for both variants, enhancement was more pronounced in the forebrain (Figure 3.3b and Table 3.4). The HN1 variant exhibited 8.48-fold greater transduction than AAV9 in forebrain, ($p < 0.001$, Welch's t-test with FDR correction), but only 3.96-fold and 2.55-fold in the brainstem and cerebellum, respectively ($p < 0.01$ and $p < 0.05$, respectively). We observed a similar degree and pattern of HN1 enhancement in the BALB/cJ forebrain (7.52-fold, $p < 0.0001$), brainstem (4.12-fold, $p < 0.0001$), and cerebellum (2.34-fold, $p < 0.01$) (Figure 3.3b, Table 3.5).

Overall, rhesus-derived variants did not demonstrate enhanced neuronal transduction efficiency in rodents.

Table 3.3 List of AAV variants identified by hSYN1-TRADE and included in the scAAV-hSYN1-GFP-BCLib

AAV Variant	ID	Origin	# VBC clones
AAV9	AAV9	Reference	15
AAV9-N272A	AAV9-N272A	Reference	5
AAV-PHP.B	AAV-PHP.B	Reference	2
AAV9-N272A-TTNLAKNS	HN1	C57BL/6J	2
AAV9-N272A-QQNGTRPS	HN2	C57BL/6J	2
AAV9-N272A-HQVTSSGA	HN3	C57BL/6J	2
AAV9-N272A-LLVTARSH	HN4	C57BL/6J	2
AAV9-N272A-VVQGEQKR	HN5	C57BL/6J	2
AAV9-N272A-SGQRVGSD	HN6	Rhesus	2
AAV9-N272A-GNINVVPH	HN7	Rhesus	2
AAV9-N272A-GSPAASSW	HN8	Rhesus	2
AAV9-N272A-KHSLTLES	HN9	Rhesus	2
AAV9-N272A-KPVSTDTF	HN10	Rhesus	2
AAV9-N272A-LDRSGSTG	HN11	Rhesus	2
AAV9-N272A-LGAQNHVV	HN12	Rhesus	2
AAV9-N272A-LRATDYGP	HN13	Rhesus	2
AAV9-N272A-MERTEPLG	HN14	Rhesus	2
AAV9-N272A-NDGLRLHL	HN15	Rhesus	2
AAV9-N272A-NLSAHSHD	HN16	Rhesus	2
AAV9-N272A-RALDLVTR	HN17	Rhesus	2
AAV9-N272A-SAGMARNS	HN18	Rhesus	2
AAV9-N272A-TAQGAAFR	HN19	Rhesus	2
AAV9-N272A-TGRPEQPK	HN20	Rhesus	2
AAV9-N272A-THSPIKLP	HN21	Rhesus	2
AAV9-N272A-TQFSQAQR	HN22	Rhesus	2
AAV9-N272A-VGDSANLR	HN23	Rhesus	2
AAV9-N272A-AVAGDRLL	HN24	Rhesus	2
AAV9-N272A-DLLTRSVS	HN25	Rhesus	2
AAV9-N272A-EWKTQLAL	HN26	Rhesus	2
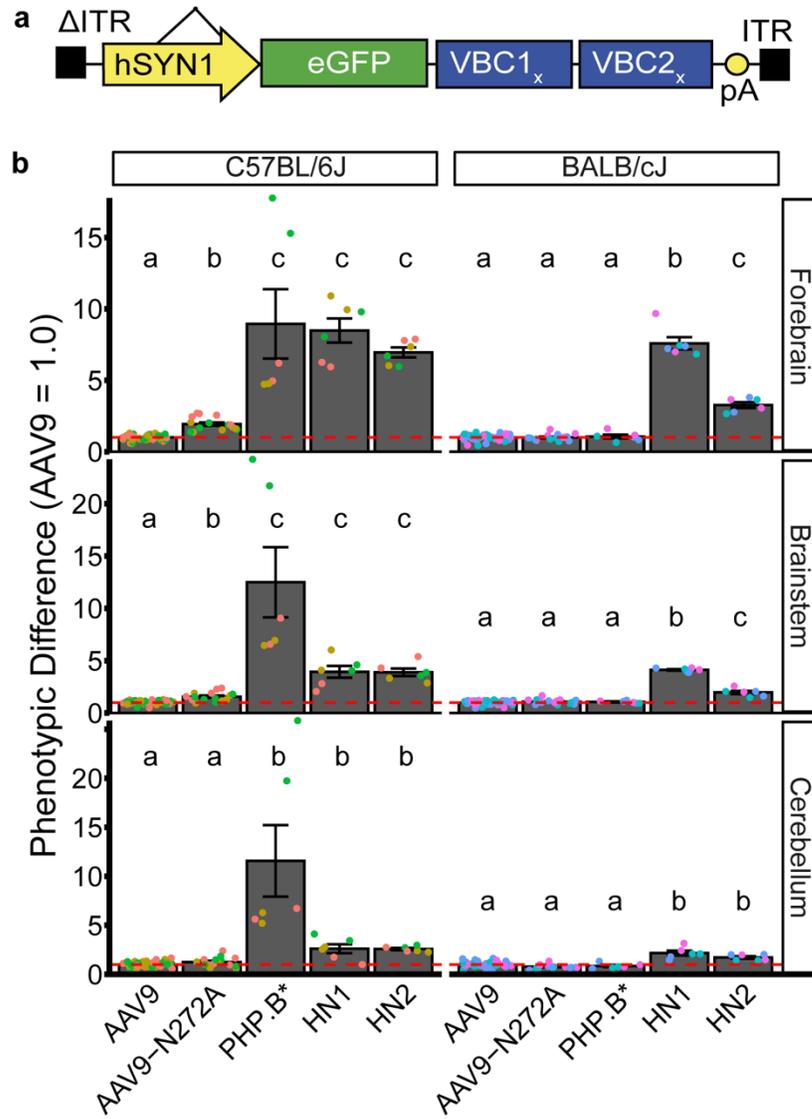
Figure 3.3 AAV RNA Barcode-Seq demonstrates enhanced neuronal transduction efficiency of TRADE-identified variants

(a) Map of the pdsAAV-hSYN1-GFP-BC construct, a self-complementary AAV genome expressing a reporter gene (GFP) and viral barcodes under the control of the neuron-specific hSYN1 promoter-enhancer sequence. (b) Mean PD value ± SEM are plotted for each variant. Each point represents a PD calculated from a single barcode clone by averaging VBC1 and VBC2. Colors represent different animals. Variants exhibiting statistically different PD values based on Welch's t-test are denoted by unique letters. Statistical comparisons were only made within a given mouse strain and tissue sample. P-values were corrected using the false discovery rate (FDR) method and compared to an alpha value of 0.05. AAV-PHP.B* transduction is underestimated.

Table 3.4 AAV Barcode-Seq Phenotypic Difference Values in C57BL/6J Mice

| | RNA | | | DNA | | | | | | | |
	Forebrain	Brainstem	Cerebellum	Heart	Kidney	Liver	Lung	Muscle	Pancreas	Spleen	Testis
AAV9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AAV9-N272A	1.94	1.56	1.23	0.81	2.07	0.06	0.50	0.17	0.24	4.41	1.42
PHP.B	8.94	12.61	11.64	0.92	2.32	0.27	0.81	0.36	0.61	2.67	1.45
HN1	8.48	3.96	2.55	0.21	0.62	0.24	3.41	0.36	0.42	0.41	0.64
HN2	6.94	3.89	2.55	0.35	0.56	0.24	2.54	0.32	0.33	0.65	0.91
HN3	1.27	1.06	0.72	0.62	1.29	0.05	0.28	0.14	0.15	2.81	0.91
HN4	0.34	0.20	0.16	0.64	3.21	0.02	0.52	0.09	0.22	4.31	1.23
HN5	0.75	0.53	0.40	0.63	2.47	0.06	0.40	0.12	0.17	2.80	1.08
HN6	2.32	1.67	1.28	0.95	1.28	0.09	0.45	0.20	0.25	2.54	1.45
HN7	1.13	0.90	0.60	0.40	0.54	0.25	0.17	0.23	0.10	0.75	0.43
HN8	0.70	0.37	0.36	0.60	2.57	0.02	0.45	0.08	0.16	3.23	1.06
HN9	1.59	1.08	0.88	0.61	1.47	0.08	0.30	0.15	0.16	2.40	0.82
HN10	1.42	1.13	0.80	0.64	1.47	0.04	0.24	0.12	0.16	2.63	0.97
HN11	1.38	1.13	0.87	0.72	1.34	0.05	0.29	0.15	0.17	3.12	1.12
HN12	0.86	0.59	0.41	0.26	0.42	0.24	0.16	0.20	0.06	0.49	0.21
HN13	1.65	1.40	1.27	0.77	1.22	0.10	0.28	0.19	0.18	2.33	0.93
HN14	0.48	0.29	0.23	0.46	1.30	0.01	0.20	0.07	0.14	1.85	0.65
HN15	0.54	0.27	0.16	0.68	3.43	0.01	0.58	0.10	0.23	4.65	1.30
HN16	1.54	0.94	0.66	0.58	1.27	0.02	0.32	0.11	0.22	3.67	1.07
HN17	0.29	0.20	0.10	0.50	2.82	0.01	0.41	0.07	0.19	3.79	1.05
HN18	1.52	1.03	0.80	0.68	1.74	0.11	0.28	0.17	0.14	2.08	0.92
HN19	0.59	0.29	0.25	0.57	3.05	0.01	0.51	0.09	0.18	3.96	1.16
HN20	1.79	1.25	0.97	0.90	1.08	0.18	0.29	0.23	0.18	1.88	0.89
HN21	0.29	0.12	0.16	0.49	2.50	0.01	0.45	0.06	0.14	3.42	1.01
HN22	0.90	0.37	0.32	0.51	2.55	0.02	0.42	0.10	0.18	3.32	1.06
HN23	2.53	1.75	1.40	0.72	1.03	0.06	0.30	0.15	0.21	2.27	1.12
HN24	0.84	0.43	0.29	0.65	2.85	0.02	0.48	0.11	0.20	3.91	1.23
HN25	0.49	0.30	0.20	0.57	2.61	0.01	0.47	0.09	0.21	3.86	1.10
HN26	0.25	0.19	0.15	0.64	3.07	0.01	0.52	0.08	0.26	4.24	1.21

**Table 3.5 AAV Barcode-Seq Phenotypic Difference Values in BALB/cJ Mice**

| | RNA | | | DNA | | | | | | | |
	Forebrain	Brainstem	Cerebellum	Heart	Kidney	Liver	Lung	Muscle	Pancreas	Spleen	Testis
AAV9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AAV9-N272A	1.02	1.08	0.85	0.44	1.22	0.03	0.64	0.29	1.22	2.75	0.17
PHP.B	1.05	1.05	0.92	0.57	1.30	0.22	0.64	0.39	1.12	1.69	0.30
HN1	7.52	4.12	2.34	0.17	0.51	0.05	3.74	0.21	0.28	0.31	0.36
HN2	3.28	1.96	1.96	0.31	0.51	0.07	3.12	0.25	0.39	0.69	0.43
HN3	0.56	0.50	0.55	0.34	1.17	0.02	0.34	0.14	0.53	1.33	0.08
HN4	0.19	0.21	0.11	0.30	1.64	0.01	0.48	0.17	1.15	2.70	0.09
HN5	0.31	0.36	0.24	0.32	1.50	0.02	0.34	0.16	0.84	1.77	0.09
HN6	1.01	0.95	0.88	0.51	1.04	0.04	0.41	0.20	0.59	1.38	0.14
HN7	0.54	0.59	0.33	0.27	0.65	0.12	0.22	0.16	0.32	0.75	0.11
HN8	0.37	0.28	0.21	0.28	1.53	0.01	0.37	0.14	0.94	2.11	0.07
HN9	0.49	0.50	0.39	0.28	1.18	0.02	0.21	0.11	0.45	1.08	0.07
HN10	0.55	0.60	0.54	0.33	1.25	0.01	0.30	0.14	0.50	1.82	0.07
HN11	0.23	0.23	0.20	0.23	1.14	0.03	0.17	0.09	0.61	0.95	0.05
HN12	0.34	0.32	0.22	0.19	0.98	0.03	0.12	0.07	0.23	0.48	0.04
HN13	0.44	0.53	0.39	0.36	1.25	0.03	0.25	0.15	0.38	1.01	0.09
HN14	0.22	0.21	0.18	0.20	0.93	0.02	0.14	0.08	0.38	0.68	0.03
HN15	0.24	0.22	0.14	0.33	1.79	0.00	0.46	0.18	1.38	3.04	0.10
HN16	0.62	0.60	0.50	0.31	0.93	0.01	0.40	0.16	0.71	1.55	0.10
HN17	0.15	0.12	0.06	0.23	1.41	0.00	0.37	0.13	0.97	2.29	0.07
HN18	0.77	0.79	0.56	0.42	1.46	0.04	0.29	0.15	0.59	1.36	0.08
HN19	0.27	0.22	0.11	0.28	1.57	0.01	0.44	0.15	1.14	2.51	0.08
HN20	0.70	0.68	0.42	0.42	0.56	0.08	0.30	0.13	0.53	0.85	0.12
HN21	0.09	0.12	0.07	0.25	1.37	0.00	0.34	0.11	0.91	2.35	0.08
HN22	0.40	0.25	0.20	0.22	1.34	0.01	0.36	0.12	1.01	2.09	0.07
HN23	0.80	0.80	0.62	0.33	0.99	0.02	0.36	0.17	0.43	1.24	0.10
HN24	0.42	0.39	0.23	0.28	1.57	0.01	0.44	0.15	1.18	2.51	0.09
HN25	0.29	0.22	0.14	0.25	1.33	0.00	0.38	0.13	1.05	2.25	0.07
HN26	0.26	0.18	0.13	0.30	1.59	0.01	0.45	0.17	1.16	2.69	0.07

HN1 exhibits reduced biodistribution to peripheral tissues and an accelerated clearance from systemic circulation

Although the scAAV9-hSYN1-GFP-BCLib was designed to focus on neuronal transduction efficiency using hSYN1-driven mRNA barcodes, it also facilitates pharmacokinetic characterization of capsid variants via analysis of DNA barcodes. We therefore evaluated the relative biodistribution of capsid variants to major organs. As expected, given that our library consisted of peptide insertions on the AAV9-N272A capsid, we observed consistent detargeting of the liver (Table 3.4, Table 3.5). In C57BL/6J mice, variants that exhibited enhanced brain transduction (i.e., HN1 and HN2) appeared to have noticeably reduced liver detargeting relative to the parental AAV9-N272A variant. However, this effect was less obvious in BALB/cJ mice (Table 3.4, Table 3.5). In addition to detargeting the liver, HN1 demonstrated a moderate level of detargeting for most other peripheral tissues, with the exception of the lung (Figure 3.6a). Interestingly, HN1 was relatively less accumulated in the spleen compared to many of the other capsid variants.

We performed an additional experiment to evaluate the relative blood clearance rate of the TRADE-identified AAV capsid variants. C57BL/6J mice were delivered a bolus dose (1×10^{13} vg/kg) of the scAAV-hSYN1-GFP-BCLib via tail vein and blood was collected after 1 minute, 10 minutes, 30 minutes, 1 hour, 4 hours, 8 hours, 24 hours, and 72 hours. AAV DNA Barcode-Seq analysis of vector genomes extracted from blood revealed that many of the variants, including AAV-PHP.B, achieved similar or higher blood concentrations than AAV9 up to 24 hours, but then relative blood concentrations rapidly fell by 72 hours compared to AAV9 (Figure 3.6b, Table 3.6). Notably, however,

the HN1 and HN2 variants exhibited a blood clearance profile that was consistently below that of AAV9, even at early timepoints.
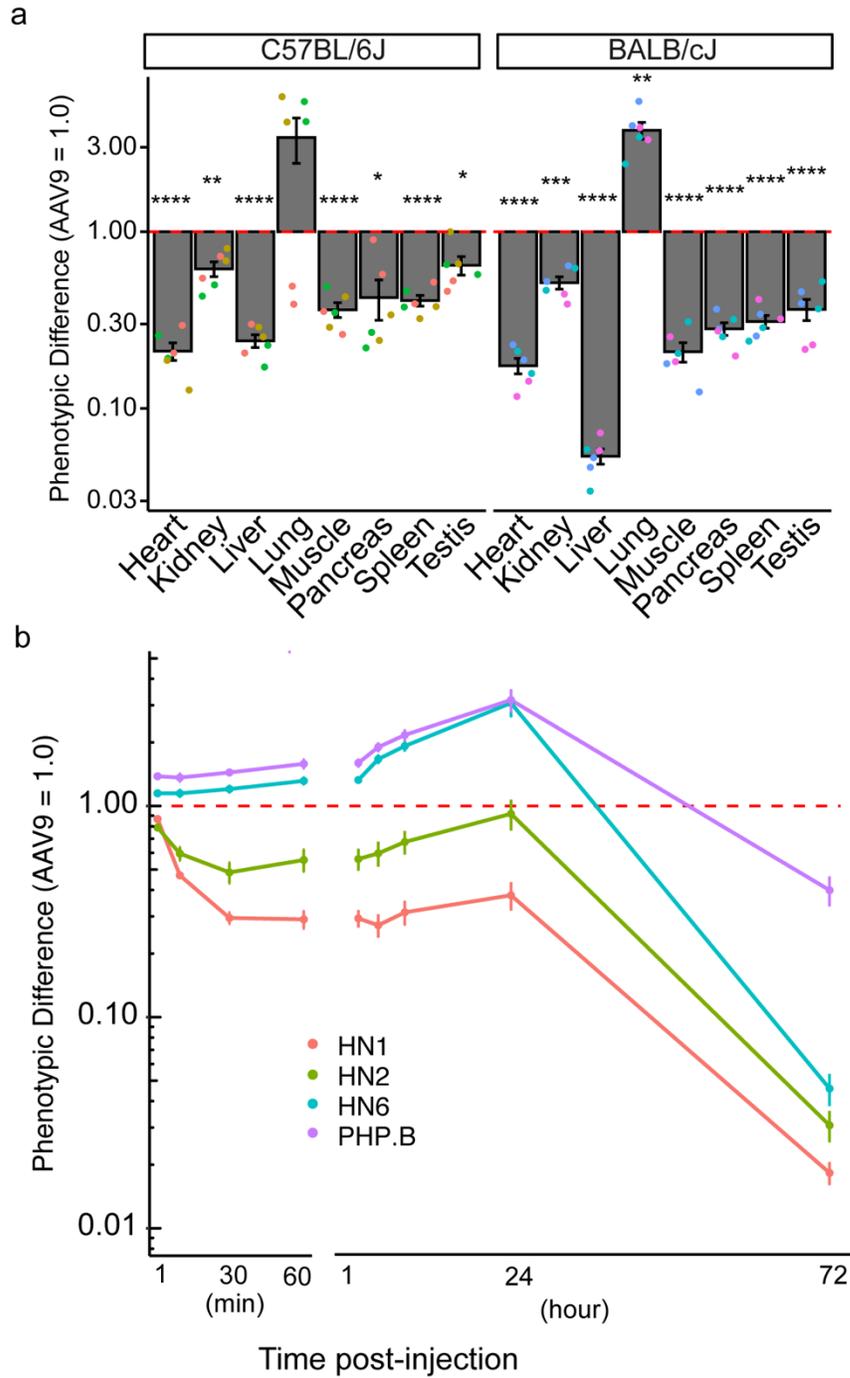
Figure 3.4 AAV DNA Barcode-seq demonstrates moderate detargeting of most non-CNS tissues and accelerated clearance

(a) Biodistribution of the HN1 capsid variant relative to AAV9 in non-CNS tissues as determined by AAV DNA Barcode-Seq using the same animals seen in Figure 3.3. (b) Concentration of AAV capsid variants in blood over time (relative to AAV9) following bolus dosing of 1×10^{13} vg/kg intravenously. Mean ± SEM are plotted across two different time scales. Welch's t-test were used to compare PD values within a given

mouse strain and tissue sample. P-values were corrected using the false discovery rate (FDR) method and compared to an alpha value of 0.05. *p < 0.05, **p < 0.01, ****p < 0.0001.

**Table 3.6 Relative blood clearance in C57BL/6J mice**

	1 min	10 min	30 min	1 hr	4 hr	8 hr	24 hr	72 hr
AAV9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AAV9-N272A	1.16	1.17	1.28	1.34	1.77	2.02	3.14	0.13
PHP.B	1.38	1.42	1.44	1.62	1.88	2.15	3.14	0.39
HN1	0.87	0.47	0.29	0.29	0.27	0.31	0.37	0.02
HN2	0.79	0.60	0.49	0.55	0.59	0.66	0.91	0.03
HN3	0.99	1.01	1.05	1.16	1.45	1.58	2.38	0.03
HN4	1.29	1.40	1.44	1.49	1.93	2.49	3.74	0.03
HN5	1.21	1.19	1.29	1.30	1.74	2.03	3.09	0.03
HN6	1.12	1.15	1.21	1.32	1.65	1.88	3.05	0.04
HN7	0.92	0.88	0.93	0.93	1.09	1.21	1.22	0.03
HN8	1.20	1.29	1.30	1.44	1.83	2.10	3.42	0.03
HN9	0.94	0.91	0.97	0.99	1.26	1.36	1.97	0.04
HN10	0.98	0.98	1.00	1.06	1.31	1.36	1.89	0.02
HN11	1.00	1.04	1.04	1.14	1.39	1.48	2.11	0.02
HN12	0.93	0.93	0.84	0.79	0.71	0.61	0.62	0.01
HN13	0.95	0.90	0.95	0.95	1.20	1.28	1.60	0.03
HN14	0.94	0.95	1.00	1.08	1.41	1.58	2.56	0.01
HN15	1.39	1.56	1.67	1.66	2.20	2.76	4.26	0.03
HN16	0.98	1.00	1.04	1.15	1.46	1.63	2.77	0.02
HN17	1.32	1.28	1.27	1.31	1.94	2.13	4.03	0.04
HN18	1.10	1.06	0.96	0.93	0.82	0.84	1.27	0.01
HN19	1.39	1.39	1.51	1.49	2.04	2.50	4.09	0.03
HN20	1.15	1.09	1.19	1.14	1.41	1.70	1.97	0.06
HN21	1.25	1.19	1.34	1.38	1.90	1.99	3.30	0.02
HN22	1.16	1.24	1.32	1.35	1.74	2.13	3.74	0.02
HN23	1.03	1.02	1.04	1.14	1.43	1.64	2.56	0.03
HN24	1.32	1.50	1.43	1.52	2.01	2.37	3.78	0.03
HN25	1.14	1.30	1.37	1.42	1.84	2.03	3.37	0.02
HN26	1.21	1.34	1.43	1.60	1.99	2.38	3.90	0.02

Validation of HN1 enhanced neuronal transduction efficiency

In order to cross-validate the enhanced neuronal transduction efficiency of the HN1 variant using a more traditional method of quantifying transduction efficiency, we generated a self-complementary hSYN1-driven GFP reporter vector (Figure 3.5a) and packaged it into the AAV9, HN1, and AAV-PHP.B capsids. These vectors were intravenously administered into separate cohorts of mice at a dose of 3 x 10^{11} vg (n = 4 mice / vector for C57BL/6J, n = 3 mice / vector for BALB/cJ). Brains were harvested three weeks post-injection for cryosectioning and transduction efficiency was quantified by immunofluorescence (Figure 3.5).

As expected given previous reports, AAV-PHP.B exhibited significantly enhanced neuronal transduction across the brain in C57BL/6J, but not in BALB/cJ mice (Hordeaux et al. 2018; Q. Huang et al. 2019). HN1 exhibited enhanced brain neuronal transduction in multiple brain regions of both mouse strains compared to AAV9. HN1 was 9.5-fold more efficient than AAV9 in the C57BL/6J motor cortex (one-way ANOVA $F(2,8) = 12.8$, $p < 0.01$; Tukey HSD $p < 0.01$) with a similar degree of enhancement observed in the visual cortex, thalamus, and striatum. In contrast to AAV-PHP.B, HN1 exhibited a 9.0-fold enhancement in the motor cortex of BALB/cJ mice (one-way ANOVA $F(2,6) = 181.1$, $p < 0.0001$; Tukey HSD $p < 0.0001$) with significant enhancement also observed in the other brain regions assayed. Notably, relative enhancement in neuronal transduction efficiency as determined by AAV RNA BC-Seq and the traditional immunofluorescence methods were highly consistent.
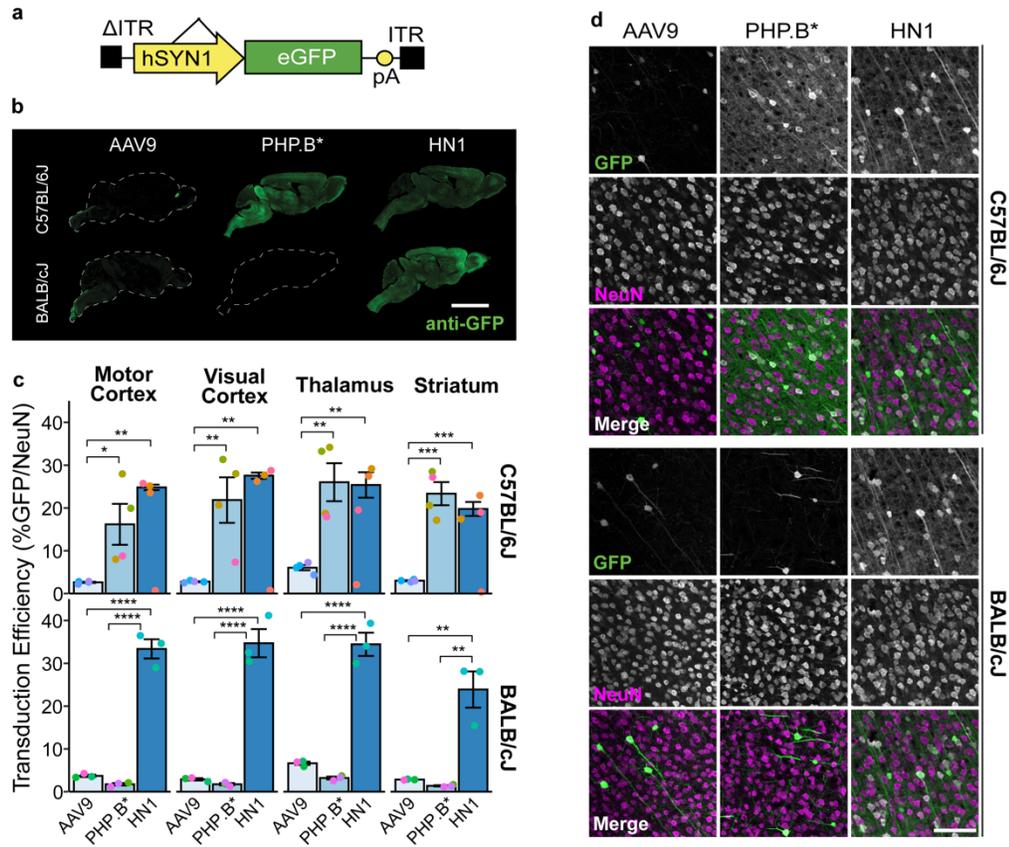
Figure 3.5 Validation of HN1's enhanced of neuronal transduction efficiency in mouse

(a) Map of a self-complementary AAV vector genome expressing a reporter gene (GFP) under control of the neuron-specific hSYN1 promoter-enhancer. (b) Tilescanned sagittal brain sections stained with an anti-GFP antibody. Scale bar = 5 mm. (c) Quantification of neuronal transduction efficiency (percent of NeuN+ cells that were also GFP+) in various brain regions of two mouse strains. Means ± SEM are plotted. Colored datapoints represent individual animals. Tukey's HSD post-hoc, *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. Note that one animal in the HN1-administered C57BL/6J cohort was designated an outlier based on Q1-1.5xIQR criteria. The datapoint is shown, but was not included in any calculations. (d) Representative high-power confocal images from each quantified brain region. Scale bar = 100 μm. PHP.B*: note that AAV-PHP.B transduction efficiency is underestimated (see below).

HN1 is highly tropic for brain neurons

In order to evaluate the tropism of the HN1 variant, we packaged a ubiquitously expressing, CAG promoter-driven, GFP reporter (Figure 3.6a) into AAV9 and HN1 capsids and administered these vectors into C57BL/6J mice at a dose of 3×10^{11} vg intravenously (n = 4 mice / vector). Brains were harvested three weeks after injection, cryosectioned, and counterstained with NeuN (marking neurons), S100 (marking a predominantly astrocytic glial population), Olig2 (marking cells in the oligodendrocyte lineage), and Iba1 (marking microglia) (Figure 3.6b-e). Notably, we observed no co-localization of GFP with Olig2+ cells or Iba1+ cells at the dose delivered.

We quantified AAV tropism as the proportion of GFP$^+$ cells that were also positive for a cell type-specific marker, and compared tropism for NeuN+ neurons to tropism for S100+ glia (Figure 3.6f). HN1 transduction was strongly biased toward neurons compared to glia across the brain. In the motor cortex, HN1 exhibited a strong neuronal tropism of 85.9% compared to a glial tropism of only 10.7% (Welch's t-test, $p < 0.0001$). In stark contrast, AAV9 demonstrated a significant bias for transduction of glia (glial tropism of 66.1% versus neuronal tropism of 26.9%, $p < 0.05$). In the thalamus, both AAV9 and HN1 exhibited a strong preference for transducing neurons with neuronal tropisms of 88.5% and 98.7%, respectively. In striatum, HN1 again exhibited robust neuronal tropism (98.1%), but AAV9 showed no significant bias towards neurons or glia.

Although it was not the primary goal of the experiment, we took advantage of this dataset to measure relative transduction efficiency by calculating the average number of transduced neurons (GFP+/NeuN+) and transduced glia (GFP+/S100+). In this context, HN1 exhibited 36-fold greater transduction than AAV9 in motor cortex (Welch's t-test, p

< 0.01), 10-fold greater in the thalamus (p < 0.01), and 13-fold greater in the striatum (p

< 0.0001). HN1-mediated glial transduction was similar to that of AAV9 in the cortex,

but significantly reduced in thalamus and striatum (p < 0.05 for each region).

125
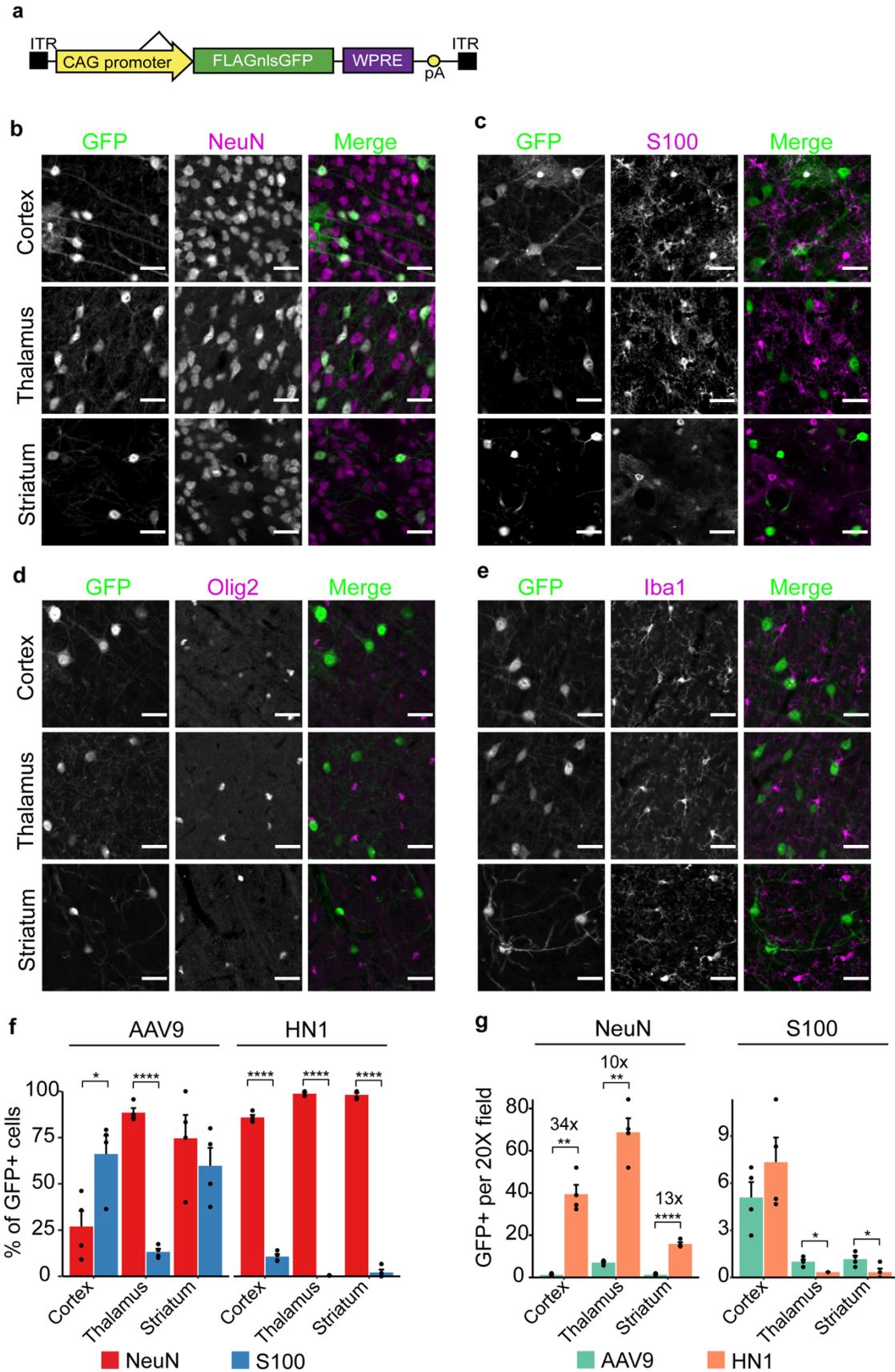
Figure 3.6 Characterization of HN1 tropism and transduction efficiency using a ubiquitously expressing single strand AAV genome

(a) Map of a self-complementary AAV genome driving expression of GFP under the control of a strong, ubiquitously expressing CAG promoter-enhancer sequence. (b-e) Representative optically sectioned images showing HN1-mediated transduction in brain sections counterstained with cell type specific markers. Note that no transduction of Olig2+ or Iba1+ cells was observed. Scale bar = 30 µm (f) Comparison of AAV9 and HN1 tropism for neurons (NeuN+) versus tropism for glia (S100+) based on the proportion of all transduced (GFP+) cells that were also positive for each cell type marker. Cell types were assayed independently. (g) Secondary analysis comparing neuronal and glial transduction efficiency, defined as the number of GFP+ cells in a 20X field. In f-g, means +/ SEM are plotted. Datapoints represent the average of 3 technical replicates for each animal. Welch's t-tests with Holm-Bonferroni correction for multiple comparisons. * p < 0.05, **p < 0.01, **** p < 0.0001.

Deep sequencing and enrichment analysis re-identifies the HN1 variant

Our initial selection of variants for validation studies was based on the most common variants identified by Sanger sequencing of 56 individual clones after three rounds of selection. However, recent advances have highlighted the power of applying next-generation sequencing methods to monitor variant evolution over each round of selection using an enrichment analysis (Kumar et al. 2020; Nonnenmacher et al. 2020). We therefore performed post-hoc deep sequencing for each round of in vivo biopanning.

We then applied an enrichment analysis, which calculates the proportion of reads in RNA extracted from tissue normalized to the proportion of reads in the AAV input stock library, to the five mouse-derived variants that were selected following three rounds of hSYN1-driven transcription-dependent selection (Table 3.7). Surprisingly, this analysis ranked HN1 as only the 34th most enriched capsid variant. HN3 and HN4, which demonstrated no enhancement based on our AAV RNA Barcode-Seq experiment (Table 3.4), actually achieved higher enrichment scores than HN1. This suggested a high level of noise in our library and we therefore sought to re-equilibrate our library and perform one additional round of in vivo biopanning.

We utilized a lenient cutoff to identify variants that were enriched after two or three rounds of selection (see Methods, Figure 3.7a). Selected variants were then grouped using a top-down algorithm to identify peptide sequences that differed by only a single amino acid. Notably, the dominant peptide sequence within each group typically had a read count several orders of magnitude greater than all of the other peptide sequences within the group (Figure 3.7b), suggesting that the less common peptide sequences were artifactual (e.g., due to PCR error, Illumina sequencing read error, etc.). We therefore,

selected only the most common peptide sequence within each group, resulting in a list of 416 unique sequences.

Each peptide was synthesized as five synonymous DNA sequences and flanked with homology arms to facilitate DNA assembly into the pAAV9-N272A-2BsaI-hSYN1-TRADE vector genome. A plasmid library and corresponding AAV library were generated and administered into C57BL/6J mice (n = 4). Brain tissue was harvested twelve days post-injection. Peptide sequences were recovered by RT-PCR and submitted for deep sequencing. For 98% of variants included in the library, we were able to recover all five synonymous DNA sequences. In the remaining cases, only four of the five DNA sequences were recovered, with the exception of just one capsid variant for which we only recovered three of the five DNA sequences.

We repeated the enrichment analysis and focused first on the HN1-5 variants. In contrast to the previous enrichment analysis, we now observed a pattern of enrichment consistent with the scAAV-hSYN1-GFP-BCLib results; HN1 and HN2 demonstrated enrichment in the re-equilibrated library whereas HN3, HN4, and HN5 were not enriched, or even selected against (Figure 3.7c).

We then attempted to identify novel capsid variants that went undetected by the previous enrichment analysis. In the absence of any reference control within this library, we attempted to identify variants that demonstrated a significant level of enrichment, taking into account variability due to different synonymous DNA sequences and biological replicates. We defined a null hypothesis that a given variant failed to demonstrate enrichment (i.e., the enrichment score was less than or equal to 1) and performed one-tailed Wilcoxon texts, adjusting p-values with the False Discover Rate

method (Figure 3.7d). This stringent analysis re-identified both the HN1 and HN2 capsid

variants. In addition, two new variants were identified: QQNGTRTP and GSNHTQSL.

Notably, the QQNGTR<u>TP</u> peptide is remarkably similar to the HN2 variant

(QQNGTR<u>PS</u>), with a Hamming distance of 2, but a common proline in the mismatch

region.

Table 3.7 Enrichment scores and ranking for mouse-derived capsid variants following three rounds of transcription-dependent selection

Variant	Enrichment Score	Enrichment Rank
HN1	25941.04	34
HN2	315794.7	1
HN3	114176.4	8
HN4	57373.74	18
HN5	14157.71	55
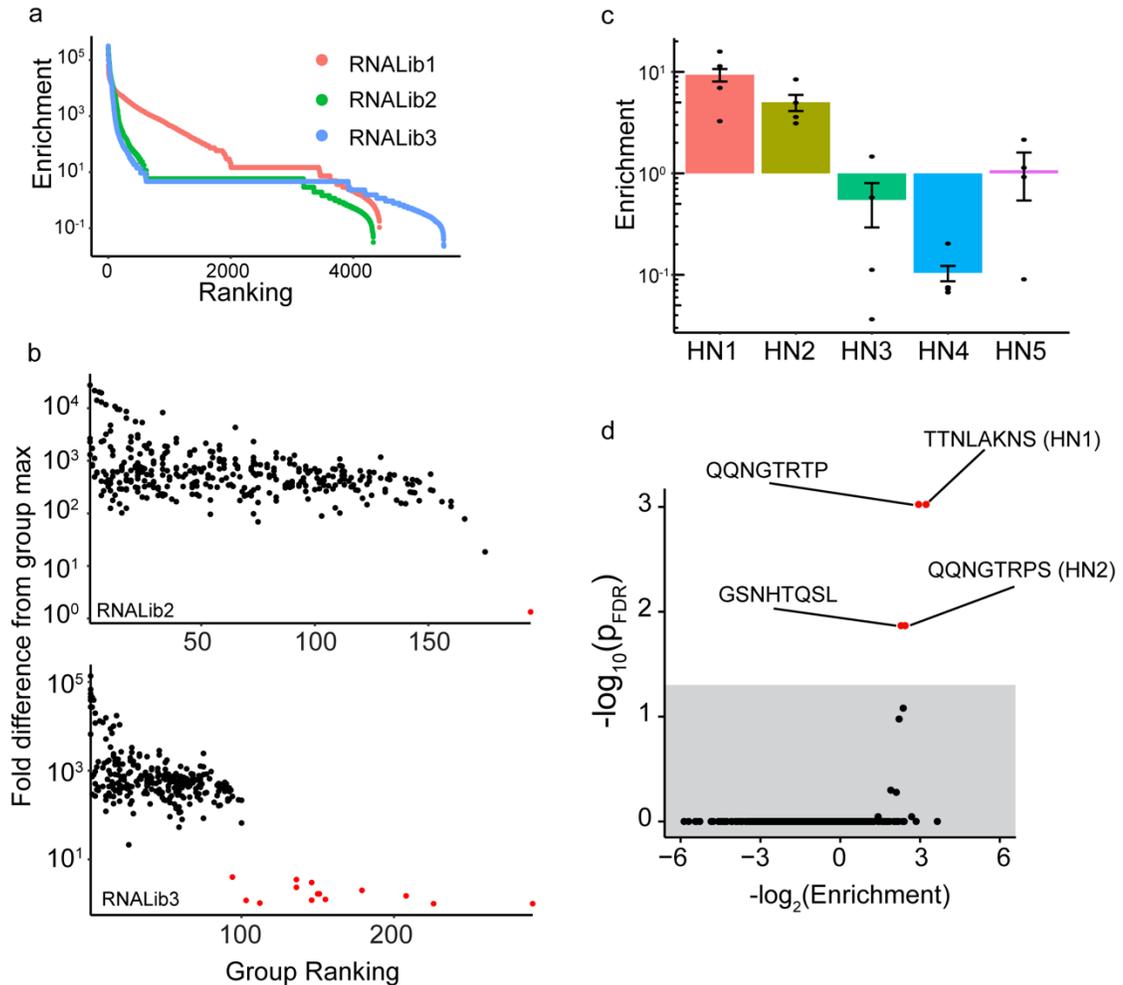
Figure 3.7 Re-identification of HN1 following library equilibration and enrichment score analysis

(a-b) Variant peptide insertion sequences were recovered by RT-PCR and deep sequenced in order to calculate enrichment scores. For each peptide sequence, the proportion of reads in RNA (for each round of selection) was divided by the proportion of reads in the AAV stock library (AAVLib0). (a) Capsid variant enrichment scores in ranked order following each round of selection. (b) Peptide sequences with a Hamming distance of 1 were grouped (top-down collapsing). Subsequently, the maximum read count in each group was divided by the read count for each peptide in the group. These fold difference values are plotted for each group, ranked by the highest enrichment score within each group. Values of 1 are not shown. Points in red highlight peptide sequences with read counts less than 10-fold different. (c-d) We used the TRADE system to biopan selected variants in C57BL/6J mice (n = 4). (c) Mean enrichment scores with SEM are shown for the five previously selected mouse-derived variants. Each point represents an enrichment score calculated from a single animal. (d) Modified volcano plot showing the

results of one-sample Wilcoxon test against the mean enrichment score. P-values are corrected using the False Discovery Rate method. Grey background highlights failure to achieve statistical significance with alpha = 0.05.

Underestimation of AAV-PHP.B transduction efficiency

Although our assays of neuronal transduction efficiency suggested that the HN1 variant was comparable to AAV-PHP.B, we were suspicious about the modest enhancement observed for the AAV-PHP.B vector, previously reported to be >40-fold (Deverman et al. 2016). We also noted a relatively lower number of reads associated with the AAV-PHP.B barcode clones in the scAAV-hSYN1-GFP-BCLib. In order to investigate if this was the result of purification bias or simply technical error at the stage of mixing crude lysates prior to purification (see Methods), we quantified the relative abundance of each AAV variant in the library before and after purification. Prior to purification, AAV variants were close to the expected equimolar ratio (Figure 3.8b, x-axis), but the relative abundance of barcoded vectors varied by more than 10-fold after bulk purification (Figure 3.8b, y-axis). We next calculated relative recovery scores by dividing the relative abundance in the post-purification stock by the relative abundance in the pre-purification crude lysate (Figure 3.8c). If bulk purification is an unbiased process and all AAV variants are recovered proportionately, then the relative recovery scores should all be similar to AAV9 (i.e., equal to 1). Indeed, we observed a number of variants with relative recovery scores similar to that of AAV9. However, we also observed that many of the variants, including AAV-PHP.B, had relative recovery scores far below that of AAV9, even more than 10-fold.

In order to objectively categorize variants in the library, we utilized unsupervised k-means clustering. We selected the number of clusters (k) based on visualization of within-group sum of squares as a function of the number of groups and determined, as expected, that two clusters optimally reduced variance without overfitting (Figure 3.8a).

We then performed k-means clustering to categorize the variants into two groups that corresponded with high or low recovery following purification. Since the pre-purification relative abundance scores were similar, we hypothesized that differences in relative abundance after purification were due to loss by precipitation and that this loss might be mediated by increased hydrophobicity.

Therefore, we compared the proportion of hydrophobic amino acids (ALIMFWV) in the 8mer peptide insertions between the High and Low recovery groups (Figure 3.8d-e). Intriguingly, we found the proportion of hydrophobic amino acids was significantly higher in the low recovery group compared to the high recovery group (42% vs 24%, two-proportion z-test, $p < 0.01$). In addition, we observed that 8 out of 12 variants in the low recovery group contained a hydrophobic amino acid in position 2 of the peptide insertion while only 1 out of 14 variants in the high recovery group had a hydrophobic amino acid at the same position (no statistical test attempted). These data strongly suggest that AAV variants with hydrophobic peptide insertions are precipitated and lost during bulk purification.

Given that the unexpectedly low relative neuronal transduction efficiency exhibited by our AAV-PHP.B vectors could be due to purification-related loss of vector activity, we sought to re-evaluate transduction efficiencies using a different purification method. We hypothesized that AAV-PHP.B might be prone to precipitation unless a high salt buffer was used during the purification, as originally described (Deverman et al. 2016; Challis et al. 2019). We therefore repeated the experiment as before, but utilized a high salt buffer to resuspend the PEG pellet (see Methods).

Neuronal transduction efficiencies for AAV9 and HN1 were similar regardless of the resuspension buffer used during vector purification. However, AAV-PHP.B transduction efficiency was vastly enhanced. Due to the use of a cytoplasmic GFP reporter, we observed a strong signal in the neuropil that obscured our ability to count GFP+ cell bodies (Figure 3.9a). Therefore, we measured the concentration of GFP in mouse brain cortex using ELISA.

Consistent with the previous experiments, HN1 exhibited a 9.34-fold enhancement in neuronal transduction efficiency compared to AAV9 in C57BL/6J mice (Welch's ANOVA $F(2,4.02) = 54.6$, $p < 0.01$; Welch's t-test, $p < 0.01$) and 7.5-fold in BALB/cJ mice (Welch's ANOVA $F(2,5.33) = 13.49$, $p < 0.01$; Welch's t-test, $p < 0.05$). As expected, AAV-PHP.B did not demonstrate enhanced transduction in the BALB/cJ brain. Surprisingly, however, AAV-PHP.B mediated a neuronal transduction efficiency that was 288-fold greater than that of AAV9 (Welch's t-test, $p < 0.01$) and 30-fold greater than that of HN1 (Welch's t-test, $p < 0.01$) (Figure 3.9b).
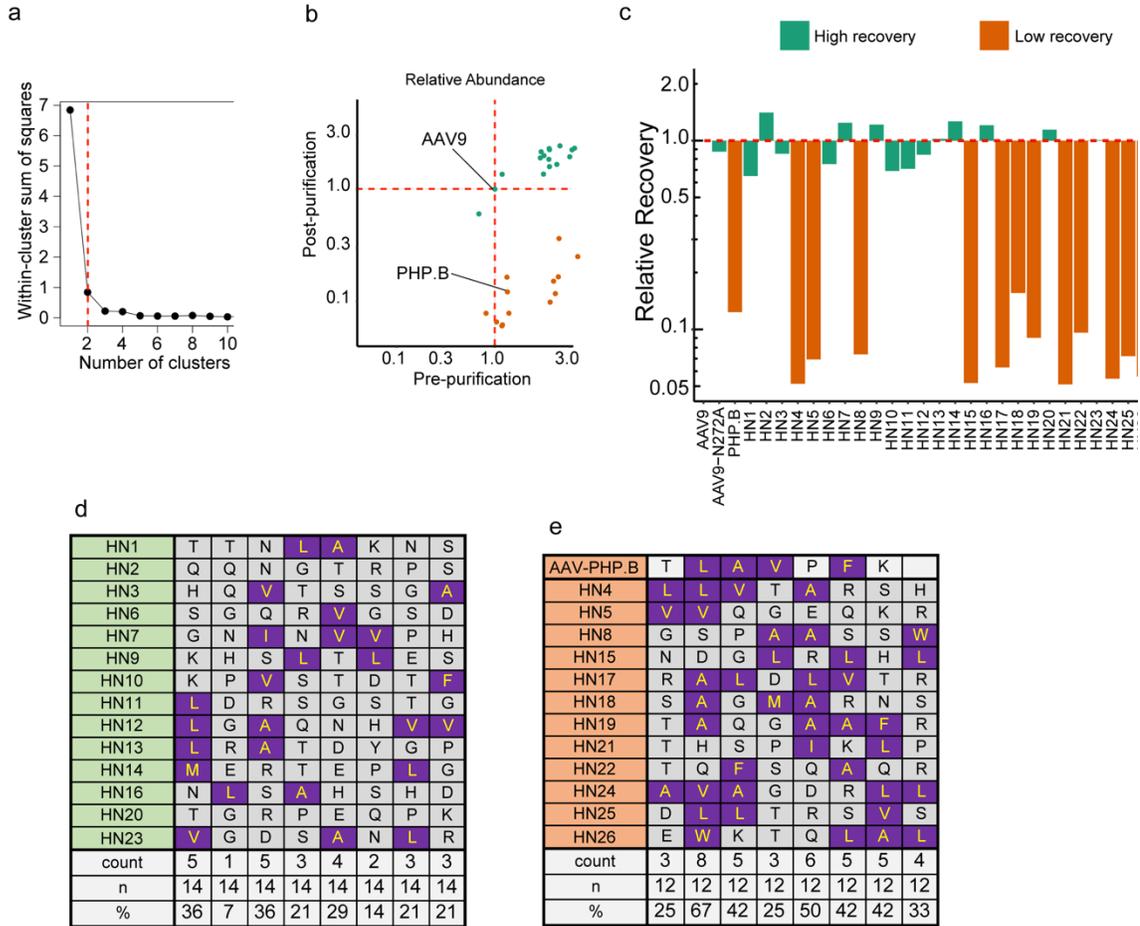
Figure 3.8 AAV-PHP.B and other variants are disproportionately lost during purification of the pooled scAAV-hSYN1-GFP-BCLib
(a) Elbow plot used to determine the optimal number of clusters for k-means clustering analysis. Dashed red line indicates selection of k = 2. Resulting clustering identified two groups, labeled "High recovery" (green) and "Low recovery" (orange). (b) Relative abundance of each AAV variant in the pre-purification mixture of crude lysate and in the final purified vector prep normalized to AAV9 calculated based on normalization of next-generation sequencing read counts. (c) Relative recovery of each variant following purification, calculated as the ratio of the relative abundance in the post-purification stock library to the relative abundance in the pre-purification crude lysate mixture. A value of 1 means that recovery was proportional to that of AAV9. (d-e) Analysis of hydrophobicity of variant peptide insertions. Amino acids with hydrophobic functional groups (ALIMFWV) are highlighted in purple and separated into the High recovery (d) and Low recovery (e) groups. Frequencies of finding a hydrophobic residue at each position of the insertion are tabulated below. In (e), AAV-PHP.B was not included in the counts.
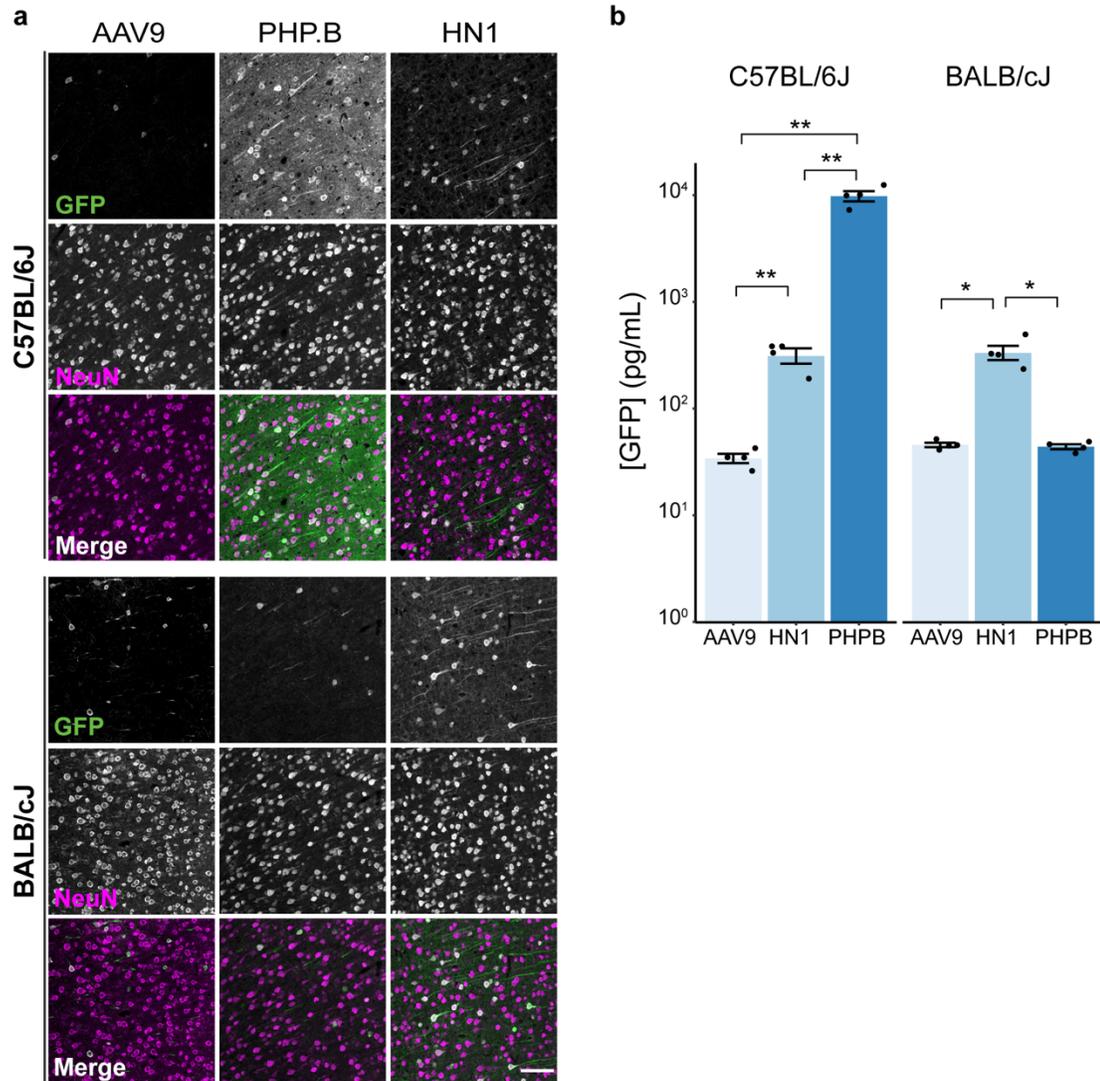
Figure 3.9 AAV-PHP.B transduces mouse brain neurons more efficiently than HN1 following intravenous administration.
The experiment in Figure 3.5 was repeated, but vectors were purified using a high salt resuspension buffer. (a) Transduction efficiency of AAV9 and HN1 vectors appear similar to those observed in the previous experiment. However, AAV-PHP.B shows a clear enhancement. Due to the high transduction efficiency, we could not confidently quantify transduction efficiency in PHP.B-administered C57BL/6J mice by immunofluorescence. Note the high level of extranuclear GFP signal. Scale bar = 100 µm. (b) We therefore quantified protein expression at the protein level using a GFP ELISA kit. Welch's ANOVA with Welch's t-test and Holm-Bonferroni correction. *$p < 0.05$, **$p < 0.01$.

AAV Barcode-Seq identifies HN1 as a potential candidate for systemically-administered CNS gene therapy

Notwithstanding the clear superiority of AAV-PHP.B in C57BL/6J mice, it is now apparent that AAV-PHP.B does not confer enhancement in the non-human primate context (Hordeaux et al. 2018; Matsuzaki et al. 2018; Liguore et al. 2019). We therefore sought to evaluate the HN1 variant in rhesus and took advantage of our scAAV-hSYN1-GFP-BCLib. We thus administered the library into a single male rhesus macaque at a dose of 2 x 10^{13} vg/kg and harvested tissues two weeks post-injection. We performed AAV RNA Barcode-Seq using tissue dissected from a variety of brain regions. Consistent with previous reports, AAV-PHP.B showed no obvious enhancement compared to AAV9 (Table 3.8). HN1 demonstrated enhanced neuronal transduction efficiency > 2-fold in several brain regions (Figure 3.10, Table 3.8). HN2 showed no evidence of enhancement in non-human primate. In contrast to the universal underperformance of rhesus-derived variants in rodents (Table 3.4, Table 3.5), some of the rhesus-derived variants, such as HN6, performed as well as AAV9 or even slightly better (Table 3.8).

AAV DNA Barcode-Seq analysis revealed similar trends in capsid variant tissue biodistribution to those observed in rodents. As expected, the liver was effectively detargeted by variants with the AAV9-N272A backbone, but the effect appeared to be attenuated for HN1 (Table 3.9). Nevertheless, HN1 exhibited a similar degree of detargeting peripheral tissues as was observed in rodents, including a relative lack of accumulation in the spleen (Figure 3.11a).

Blood was collected at various timepoints following infusion in order to evaluate the relative blood clearance rate of the TRADE-identified AAV capsid variants. HN1 exhibited a trend toward more rapid blood clearance than AAV9, though the effect was not as robust as seen in C57BL/6J mice (Figure 3.6c, Table 3.10).

In addition to analyzing the non-human primate samples, we also evaluated biodistribution to human hepatocytes using a chimeric model of mice with humanized livers (Azuma et al. 2007). We administered the scAAV-hSYN1-GFP-BCLib via tail vein at a dose of 5 x 10^{11} vg (n = 2 mice). Two weeks later, human hepatocytes were isolated by collagenase perfusion and flow cytometry. Intriguingly, AAV DNA Barcode-Seq demonstrated a loss of relative detargeting for AAV-PHP.B and HN2 compared to detargeting of rodent and NHP livers. However, HN1 retained a similar level of liver detargeting in human hepatocytes (Figure 3.6b).
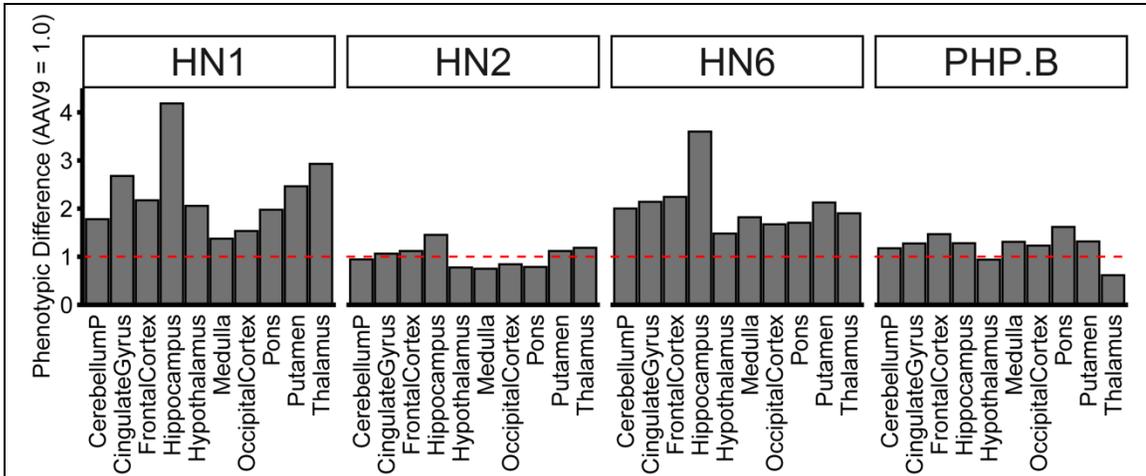
Figure 3.10 hSYN1-GFP-BC-seq demonstrates a moderate enhancement in NHP brain
The scAAV-hSYN1-GFP-BCLib was administered to a single male rhesus macaque at a dose of 2×10^{13} vg/kg. AAV RNA Barcode-Seq determined phenotypic differences, representing relative neuronal transduction efficiencies, in various brain regions

Table 3.8 AAV RNA Barcode-Seq in rhesus brain

	Cerebellum	Cingulate Gyrus	Frontal Cortex	Hippocampus	Hypothalamus	Medulla	Occipital Cortex	Pons	Putamen	Thalamus
AAV9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AAV9-N272A	1.63	1.76	1.75	2.08	1.00	1.45	1.51	1.45	1.99	1.19
PHP.B	1.21	1.27	1.50	1.39	0.94	1.33	1.23	1.62	1.29	0.69
HN1	1.78	2.68	2.18	4.07	2.05	1.39	1.54	1.99	2.54	2.84
HN2	0.96	1.06	1.13	1.45	0.78	0.77	0.84	0.79	1.15	1.16
HN3	1.18	1.19	1.33	1.71	0.70	1.17	1.06	1.04	1.39	1.21
HN4	0.14	0.06	0.32	0.21	0.12	0.23	0.24	0.07	0.28	0.08
HN5	0.52	0.42	0.60	0.88	0.40	0.50	0.44	0.36	0.61	0.28
HN6	2.03	2.14	2.28	3.54	1.48	1.87	1.67	1.71	2.20	1.87
HN7	1.40	1.44	1.45	1.64	1.13	1.48	1.18	1.01	1.39	1.29
HN8	0.61	0.59	0.67	0.73	0.42	0.56	0.42	0.45	0.57	0.50
HN9	1.45	1.91	1.66	2.15	0.92	1.65	1.30	1.27	1.96	1.48
HN10	1.24	1.40	1.45	1.78	0.87	1.33	1.16	1.11	1.50	1.00
HN11	1.22	1.37	1.42	1.62	1.00	1.28	1.15	1.12	1.41	1.19
HN12	1.43	1.64	1.70	1.98	0.97	1.49	1.26	1.09	1.74	1.44
HN13	1.74	1.86	1.82	2.17	1.20	2.38	1.54	1.97	2.14	1.77
HN14	0.45	0.27	0.26	0.28	0.14	0.30	0.32	0.19	0.23	0.17
HN15	0.43	0.19	0.46	0.44	0.04	0.28	0.19	0.23	0.36	0.09
HN16	0.65	0.79	0.82	0.89	0.35	0.75	0.77	0.64	0.72	0.54
HN17	0.18	0.24	0.14	0.08	0.28	0.16	0.11	0.07	0.19	0.03
HN18	1.23	1.62	1.70	2.60	1.17	1.47	1.15	1.09	1.46	1.13
HN19	0.24	0.21	0.59	0.50	0.13	0.14	0.26	0.26	0.31	0.17
HN20	1.42	1.60	1.81	2.28	1.45	1.51	1.22	1.41	2.27	1.34
HN21	0.11	0.05	0.15	0.04	0.49	0.27	0.19	0.08	0.10	0.14
HN22	0.17	0.48	0.59	0.49	0.24	0.24	0.27	0.19	0.23	0.12
HN23	1.01	1.21	1.11	1.55	0.61	1.23	0.92	0.95	1.25	0.76
HN24	0.61	0.70	0.68	0.37	0.39	0.58	0.54	0.27	0.49	0.62
HN25	0.27	0.35	0.22	0.15	0.01	0.27	0.17	0.27	0.31	0.15
HN26	0.28	0.36	0.12	0.10	0.27	0.04	0.12	0.12	0.26	0.10
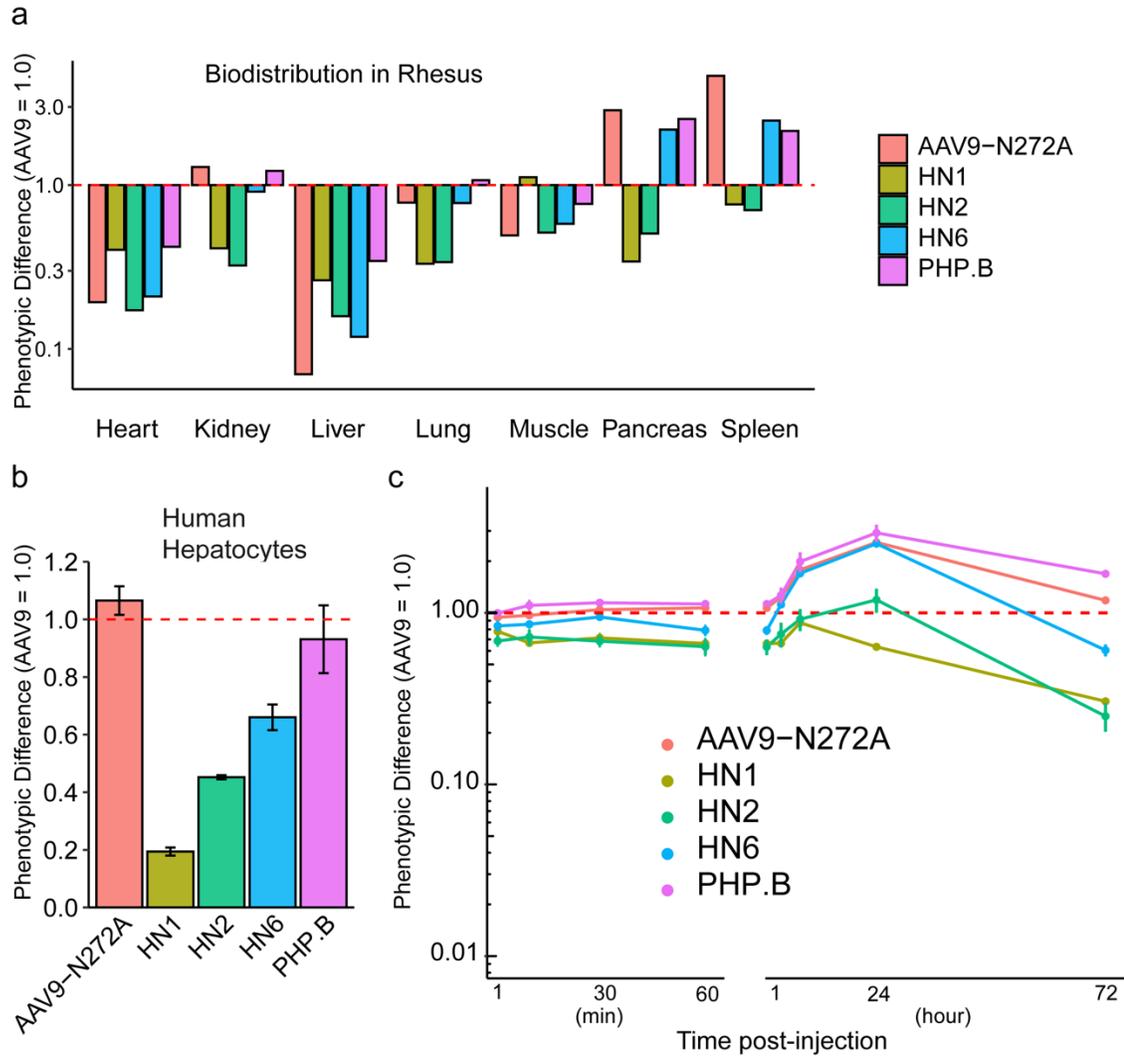
Figure 3.11 Biodistribution and pharmacokinetic profile of the TRADE-identified capsids

(a) Biodistribution of the HN1 capsid variant relative to AAV9 in non-CNS tissues as determined by AAV DNA Barcode-Seq (n = 1). (b) Relative biodistribution to human hepatocytes following intravenous administration to chimeric mice harboring humanized livers (n = 2). (c) Concentration of AAV capsid variants in blood over time, relative to AAV9 (n = 1).

Table 3.9 AAV DNA Barcode-Seq in non-CNS rhesus tissues

	Liver	Heart	Kidney	Lung	Muscle	Pancreas	Small Intestine	Stomach	Spleen	Bone Marrow
AAV9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AAV9-N272A	0.07	0.20	1.29	0.79	0.26	2.91	0.64	0.82	4.67	0.06
PHP.B	0.35	0.42	1.22	1.09	0.54	2.55	0.40	0.80	2.14	0.43
HN1	0.26	0.41	0.41	0.34	1.04	0.35	0.24	0.33	0.76	0.33
HN2	0.16	0.17	0.32	0.35	0.32	0.51	0.11	0.30	0.71	0.07
HN3	0.06	0.14	0.74	0.66	0.18	1.99	0.13	0.50	1.98	0.03
HN4	0.06	0.07	1.21	0.90	0.22	2.76	0.10	0.56	3.28	0.03
HN5	0.17	0.10	0.90	0.64	0.12	1.84	0.12	0.56	2.38	0.02
HN6	0.12	0.21	0.91	0.78	0.25	2.19	0.18	0.63	2.48	0.05
HN7	0.27	0.19	0.43	0.36	0.18	0.47	0.16	0.37	0.51	0.12
HN8	0.04	0.08	0.98	0.71	0.12	2.28	0.11	0.44	2.57	0.02
HN9	0.12	0.13	0.36	0.32	0.13	0.71	0.11	0.29	0.79	0.04
HN10	0.08	0.13	0.63	0.63	0.16	1.39	0.11	0.28	2.53	0.02
HN11	0.06	0.16	0.74	0.63	0.18	1.62	0.16	0.47	1.68	0.03
HN12	0.19	0.11	0.38	0.28	0.10	0.56	0.09	0.30	0.61	0.05
HN13	0.18	0.23	0.44	0.58	0.21	0.88	0.18	0.49	1.12	0.06
HN14	0.00	0.04	0.49	0.30	0.06	0.59	0.05	0.34	0.83	0.01
HN15	0.01	0.06	1.43	1.01	0.13	2.43	0.10	0.69	3.80	0.02
HN16	0.02	0.07	1.00	0.37	0.08	0.70	0.06	0.28	0.87	0.02
HN17	0.01	0.04	1.01	0.85	0.08	2.38	0.10	0.32	3.25	0.01
HN18	0.15	0.13	0.52	0.36	0.13	0.90	0.09	0.33	1.17	0.01
HN19	0.06	0.08	1.14	0.93	0.15	2.98	0.10	0.36	3.41	0.02
HN20	0.36	0.22	0.50	0.34	0.21	0.88	0.14	0.31	0.96	0.05
HN21	0.02	0.05	1.03	0.74	0.12	2.17	0.06	0.57	3.08	0.01
HN22	0.06	0.07	0.89	0.72	0.14	1.73	0.09	0.25	2.68	0.02
HN23	0.04	0.08	0.28	0.61	0.09	0.74	0.10	0.24	0.71	0.01
HN24	0.02	0.08	1.12	0.78	0.17	2.25	0.15	0.39	3.42	0.03
HN25	0.01	0.05	0.98	0.73	0.10	2.24	0.07	0.37	2.85	0.01
HN26	0.01	0.05	1.11	0.92	0.11	2.15	0.12	0.30	3.50	0.01

Table 3.10 AAV DNA Barcode-Seq: Blood Clearance in Rhesus

	1 min	10 min	30 min	1 hr	4 hr	8 hr	24 hr	72 hr
AAV9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AAV9-N272A	0.95	0.96	1.06	1.05	1.24	1.79	2.64	1.18
PHP.B	1.00	1.10	1.16	1.11	1.27	1.99	2.94	1.69
HN1	0.78	0.66	0.71	0.65	0.67	0.88	0.65	0.30
HN2	0.69	0.72	0.69	0.62	0.75	0.92	1.24	0.25
HN3	0.93	1.00	0.96	1.00	1.41	1.48	2.31	1.41
HN4	1.05	1.16	1.08	1.19	1.18	2.00	3.52	1.54
HN5	1.03	0.98	1.08	1.07	1.18	1.77	2.72	1.56
HN6	0.85	0.86	0.96	0.77	1.12	1.70	2.60	0.60
HN7	0.75	0.78	0.81	0.79	0.95	1.33	1.78	0.27
HN8	1.05	1.06	1.02	1.07	1.27	1.82	3.02	1.67
HN9	0.83	0.76	0.87	0.78	0.92	1.36	1.91	0.29
HN10	0.84	0.87	0.90	0.87	1.14	1.32	2.32	0.28
HN11	0.88	0.89	0.90	0.86	1.20	1.55	2.42	0.61
HN12	0.81	0.80	0.86	0.80	0.99	1.41	1.90	0.27
HN13	0.76	0.71	0.82	0.76	0.90	1.31	1.86	0.29
HN14	0.76	0.75	0.83	0.78	0.99	1.40	2.21	0.18
HN15	1.31	1.07	1.27	1.23	1.36	2.37	4.46	2.03
HN16	0.76	0.80	0.91	0.84	0.98	1.41	1.90	0.26
HN17	1.02	1.07	1.21	1.00	1.29	1.88	2.91	1.84
HN18	0.88	0.88	0.96	0.90	1.10	1.57	2.44	0.58
HN19	1.08	1.04	1.15	1.14	1.26	2.04	3.69	2.05
HN20	0.89	0.82	0.88	0.81	0.91	1.62	2.28	0.39
HN21	1.00	1.00	1.11	0.98	1.20	1.54	2.41	2.10
HN22	0.96	0.97	1.09	1.03	1.24	1.74	2.90	1.82
HN23	0.76	0.76	0.86	0.80	0.97	1.50	2.14	0.35
HN24	1.07	1.07	1.15	1.15	1.30	1.88	3.39	2.15
HN25	0.99	0.98	1.14	1.06	1.19	1.62	2.83	1.39
HN26	1.13	1.09	1.18	1.09	1.32	1.97	3.32	1.90

Characterization of HN1 transduction efficiency and tropism in the non-human primate CNS

We next sought to validate the enhanced transduction efficiency phenotype of the HN1 variant in rhesus macaque. In order to overcome the challenges of using a small number of non-human primates that are susceptible to highly variable results (Gray et al. 2011), we developed a method that allowed for within-subjects comparisons of AAV9 and HN1 transduction on a per cell basis by using an epitope-tagged GFP reporter approach. AAV9 and HN1 capsids were packaged with CAG promoter-driven transgene cassettes encoding a GFP transgene fused with either an HA or a FLAG epitope at the N-terminus. In addition, we incorporated DNA/RNA barcode cassettes into the 3'UTR of these vector genomes. We generated four vectors: AAV9-CAG-FLAGnlsGFP-BCLib, HN1-CAG-HAnlsGFP-BCLib, AAV9-CAG-HAnlsGFP-BCLib, and HN1-CAG-FLAGnlsGFP-BCLib. We then mixed two libraries with reciprocal capsid-epitope relationships and co-administered each library into one of two rhesus macaques at a total dose of 3 x 10^{13} vg/kg intravenously. Animals were sacrificed three weeks post-injection and tissues were processed for immunofluorescent staining or frozen on dry ice for molecular studies. Overall transduction in the brain drastically differed in the two animals (Figure 3.12). While unfortunate, this finding supports our rationale for developing a within-subjects methodology.

Although our study was designed to use immunofluorescent staining of both the HA and FLAG epitope tags to confidently distinguish between AAV9- and HN1-mediated transduction, we were unable to achieve reliable anti-FLAG staining Figure 3.13. On the other hand, anti-HA staining was robust and we utilized this marker to

distinguish between AAV9 and HN1 transduction, focusing on the animal with higher transduction efficiency (Animal 1). Overall, we observed a strong bias towards HN1-transduced cells across the brain. The bias was particularly evident in the motor cortex, thalamus, and striatum (Figure 3.14a).

We quantified the proportion of neuronal transduction that was mediated by HN1 (GFP+/NeuN+/HA+) versus AAV9 (GFP+/NeuN+/HA-) and found that the bias toward HN1 was significantly deviated from a null hypothesis of equal transduction efficiencies (one-proportion z-tests with Bonferroni corrections for multiple comparisons, Figure 3.14b). The degree of HN1 enhancement ranged from ~5-fold (83% of transduced neurons) in the thalamus ($p < 0.0001$) to >10-fold (92% of transduced neurons) in the striatum ($p < 0.0001$). Conversely, HN1 transduced fewer glial cells than AAV9 in multiple brain regions. Indeed, HN1 transduced only 28% of S100+ glia in the striatum and thalamus, or ~2.5 fold less than AAV9 ($p < 0.0001$ and $p < 0.01$, respectively), and only 17% of glia in primary somatosensory cortex (~5-fold less than AAV9, $p < 0.001$).

We also quantified the tropism of AAV9 and HN1 in rhesus by calculating the proportion of transduction mediated by each vector that was neuronal (NeuN+) or glial (S100+). In most brain regions assayed, with the exception of hippocampus, HN1 exhibited a strong bias for neuronal transduction over glial transduction (Figure 3.14c). In the motor cortex, HN1 demonstrated a significant bias for transducing neurons compared to glia such that 88% of HN1-transduced cells were neurons and 11% were glia (two-proportion z-test with Bonferroni correction for multiple comparisons, $p < 0.0001$). In the thalamus, 95% of HN1-transduced cells were neuronal, compared to only 6% of transduced cells that were glial ($p < 0.0001$). In the striatum, we measured neuronal

tropism as 91% neuronal and 14% glial (p < 0.0001). Note that percentages do not add to exactly 100% because neuronal tropism and glial tropism were measured using independently stained sections. In contrast, AAV9 tropism was generally biased toward glia or demonstrated no statistically significant bias (Figure 3.14c).

In the spinal cord, HN1 mediated transduction of choline acetyltransferase (ChAT)-positive lower motor neurons. Transduction was observed in cervical, thoracic, and lumbar segments, but was greatest in the lumbar spinal cord (Figure 3.15a). We quantified relative transduction across all spinal segments and found that HN1 was ~4-fold more efficient than AAV9 (Figure 3.15b, p < 0.01). Notably, transduction appeared highly specific to lower motor neurons and we did not observe any obvious GFP signal in ChAT(-) cells.

Given recent reports of dorsal root ganglia (DRG) pathology following systemic AAV administration (Hordeaux et al. 2020; Hinderer et al. 2018), we histologically evaluated sensory neurons in the DRG. Transduction was observed in both animals, but was highly variable in Animal 2. However, observed variability in Animal 2 was not limited to GFP signal, but also apparent with staining of the neuronal PGP9.5 marker that showed a consistently robust signal in Animal 1 (Figure 3.16a). The representative figure for Animal 2 highlights the variability of PGP9.5 signal in contrast to robust and in-focus Hoechst signal. In addition, we observed a GFP+/HA+/PGP9.5+ locus with aberrant morphology associated with nuclear infiltration in Animal 2 (Figure 3.16b). Hematoxylin & Eosin staining of frozen DRG sections revealed a dense infiltration of mononuclear cells in Animal 2, but not Animal 1 (Figure 3.16c). We therefore focused quantification

of relative DRG sensory neuron transduction on Animal 1 and found that HN1 was

responsible for transduction in 71% of cases (Figure 3.16d, p < 0.001).
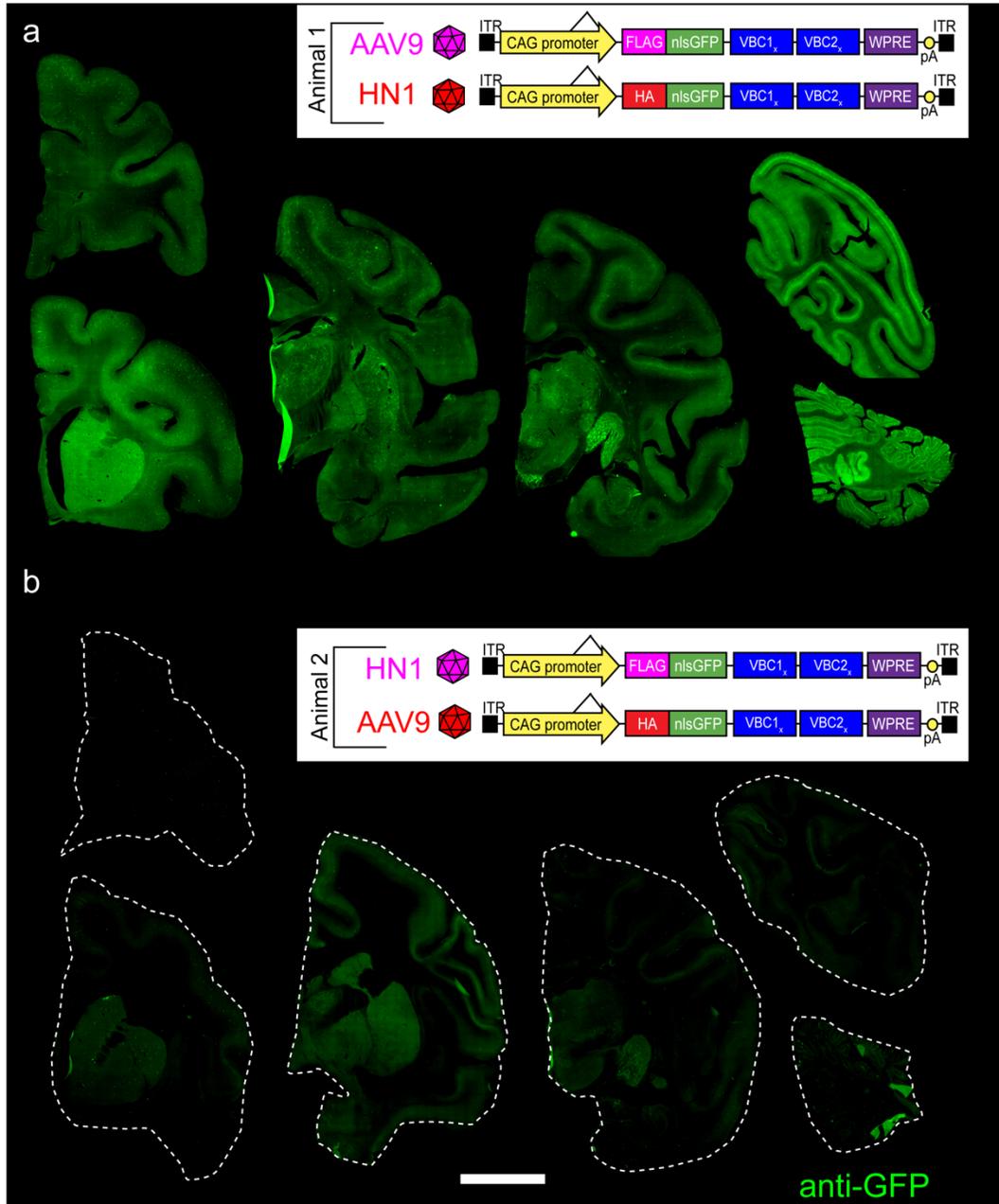
Figure 3.12 Widespread transduction of the rhesus brain following intravenous co-delivery of AAV9 and HN1 vectors

Within-subjects study design for comparison of AAV9 and HN1 capsid variants using immunofluorescence. Epitope-tagged transgene cassettes were individually packaged into AAV9 and HN1 capsids, then co-delivered into rhesus macaques such that staining for epitopes would allow identification of the transducing capsid. Overall transduction efficiency across whole brain hemisphere sections is represented by anti-GFP staining. Animal 2 exhibited a much lower overall transduction efficiency. Scale bar = 10 mm.
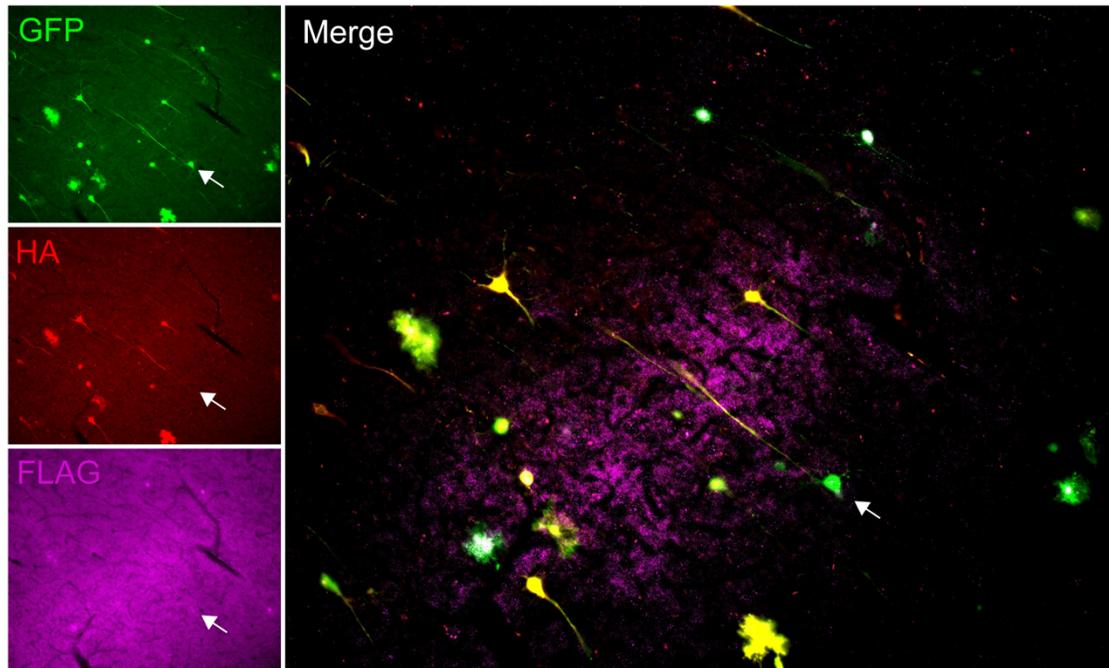
Figure 3.13 Reliable detection of the HA, but not the FLAG, epitope
Multiplexed immunofluorescent detection of GFP, HA, and FLAG.
Staining for GFP and HA was robust, but FLAG staining resulted in low
confidence signal. White arrow points to a GFP+/HA-/FLAG- cell.

a

b

c
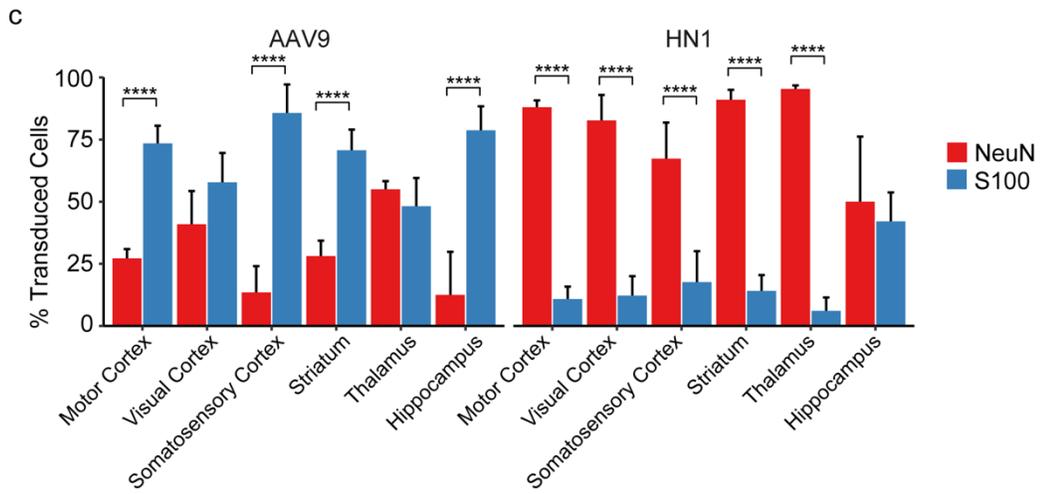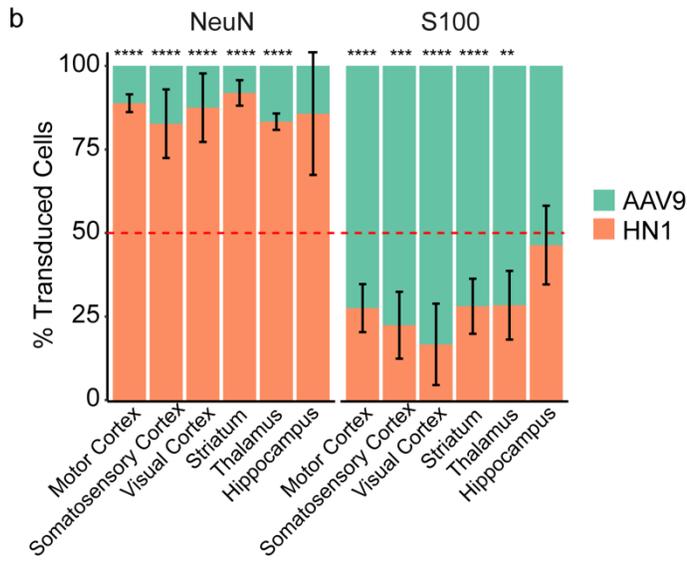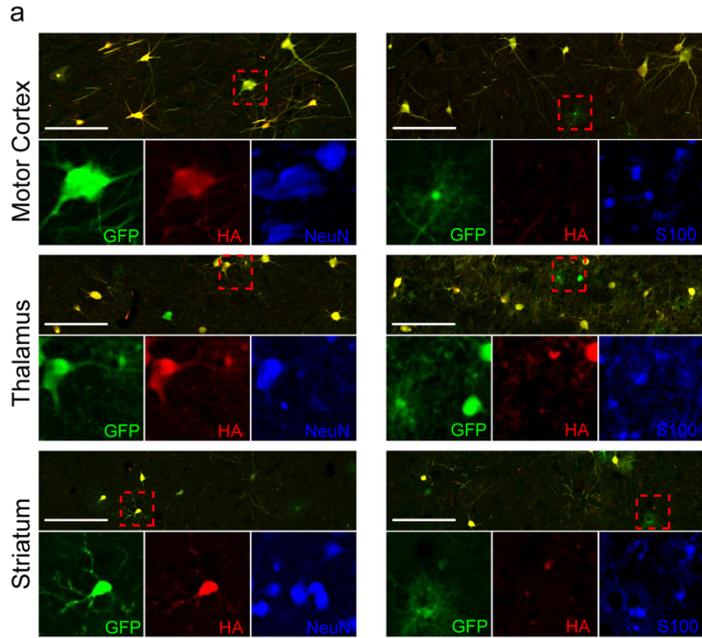
Figure 3.14 HN1 exhibits enhanced neuronal transduction efficiency and a strong preference for neurons over glia in the rhesus brain
(a) Representative optically sectioned images demonstrating enhanced neuronal transduction efficiency of HN1 in specific brain regions of Animal 1. Scale = 200 μm, inset = 100 μm x 100 μm. (b) Quantification of relative transduction efficiency. Of all transduced (GFP+) cells, we determined the proportion that were transduced by HN1 vs AAV9 (HA+ vs HA-, respectively). Percent ± the margin of error (95% confidence level) are plotted. Significance determined by one-proportion z-test and p-values were corrected using the Bonferroni method. (c) Quantification of tropism. The percentage of transduced cells that were neurons (NeuN+) or glial (S100+) are plotted for both AAV9 and HN1 with margin of error (95% confidence). Two-proportion z tests were used to determine if each capsid preferentially transduced one cell type over the other. P values corrected using Bonferroni method for multiple comparisons (n = 14, including spinal cord and dorsal root ganglia).
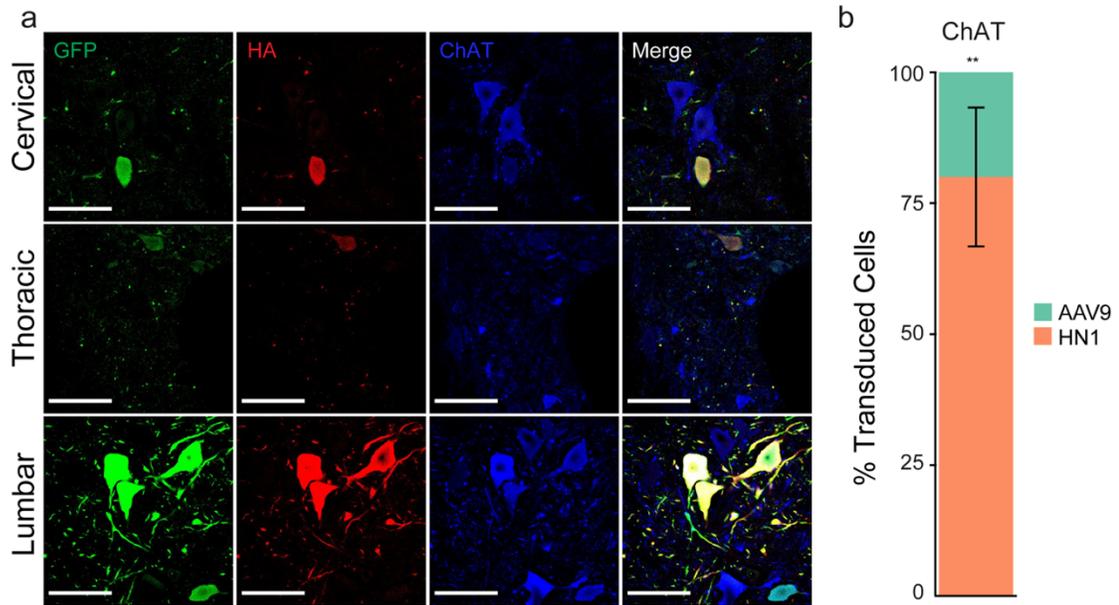
Figure 3.15 HN1 transduces lower motor neurons in the rhesus spinal cord

(a) AAV9-N272A-HN1 transduces lower motor neurons (ChAT+) in the spinal cord of Animal 1. Transduction efficiency was highest in the lumbar spinal cord, which also demonstrated much more robust expression per cell. Scale bar = 100 μm. (b) Quantification of the proportion of lower motor neuron transduction (GFP+/ChAT+) that was mediated by HN1 (HA+) versus AAV9 (HA-) in Animal1. Percent ± margin of error (95% confidence) is shown with one-proportion z test against a null hypothesis of 50% (**$p < 0.01$). P-value is Bonferroni corrected for 14 multiple comparisons, including those made in brain and dorsal root ganglia.
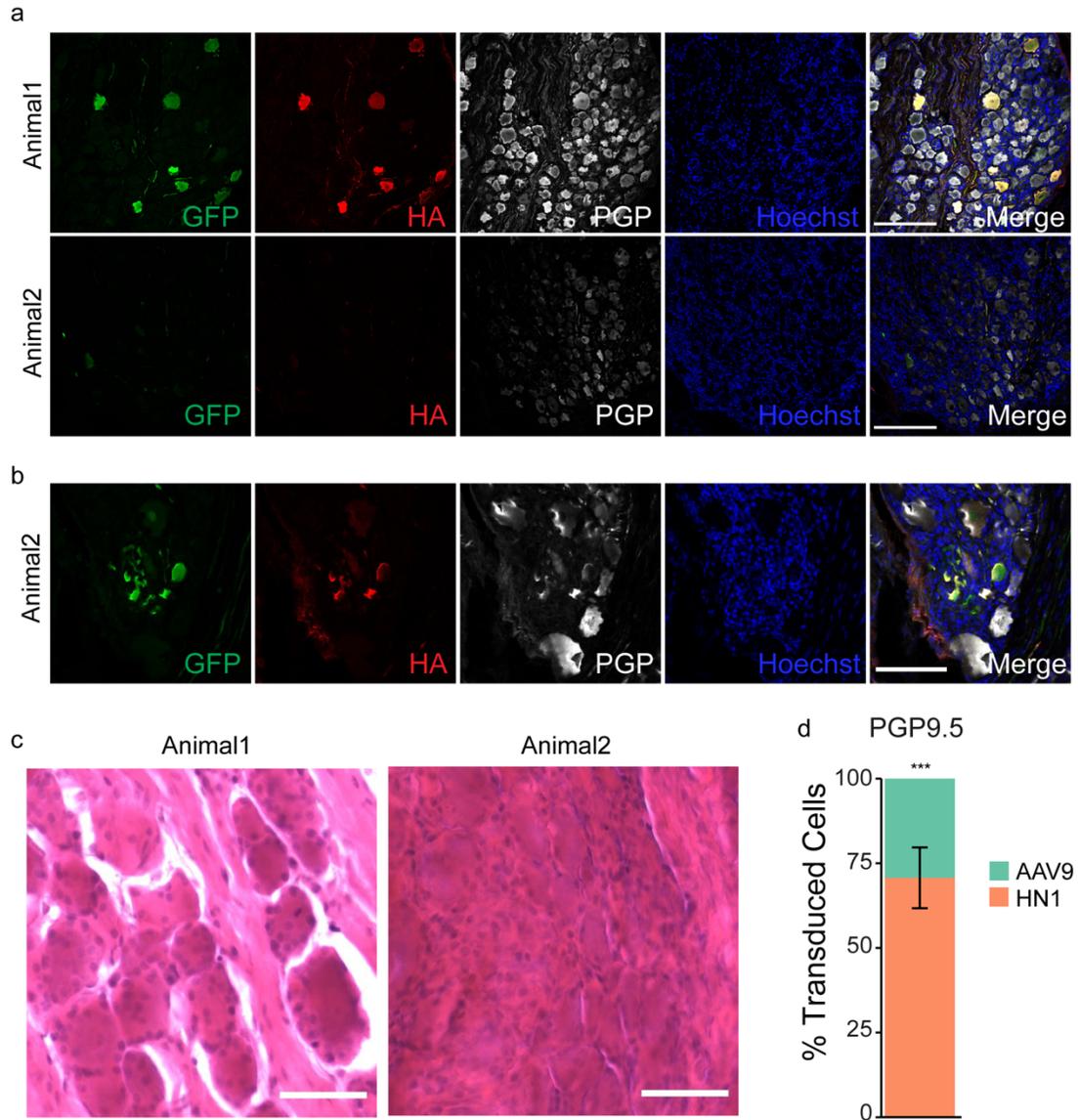
Figure 3.16 HN1 transduces sensory neurons in the dorsal root ganglia of rhesus macaque
(a) HN1 transduces sensory neurons in rhesus dorsal root ganglia. Low apparent transduction efficiency in Animal 2 was associated with more variable staining of the neuronal PGP9.5 marker. Scale bar = 200 µm. (b) Apparent degeneration of an AAV9-transduced sensory neuron consistent with previously described neuronophagia (Hordeaux et al. 2020). Note also the Hoechst+ nuclear infiltration into the neuronal soma. Scale bar = 100 µm. (c) H&E-stained rhesus dorsal root ganglia demonstrating a clear mononuclear infiltration in Animal 2. Scale bar = 50 µm. (d) Quantification of the proportion of sensory neuron transduction mediated by HN1 (HA+) in Animal 1. Percent ± margin of error (95% confidence) is shown with one-proportion z test (***$p < 0.001$, Bonferroni corrected).
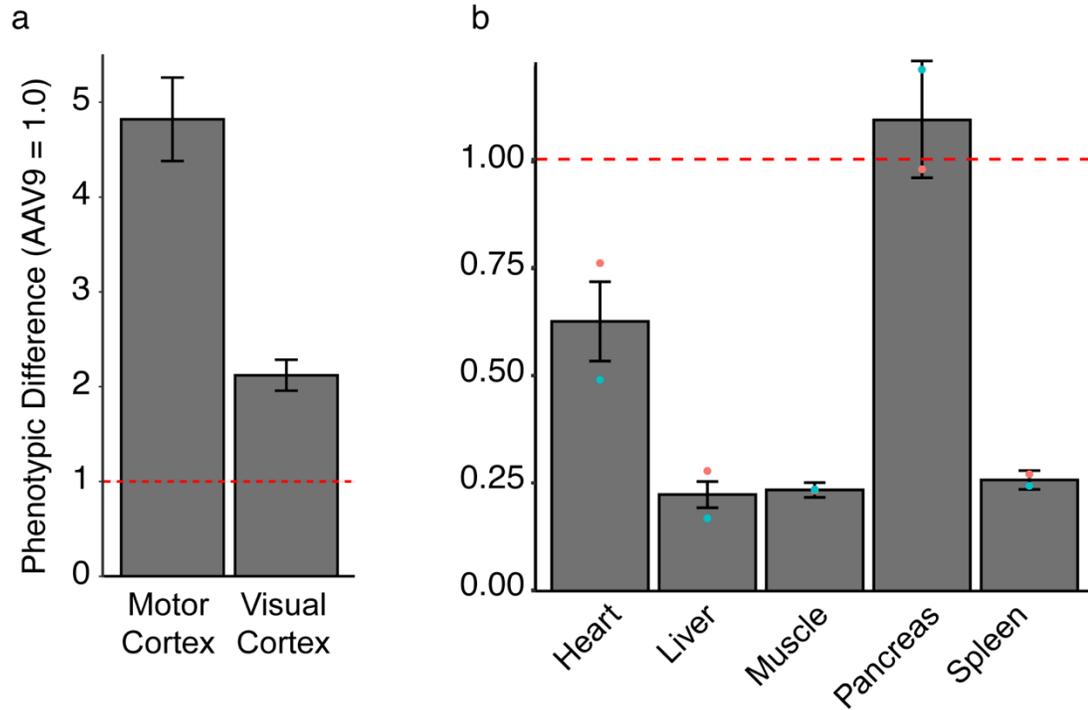
Cross-validation using CAG-driven RNA barcodes

We used AAV RNA Barcode-Seq as an orthologous approach to confirm the enhanced transduction efficiency based on CAG-expressed mRNA barcodes. We attempted to analyze 27 different CNS tissues, including numerous brain regions, three spinal cord segments, and the neural retina. Unexpectedly, different barcode sequences representing the same capsid exhibited a large degree of variation; we sometimes detected a high number of reads for one barcode sequence and completely failed to detect another barcode sequence representing the same capsid variant despite similar read counts in the AAV input vector.

We hypothesized that the high degree of variation may be due to a bottlenecking effect induced by a low transduction efficiency in the CNS. We therefore carried out AAV RNA Barcode-Seq analysis of peripheral tissues, including the liver, where we expected a high degree of transduction for at least AAV9. This analysis resulted in consistent read numbers across barcodes representing the same capsid variant. Furthermore, the relative transduction efficiencies of HN1 in these tissues (Figure 3.17b) were largely in line with the previously calculated relative biodistributions. Notably, we were unable to PCR amplify RNA barcodes from the lung and kidney.

We therefore sought to overcome this bottlenecking phenomenon by increasing the input for the RT-PCR reaction. We concentrated large amounts of cortical tissue using an mRNA purification kit and overloaded the RT-PCR reaction. This approach successfully allowed for robust recovery of barcode amplicons that did not demonstrate a bottlenecking effect following AAV RNA Barcode-Seq. Instead, analysis confirmed an

overall enhancement of HN1-mediated transduction efficiency over AAV9 in the motor

cortex (4.8-fold) and visual cortex (2.1-fold) of Animal 1 (Figure 3.17a).

**Figure 3.17 Cross-validation of immunofluorescence data using CAG-driven RNA barcodes.**
(a) Relative overall transduction efficiency of HN1 compared to AAV9 in motor and visual cortex of Animal 1 as determined by AAV RNA Barcode-Seq. (b) Relative overall transduction efficiency of HN1 compared to AAV9 in peripheral tissues of both animals. Barcode amplicons failed to amplify in kidney and lung.

HN1 produces at a level similar to that of AAV9

Finally, an important characteristic of vectors for clinical use, particularly those that will be delivered systemically, is the ability to produce at high titers. Therefore, we assayed HN1 production efficiency and directly compared it to production of AAV9 using a standard recombinant AAV production method (Figure 3.18). HN1 produced at a similar level to AAV9 (uncorrected Welch's t-test, p = 0.399).
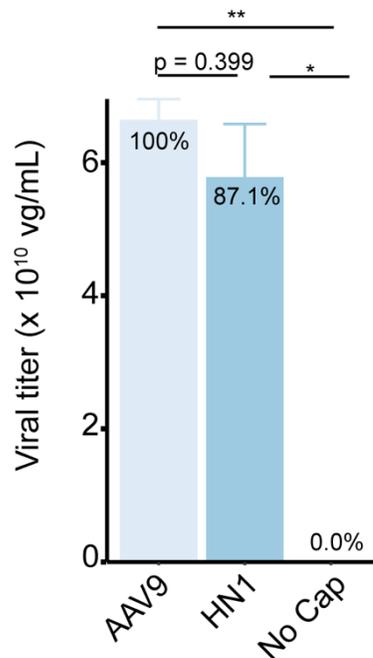
Figure 3.18 HN1 production is comparable to AAV9
pdsAAV-hSYN1-GFP vectors were packaged into either an AAV9 or HN1 capsid using a standard triple transfection protocol. As a negative control, the pRepCap was replaced with pHLP-Rep, a plasmid that did not express *cap*. Benzonase-resistant titers were quantified by radioactive dot blot assay that recognized the hSYN1 sequence and are plotted as mean ± SEM (n = 3 independent transfections / condition). Uncorrected Welch's t-test, *p < 0.05, **p < 0.01.

3.5   Discussion

Successful application of TRADE

The primary goal of this study was to demonstrate that the TRADE technology could be successfully employed to identify enhanced AAV variants. Regarding the primary goal, we report the successful application of the TRADE system to identify novel AAV capsid variants with enhanced transduction of brain neurons following intravenous vector administration. Transcription-dependent evolution utilizing a cell type-specific promoter-enhancer sequence to drive expression of the anti-sense cap sequence in rodents and rhesus led to the selection of 5 mouse-derived and 21 rhesus-derived capsid variant candidates. We utilized orthologous approaches to assay enhanced neuronal transduction efficiency including (1) AAV barcoding using a hSYN1-driven barcode cassette, and (2) a standard histological assay using immunofluorescence to detect hSYN1-driven GFP expression. Measurements of relative enhancement over AAV9 were highly concordant and showed a ~10-fold effect in rodents.

Nevertheless, we identified a major limitation of this study. Our analyses were focused on identifying 8mer peptide sequences flanked by the glycine-serine linkers and we overlooked a substantial amount of AAV-PHP.B contamination within the library. Importantly, our NGS analysis failed to identify this contamination because the PCR primers we used bind specifically to the glycine-serine linkers that are not present in the AAV-PHP.B capsid. We therefore repeated the NGS experiment using primers that would recognize both AAV-PHP.B and 8mer variants. We identified AAV-PHP.B contamination in the original AAVLib0 (0.066% of total reads). Following in vivo biopanning we observed that the proportion of AAV-PHP.B reads increased to 38.9% in

RNALib1, 14.0% in RNALib2, and 8.2% in RNALib3. This data suggested that AAV-PHP.B exhibited strong enrichment after the first round of selection, but the enrichment effect was diluted in subsequent selection rounds.

One potential explanation for this could be the known issues of cross-packaging and mosaicism. Cross-packaging occurs when a viral capsid packages a genomic sequence encoding some other protein sequence. Mosaicism occurs when different viral protein monomers are assembled into a chimeric viral particle. For in vivo biopanning, it is critical that the screened AAV library demonstrate a strong genotype-phenotype correlation between the packaged AAV genome (genotype) and capsid shell (phenotype). The effect of cross-packaging and mosaicism have been measured and their effect on in vivo selection is minimal if the AAV library production is carried out with limiting amounts of library plasmid (Nonnenmacher et al. 2015; Schmit et al. 2020). Importantly, it appears that capsid mosaicism is common, but the expected deleterious effect on genotype-phenotype correlation is attenuated because of a propensity for capsids to package the genome that encodes the majority of its viral protein monomers (Schmit et al. 2020). Based on this, we speculate that high levels of a strongly enhanced variant, such as AAV-PHP.B, may result in many capsid mosaics with a minority of AAV-PHP.B viral protein monomers. Such mosaic particles would be expected to package a non-AAV-PHP.B genome, but may still exhibit an enhanced phenotype associated with increased binding to the LY6A receptor and traversing the blood-brain-barrier (Q. Huang et al. 2019). In this scenario, further rounds of selection would lead to increased amounts of noise in enrichment scores for both AAV-PHP.B as well as other variants in the library. This mechanism may account for recent observations that more rounds of in vivo

selection can be deleterious (De Alencastro et al. 2020), and it also provides a strong

rationale for re-equilibrating libraries using oligo synthesis (Kumar et al. 2020;

Nonnenmacher et al. 2020).

We applied this theory to explain the unexpected results of our post-hoc

enrichment analysis that ranked our HN1 variant below that of HN3 and HN4, which had

failed to demonstrate enhancement by AAV RNA Barcode-Seq. Thus, we re-equilibrated

our library, performed an additional round of in vivo biopanning, and repeated the

enrichment analysis. These results fell much more in line with our expectation,

reidentifying both HN1 and HN2, demonstrating a lack of enrichment for HN3-5, and

also identifying two more potential AAV capsid variants for future study.

We therefore conclude that TRADE is a powerful approach for in vivo

biopanning, capable of stringently removing background signal while providing a flexible

platform for different cell type-specific promoters, different routes of administration, and

use of different biological models, including large pre-clinical models. On the other hand,

it is important to note that evidence of successful application in non-human primate is

preliminary.

We administered a low dose of only 2.8×10^{12} vg/kg intravenously for in vivo

biopanning. At the time, we reasoned that RT-PCR is a highly sensitive technique and

use of a low dose would provide more stringent screening conditions. However, in

retrospect, considering the diversity of the library and limited efficiency demonstrated by

the parental capsid, it was not entirely surprising that we were only able to recover 1-2

variant sequences from each nested RT-PCR reaction. The inability to recover a diversity

of sequences coupled with the recognition that most in vivo selection paradigms require

at least two rounds of selection did not give us a high degree of confidence that we would identify an enhanced, rhesus-derived, variant for use in non-human primate. Fortunately, our AAV Barcode-Seq approach allowed us to evaluate all of the selected variants. Notably, this analysis identified two variants, HN1 and HN6, that exhibited enhanced neuronal transduction efficiency in various rhesus brain regions. Although we focused our efforts on characterizing the HN1 variants, we note that the rhesus-derived HN6 variant exhibited a similar profile to HN1. Thus, while identification of HN1 demonstrates successful application of TRADE in mice with translation to the non-human primate context (see below), identification of HN6 may demonstrate the successful application of TRADE in non-human primate.

Characterization of the HN1 capsid variant in rodents

Comparison to AAV-PHP.B

We designed our experiments to not only show enhancement over the clinically-relevant AAV9 capsid, but also to benchmark HN1 against the now well-described AAV-PHP.B capsid (Deverman et al. 2016). We observed a clear difference regarding generalizability across mouse strains. As previously reported, AAV-PHP.B demonstrated enhancement in the LY6A(+) C57BL/6J strain, but not the LY6A(-) BALB/cJ strain (Hordeaux et al. 2018, 2019; Matsuzaki et al. 2019; Q. Huang et al. 2019; Batista et al. 2020). In contrast, HN1 showed no obvious strain-specific mechanism of enhancement. Notably, a number of variants have recently been reported that also do not exhibit a strain-dependent enhancement including AAV-F (Hanlon et al. 2019), AAV-PHP.C1-C3 (Kumar et al. 2020), and others (Nonnenmacher et al. 2020). We also observed a

difference in transduction pattern across the brain. Whereas AAV-PHP.B exhibited a high level of transduction throughout the entire brain, the degree of enhancement for HN1 was greater in the forebrain than the hindbrain. This is particularly interesting given that the first round of selection utilized RNA derived from only cortical forebrain.

Although both of our assays demonstrated an apparent equivalence for neuronal transduction efficiency in the forebrain between HN1 and AAV-PHP.B, we were suspicious of the relatively low fold-enhancement observed for AAV-PHP.B that was originally reported to be >40-fold (Deverman et al. 2016). Interestingly, we are not the first group to report unexpectedly low transduction efficiencies when using the PHP family of variants (Mathiesen et al. 2020; Hanlon et al. 2019). Understandably, such differences in observed efficiency may be overlooked in lieu of lab-to-lab differences with respect to assessing transduction efficiency (e.g., reporter genes, instrumentation, staining protocols, vector titration, etc.). However, we were intrigued by the high consistency exhibited by our two assays, one measuring relative transduction at the level of mRNA and the other at the level of neuronal cells expressing protein. Such consistency between orthologous assays suggested to us that the observed attenuated enhancement of AAV-PHP.B was not due to differences in assays, but perhaps due to our method of vector preparation.

An early observation came from our AAV Barcode-Seq data. As described in Methods, the AAV Barcode-Seq pipeline involves the independent production of each barcoded AAV clone so that a one-to-one relationship between capsid and viral barcode sequence can be established. Subsequently, crude lysates containing each barcoded AAV vector are titered and mixed into an equimolar pool that is purified as a single lot.

Because the AAV Barcode-Seq analysis takes into account the relative abundance of AAV variants in the AAV stock library, a certain amount of variation during titering and mixing of the crude lysates does not adversely affect the end results. However, the underlying assumption is that purification is an unbiased process. In the case of the scAAV-hSYN1-GFP-BCLib, however, AAV-PHP.B barcode clones were, on average, present at nearly 1/10th the level of AAV9 barcode clones following purification despite showing a nearly equal number of reads in the crude lysate pool. This clearly demonstrated that purification of AAV libraries is not always an unbiased process and provided a hint as to the unexpectedly low enhancement of AAV-PHP.B that we observed.

To the best of our knowledge, there have been no formal, peer-reviewed studies of factors that influence AAV-PHP.B purification, but the CalTech CLOVER (CLARITY, Optogenetics, & Vector Engineering) Center AAV FAQ (https://clover.caltech.edu/aav/faq) does warn that AAV-PHP.B may be prone to aggregation and recommends use of a high-salt buffer during purification. We thus repeated the experiment, but utilized the high-salt SAN buffer for purification (Challis et al. 2019), and observed a drastic difference in the performance of AAV-PHP.B. Indeed, our scAAV-PHP.B-hSYN1-GFP vector demonstrated such a high level of transduction efficiency that it initially appeared to be a homogeneous background signal. Due to this high signal, we were unable to confidently count individual GFP+ cells and instead utilized a GFP ELISA kit that still allowed us to assay transduction at the level of protein expression. This assay demonstrated a nearly 300-fold enhancement over AAV9. In contrast to the original report of AAV-PHP.B, we used self-complementary vectors,

which may account for the even larger observed enhancement effect. Notably, purification of AAV9 and HN1 vectors with the high-salt SAN buffer did not appear to affect performance. Therefore, the optimal performance of AAV-PHP.B was strongly dependent on the use of high-salt SAN buffer used to resuspend the AAV PEG pellet during vector purification (see Methods), and poor performance was observed when a HEPES-based containing the anionic surfactant Sarkosyl, was used instead. Note, the HEPES-based resuspension buffer has been previously reported (Grimm et al. 2003) and is commonly used in the field.

Further investigation of relative recovery in the scAAV-hSYN1-GFP-BCLib identified two groups of variants, those that were recovered efficiently (high recovery) and those that were lost (low recovery) during the purification process. A simple analysis suggested that a difference in hydrophobicity of the peptide insertion at the 3-fold axis of the capsid could predict differential recovery when purifying a mixed pool of AAV variants. Taken together with the clear difference in AAV-PHP.B performance based on resuspension buffer used, we agree with the speculation by CLOVER that failure to use a high-salt buffer may result in aggregation of the AAV-PHP.B capsid. We further posit that large aggregates of vector particles can precipitate, leading to precipitation and observable losses during titration. However, smaller aggregates may form that remain soluble in the final vector prep. These small aggregates would contribute to measurable titer, but could exhibit attenuated infectious activity, leading to observations of poor performance.

These data provide clear evidence that bulk purification of pooled AAV crude lysates is not necessarily an unbiased process, which may have important implications for

the field given that use of highly diverse AAV libraries for capsid evolution and high-throughput barcoding is becoming more common. In addition, although the field has been aware that purification methods can drastically alter vector performance (Klein et al. 2008; Sheridan 2019), our data provide insight specifically into the performance of the widely used AAV-PHP.B variant and may help address issues of reproducibility.

Finally, we return to the original intent of these studies – benchmarking the HN1 variant against AAV-PHP.B. We conclude that although the HN1 variant is enhanced relative to AAV9, it falls far short of the enhanced neuronal efficiency exhibited by AAV-PHP.B in the C57BL6 context. On the other hand, HN1 mediates cross-strain enhancement not observed by AAV-PHP.B.

Pharmacokinetic profiling: Biodistribution and blood clearance

Although achieving higher levels of neuronal transduction efficiency was the primary goal of our in vivo screen, we were also interested in evaluating other properties of novel variants related to their potential as clinical gene therapy vectors. Controlling vector tissue distribution is critical for both the safety profile of a gene therapy vector, as well as enhancing efficiency because the effective dose delivered to the CNS can be limited if vector particles rapidly accumulate in off-target peripheral tissues. We therefore utilized AAV DNA Barcode-Seq to evaluate the pharmacokinetic profile HN1, as well as all of the other variants included in the library.

For high-dose systemic administration, biodistribution to the liver can be a major dose-limiting factor, often necessitating the use of immunosuppressing regimens (REF). We therefore combined our directed evolution approach with a rationally designed liver-

detargeted parental capsid, AAV9-N272A (S. J. Huang et al. 2017). As expected,

biodistribution to the liver was reduced for all of the variants in the library. Interestingly,

the degree of liver detargeting appeared attenuated in variants that exhibited enhanced

brain transduction. This was unexpected given that selection of AAV variants that

efficiently cross the blood-brain-barrier and transduce the brain often exhibit lower levels

of biodistribution to the liver, independent of the N272A liver-detargeting mutation. For

example, our data show that AAV-PHP.B exhibits a similar degree of liver-detargeting as

HN1 in rodents despite not having the N272A mutation. In addition, the vast majority of

CNS-tropic variants identified by a similar transcription-dependent directed evolution

study exhibited liver detargeting (Nonnenmacher et al. 2020). This suggests that the liver

detargeting and enhanced CNS phenotypes interact. A simple explanation is that merely

detargeting the liver can lead to increased vector bioavailability in the brain.

In addition to liver detargeting, HN1 exhibited a moderate level of detargeting in

most other major tissues. Although many of the variants demonstrated a propensity to

accumulate in the liver, this was not the case for HN1. The only exception to the general

rule of peripheral detargeting for HN1 was in the lung. We briefly attempted to assay

HN1 transduction efficiency in the lung using mice injected with AAV9 and HN1 vectors

with a ssCAG-FLAGnlsGFP payload in combination with a whole tissue optical clearing

technique (Pincus et al. 2020), but observed very limited evidence of transduction (data

not shown). Notably, lack of GFP expression is not necessarily inconsistent with

increased biodistribution based on DNA barcode reads. We decided that the scope of our

study did not warrant further pursuit.

Previously, the blood clearance profile of AAV9 following intravenous dosing was shown to be delayed relative to other common serotypes. Importantly, the half-life was described as biphasic, including an early (i.e., within about 30 minutes) and rapid biodistribution, followed by a slower elimination phase (after 1 hour) (Kotchey et al. 2011). Our data suggest that HN1 both biodistributes more rapidly and is also eliminated more quickly than AAV9 in C57BL/6J mice. Coupled with our observations of detargeting from major peripheral organs, this suggests that the HN1 variant rapidly and preferentially biodistributes to the CNS following systemic administration.

Cell type tropism in the CNS

In addition to characterizing tissue tropism, we also evaluated the cell type tropism of HN1 within the CNS compared to AAV9 following intravenous delivery in adult mice. To establish a baseline, we evaluated AAV9 tropism and found that it varied by region, favoring glial transduction in the cortex, neuronal transduction in the thalamus, and showing no preference in the striatum. This is qualitatively in-line with previous reports that AAV9 transduces both neurons and astrocytes following intravenous administration in adult mice (Duque et al. 2009; Gray et al. 2011; Mathiesen et al. 2020).

In contrast, HN1 demonstrated a strong bias for neuronal transduction across the same brain regions. Intriguingly, the HN1 variant was selected to efficiently transduce brain neurons using a positive selective strategy with no additional negative selective pressures applied. Put another way, we expected our screen to identify variants with enhanced neuronal transduction efficiency, but not necessarily for the enhancement to be specific to neurons.

Notably, in our hands, transduction was limited to neurons, astrocytes, and likely endothelial cells that were not formally evaluated. We observed no transduction of oligodendrocytes (Olig2+) for either capsid at the dose administered. AAV9 is known to transduce oligodendrocytes, but efficiency is low even at a dose ~10-fold higher than that used in our study (Deverman et al. 2016; Powell, Samulski, and McCown 2020). We also did not observe transduction of microglia. Indeed, reports of AAV-mediated microglial transduction are limited in the literature and questionable in light of the many studies that report failure to transduce this cell type (reviewed by (Maes et al. 2019)).

Although the primary goal of administering CAG-GFP vectors into mice was to evaluate tropism, we took advantage of the opportunity to evaluate neuronal transduction efficiency using a different vector genome. In the context of a single-strand CAG-driven reporter genome, the relative neuronal transduction efficiency of HN1 over AAV9 in the murine cortex increased to >30-fold, compared to ~10-fold when using a self-complementary hSYN1-driven reporter. This observation intimates that at least part of HN1's mechanism of enhancement acts at the level of second-strand synthesis or later.

Finally, it is important to note that vector tropism is not solely dependent on the capsid sequence. Recent studies have found that the configuration of the AAV genome modifies vector tropism. Indeed, AAV-PHP.B, and the closely related AAV-PHP.eB, exhibited an increased tropism for astrocytes when a self-complementary genome was used compared to a single-stranded vector (Rincon et al. 2018; Mathiesen et al. 2020). In addition, it has become more clear that the capsid interacts with the promoter-enhancer sequence in a way that is still unclear (Powell, Samulski, and McCown 2020). Thus, our observations of a highly neurotropic capsid need to consider these recent discoveries.

Evaluation of HN1 in non-human primate

The secondary goal of this study was to identify novel capsid variants that may have clinical utility as gene therapy vectors for diseases of the central nervous system following non-invasive intravenous administration. Currently, AAV9 remains the standard for AAV-mediated CNS gene therapy as it is the only vector to receive regulatory approval following its paradigm-shifting demonstration in children with spinal muscular atrophy (Mendell et al. 2017). However, given recent and serious safety concerns regarding the administration of high-dose systemic AAV, including dose-limiting inflammation within the dorsal root ganglia (Hordeaux et al. 2020) and even death (Wilson and Flotte 2020), it is clear that enhanced vectors with improved therapeutic indices must be developed to realize the potential of AAV-mediated gene therapy. As many of the serious safety concerns only present in the context of super-high doses (Srivastava 2020), such improved therapeutic indices may be achieved by enhancing transduction efficiency for the target tissue and/or limiting biodistribution to off-target tissues. We therefore evaluated the HN1 variant to see if it conferred any potential clinical benefit over the gold standard AAV9 and found that many of the HN1 phenotypes observed in mice did translate.

A strength of the AAV Barcode-Seq technology is that allows for high-confidence within-subject comparisons. This is especially salient for phenotyping of variants using a small number of large animals that may be heterogeneous with respect to their genetics, exposures, age, sex, and other factors. Thus, using this technology and a single animal, we found that the HN1 variant showed evidence of a modest enhancement of neuronal

transduction efficiency in a variety of brain regions, up to 4.1-fold. The HN2 and AAV-PHP.B variants did not demonstrate enhancement, providing further evidence for the potential to select species-specific mechanisms of enhancement (Matsuzaki et al. 2018; Hordeaux et al. 2018; Liguore et al. 2019). Furthermore, we found that the general trend of detargeting peripheral tissues was consistent between mice and rhesus, notably exhibiting a 4-5 fold lower biodistribution to the liver compared to AAV9.

Additionally, we utilized xenograft FRG mice with chimeric mouse-human livers (Azuma et al. 2007) to assess biodistribution to human hepatocytes. We found that the HN1 liver detargeting phenotype translated in this context as well, but was not exhibited by other variants with enhanced brain transduction (i.e., HN2, HN6, or AAV-PHP.B). Intriguingly, the parental AAV9-N272A variant also did not demonstrate a liver detargeting phenotype in human hepatocytes despite robust liver detargeting in rhesus, suggesting a disconnect in how these models predict the clinical context. Given recent interest in using this chimeric xenograft model to predict AAV phenotypes in human (Lisowski et al. 2014; Pekrun et al. 2019; Vercauteren et al. 2016; Biswas et al. 2020; Havlik et al. 2020), it will be important for future studies to weigh the advantages of this model against non-human primates.

Regarding HN1 blood clearance following intravenous administration in rhesus, we observed a similar pattern to that in mice. HN1 appears to be cleared more rapidly than AAV9 during both the early biodistribution and late elimination phases, although the observed difference in rhesus was less than in mice.

In contrast to taking a barcode approach, standard histological assays of vector performance in large animals rely on between-subjects comparisons despite limited

sample sizes. This approach is susceptible to considerable variance between biological replicates that can complicate data interpretation. For example, a seminal study characterizing AAV9 in rhesus macaques observed a stark difference between animals that was not due to pre-existing neutralizing antibody titers (Gray et al. 2011). However, histological evaluation provides important information that cannot be easily obtained using current AAV barcoding technology, such as the expression of protein on a per cell basis. Therefore, we developed a method that would allow for histological evaluation of HN1 while facilitating a within-subjects comparison to AAV9.

Our method compares different AAV variants by associating each variant with a different epitope-tagged GFP transgene cassette, in this case either a FLAG-tagged GFP or an HA-tagged GFP associated with either HN1 or AAV9. In order to overcome potential biases with respect to the detection of each epitope, we generated two vector mixtures with reciprocal epitope-capsid relationships and administered each independently in our n = 2 experiment. Unfortunately, we observed a stark disparity in overall CNS transduction efficiency between the two animals, despite the theoretical equivalence of the two vector mixtures and lack of evidence for pre-existing neutralizing antibody titers or AAV binding antibodies. Notably, a recent study suggests that in vitro assays for detection of neutralizing antibodies may not be sufficient to predict in vivo performance following systemic vector delivery (D. Wang et al. 2018). Indeed, observable transduction in Animal 2 was so sparse that quantification was not feasible. Notably, the low transduction efficiency in Animal 2 was coincident with an apparent inflammatory response in the dorsal root ganglia. DRG pathology has been reported as a dose-limiting toxicity following delivery of systemic AAV vectors (Hordeaux et al. 2020;

Hinderer et al. 2018), and GFP+ signal in degenerating primary sensory neurons of the DRG has been previously reported (Schuster et al. 2014). Although this observation effectively limited the experiment to a sample size of one, it also underscores importance of pursuing the within-subjects methodology because, had we taken the standard approach and injected a single capsid variant into each animal, we might have incorrectly attributed differences in apparent transduction to the capsid variant.

An additional limitation of this experiment was a failure to detect robust FLAG signal. We utilized two epitopes so that we would not have to rely on the sensitivity of a single antibody to confidently identify the variant that mediated transduction of any given cell. Despite many different attempts to obtain robust FLAG staining, including seven different anti-FLAG antibodies and a variety of different methodologies including antigen retrieval and tyramide signal amplification, we consistently observed the same pattern – a generally low signal to noise ratio with a few identifiable cells. In review of the vector design, we found that the original GFP start codon was not removed when the original lab added the N-terminal nuclear localization signal (Addgene #104061). Our cloning strategy to fuse the N-terminal FLAG epitope replaced the start codon preceding the nuclear localization signal, but neglected to take into account the original GFP start codon. Thus, we speculate the addition of the FLAG tag, but not the HA tag, somehow favored use of the original start codon in the majority of cells, leading to GFP expression without the FLAG tag.

We also note that although the constructs contain a nuclear localization signal that sometimes concentrated GFP signal in the nucleus, expression was clearly observed in the cytoplasm as well. Given the robust detection of HA, which is upstream of the

nuclear localization signal, the poor nuclear localization cannot be explained by our speculative mechanism of a downstream start codon.

Due to these limitations, our quantification ultimately relied on the presence/absence of anti-HA staining in Animal 1. Notably, in this animal, the HA epitope was linked to the HN1 variant. Thus, if the sensitivity of HA staining was imperfect, we would underestimate the performance of HN1. A caveat of this analysis is that we assume the absence of co-infection by both vectors. We believe this to be the likely scenario because of the low transduction efficiency observed. However, this rationale requires the further assumption that transduction is not biased toward a specific population of cells. It is conceivable that a subset of cells is more susceptible to AAV transduction and co-infection by both HN1 and AAV9. In this scenario, we would grossly overestimate the relative efficiency of HN1 over AAV9 since AAV9 is quantified only as absence of signal. While we cannot definitively rule this possibility out, we note that following co-administration of two vectors in another CNS tissue, the retina, co-localization of both vector genomes in the same cell was extremely rare (S. K. Wang et al. 2020) and also address this with our barcode analysis (see below).

Notwithstanding these important caveats to interpretation of our immunofluorescence data, we found that HN1 outperformed AAV9 in the rhesus brain following intravenous administration. In certain brain regions, HN1 mediated >90% of all transduction events in neurons and, conversely, >90% of all HN1-mediated transduction events were neuronal, demonstrating a clear enhancement in both neuronal transduction efficiency and tropism. Additionally, HN1 demonstrated significantly more efficient transduction of the spinal cord, albeit with a high margin of error.

In order to cross-validate our immunofluorescence data we again utilized our AAV Barcode-Seq technology, this time taking advantage of CAG-driven barcodes cassettes. We ambitiously attempted to quantify relative transduction efficiency in 27 CNS tissues using AAV RNA Barcode-Seq. Unfortunately, we experienced a technical limitation during analysis such that barcodes representing the same capsid variant returned highly variable read counts, as low as being undetected. We observed no such issue in the assayed peripheral tissue. Given the relatively low transduction efficiency in the CNS tissues compared to the peripheral tissues, we interpreted this as a bottlenecking phenomenon whereby a limited number of vector genome-derived transcripts for RT-PCR resulted in a biased readout. This situation is analogous to our limited ability to recover capsid variants following a low dose library injection in rhesus. We hypothesized that we could overcome this issue by increasing the input to the RT-PCR reaction. We therefore concentrated RNA from large sample of tissue (~1g) using an mRNA purification kit and repeated the AAV Barcode-Seq analysis. This led to an expected readout and demonstrated an approximately 5-fold and 2-fold enhancement in overall transduction efficiency in motor and visual cortex, respectively.

Importantly, the CAG-driven barcodes are unable to distinguish between neuronal and glial transduction as we were able to using histological quantification. Since HN1 exhibited greater transduction efficiency in neurons, but lesser transduction efficiency in glia, the overall difference in CAG-driven barcodes should fall somewhere in between. As a rough estimate, we re-analyzed the data and found HN1 transduces only 2.5-fold more total cells than AAV9 in motor cortex and 1.6-fold more cells in visual cortex. The remaining discordance between the AAV Barcode-Seq data and the

immunofluorescence data may be accounted for by differences in the relative difference of CAG activity in different cell types consistent with a recent study reporting that CAG-driven GFP expression was stronger in neurons than glia in the rat brain (Chatterjee et al. 2021).

To summarize this section, we faced multiple challenges, including unexpectedly low overall transduction efficiency in a pre-screened animal, difficulty detecting (or perhaps expressing) the FLAG tag in vivo, and barcode bottlenecking. Notwithstanding the generalizability of a study using a sample size of 1, our within-subjects and orthogonal assays allow us to confidently conclude that HN1 neuronal transduction efficiency is enhanced in this test subject.

Conclusion

In conclusion, we successfully utilized the TRADE system to identify a novel AAV capsid variant that exhibits enhanced neuronal transduction efficiency after intravenous administration, on the order of 10-fold greater than the current clinical-grade standard AAV9 vector in both rodents and non-human primate. Additionally, HN1 is relatively detargeted from many peripheral tissues, including the liver, and exhibits a strong neuronal tropism. Taken together, these studies demonstrate the potential of HN1 as a clinical candidate for systemically-delivered CNS gene therapy and pave the way for the development of next-generation clinically-relevant AAV variants.

4 Chapter 4. Discussion and Conclusion

4.1 Discussion

Development of TRADE

In this age of molecular medicine, the critical role of gene delivery methods has never been more apparent. Rapid developments in uncovering the molecular basis of disease or manipulating the genome (i.e., CRISPR/Cas and base editing systems) are only effective if they can be safely and efficiently delivered into the relevant target cell type. To this end, AAV vectors have invigorated the field of gene therapy by safely and efficiently delivering genes to a wide variety of clinically-relevant targets.

Despite its origin as a contaminant deemed "defective" for its inability to replicate (Atchison, Casto, and Hammon 1965), over fifty years of dedicated research has resulted in the current reality of clinical gene therapy mediated by AAV vectors. It is interesting to note, however, that the majority of clinical trials continue to use the first AAV isolate, AAV2 (Kuzmin et al. 2021). Furthermore, AAV vectors continue to be produced by simple transfections of human cell lines, an approach developed more than 20 years ago (Matsushita et al. 1998). Recently, AAV9, another natural AAV isolate, has risen to stardom due to the profound demonstration that it could, with a single administration, mediate sufficiently safe and efficacious gene delivery to treat (and potentially cure) the most common genetic cause of infant mortality, spinal muscular atrophy (Mendell et al. 2017).

Nevertheless, current applications are limited to diseases in tissues that can be transduced by naturally-occurring vectors. Furthermore, recent events in clinical trials highlight unaddressed safety concerns that were observed years ago (Flotte 2020). To elaborate on an analogy I first heard from Dr. Arun Srivastava, a pioneer of AAV biology, current clinical applications of the naturally-occurring AAV serotypes are tantamount to traveling by horse and buggy. The road may be bumpy and slow, but that vehicle can certainly take you many places. Current engineering approaches are developing novel capsids that may be likened to the invention of the car. Cars may get us places faster and safer, but ultimately, we are still limited to land travel. To discover new areas, we need to develop air travel, or even space travel; thus, unlocking new gene therapy applications will require drastically better enhanced vectors.

In order to address this need and overcome the challenges of previous methods, we developed the transcription-dependent directed evolution system (TRADE). The goal of the TRADE system was to mediate highly stringent selection of functional AAV variants using a flexible platform that did not depend on transgenic animals, thus facilitating application in clinically-relevant models.

We stratify recent technological advances for in vivo biopanning of AAV capsid libraries into 4 generations (Table 4.1) that are distinguished by the level of capsid sequence selection and mechanism for targeting specific cell types (Figure 4.1). The first demonstration of in vivo directed evolution was carried out using adenovirus recovery of replicative AAV genomes (Grimm et al. 2008). Although this method theoretically mediates a high degree of functional selectivity, it is severely limited by the tropism (and virulence) of adenovirus. Given its limited application, it is designated Generation 0.

Generation 1 refers to directed evolution strategies that indiscriminately recover all of the vector genome DNA in the tissue of interest. In theory, this approach is susceptible to a high background of non-functional AAV variants. Furthermore, targeting of a specific cell type requires some additional mechanism for cell sorting. For example, a transgenic line expressing GFP in photoreceptors was used to enrich for AAV variants that targeted this cell type (Dalkara et al. 2013). Generation 2 includes Cre-dependent selection strategies such as CREATE (Deverman et al. 2016). These methods greatly reduce the background of non-functional variants by selecting at the level of double strand vector genome DNA. However, selection and cell type specificity are dependent on Cre-driver animals. Generation 2.5 represents the iTransduce system (Hanlon et al. 2019). Although this system increases the stringency of selection to the level of vector-mediated protein expression, it retains the requirement for Cre and a transgenic animal (Ai9). Furthermore, the requirement for cell sorting may serve as a barrier to capturing maximal diversity as it was for certain applications of Generation 1 designs. Therefore, its utility over the CREATE system may be incremental. Finally, Generation 3 consists of transcription-dependent systems including TRACER (Nonnenmacher et al. 2020) and the TRADE system described here. These platforms select at the level of transgene expression and do not require any transgenic animals. Cell type specificity is dictated by selection of a promoter-enhancer sequence. Together, these systems represent the most advanced systems for in vivo directed evolution.

The key differences between TRACER and TRADE were discussed in Chapter 2. In brief, the design of TRADE minimizes co-expression of the p40-driven sense *cap* transcript and the cell type-specific promoter-driven antisense *cap* transcript. Doing so

may benefit vector production and limit the production of immunogenic capsids during

biopanning. Also of note is the potential to utilize the optional reporter gene in the

TRADE system to select at the level of protein-expression, as in iTransduce. Notably, the

use of Cre in the iTransduce platform does increase sensitivity, but perhaps does so at the

cost of selecting transiently-expressing AAV variants (Lang et al. 2019).

Table 4.1 Advances in *in vivo* AAV capsid directed evolution technologies

Generation	Platform	Cre-dependent	Level of selection	Cell specificity	Reference
0	Adenovirus rescue	No	productive replication	limited to adenoviral tropism	(Grimm et al. 2008)
1	DNA PCR	No	any vector genome DNA	requires cell sorting	(Dalkara et al. 2013)
2	CREATE	Yes	double strand vector genome DNA	requires Cre driver	(Deverman et al. 2016)
2.5	iTransduce	Yes	protein expression	requires cell sorting	(Hanlon et al. 2019)
3	TRACER	No	sense mRNA expression	requires specific promoter-enhancer	(Nonnenmacher et al. 2020)
3	TRADE	No	anti-sense mRNA expression	requires specific promoter-enhancer	unpublished
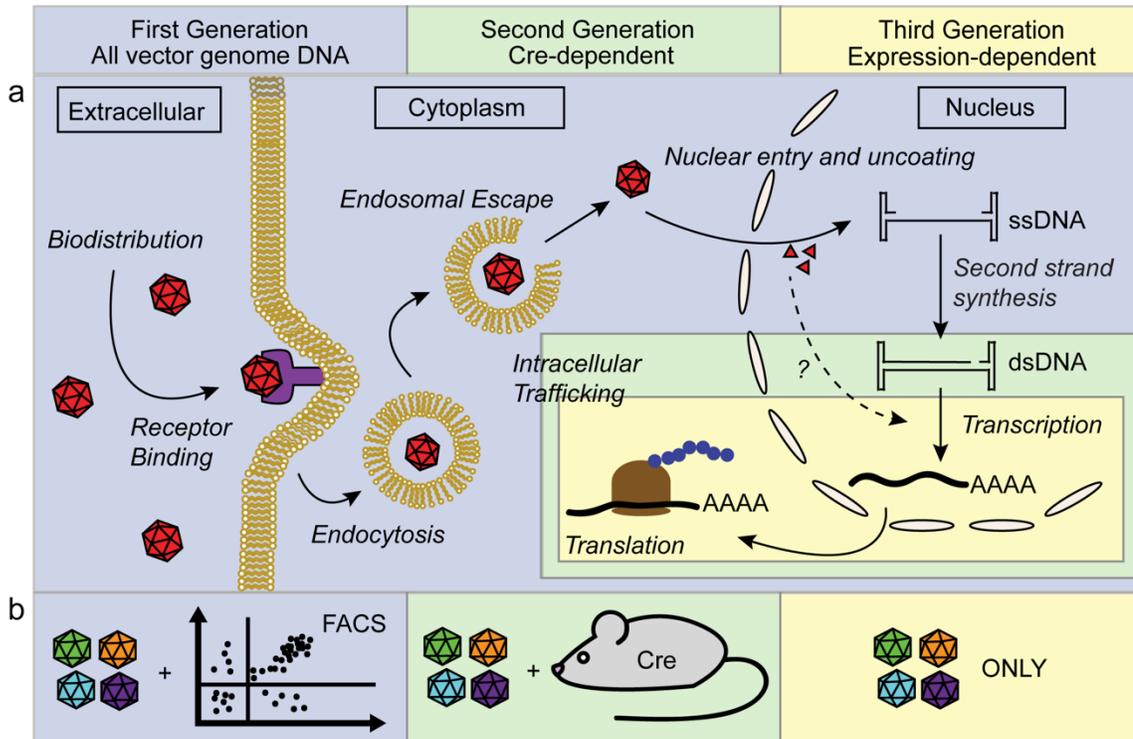
Figure 4.1 Classification of recently developed in vivo AAV capsid directed evolution systems
(a) Overview of the complex, multi-step process of AAV vector mediated transduction of a single cell. Following administration, the AAV vector must biodistribute to the target tissue and cell type. Subsequently, the vector is internalized by receptor-mediated endocytosis. Acidification of the endosome mediates externalization of a phospholipase A2 domain that allows the vector to escape and traffic to the nucleus. Following nuclear entry and breakdown of the capsid, the single strand DNA genome is converted to a double strand form that is transcriptionally active, thus allowing expression of transgene mRNA and protein. (b) Specific cell types that have been transduced can be isolated by cell sorting, use of Cre-driver animals to limit recombination events, or driving expression of the transgene using. a cell type-specific promoter-enhancer. Background highlighting identifies the level of selection for each technological generation (a) and mechanism for mediating cell type specific selection.)

181

Application of TRADE

We applied our TRADE system to identify novel variants with enhanced brain

neuronal transduction efficiency following intravenous administration into C57BL/6J

mice or rhesus macaque. We screened a capsid library that contained elements of both a

rationally designed liver detargeting mutation (N272A) on the AAV9 capsid and a

randomized 8 amino acid peptide displayed on the capsid surface at the three-fold

protrusions. Rather than selecting only 1 or 2 lead candidates for further validation, we

utilized AAV Barcode-Seq technology to simultaneously phenotype 26 potential hits,

again utilizing both rodents and rhesus macaque. AAV RNA Barcode-Seq using hSYN1-

driven barcodes identified an ~10-fold enhancement in neuronal transduction efficiency

in two mouse strains that was highly consistent with subsequent validating experiments

using immunofluorescence and ELISA to detect GFP. Notably, however, when the

promoter-enhancer sequence was changed from hSYN1 to CAG, the measured relative

neuronal transduction efficiency increased up to ~30-fold. This observation may fit with

recent reports that the AAV capsid plays a role in driving promoter activity (Powell,

Samulski, and McCown 2020; Bohlen et al. 2020), and further provides a rationale for

utilizing a transcription-dependent platform for AAV directed evolution.

AAV RNA Barcode-Seq using hSYN1-driven barcodes in the non-human primate

brain demonstrated a moderate level of enhancement across multiple brain regions. Use

of a within-subjects immunofluorescence strategy allowed us to validate the enhancement

in neuronal transduction efficiency using a CAG-GFP construct. We again observed a

discrepancy such that the observed relative neuronal transduction efficiency under a

hSYN1 promoter-enhancer was 2~4x, but closer to ~10x using a CAG promoter-

enhancer. Notably, however, each experiment contained a single animal and differences could be entirely due to individual differences.

We utilized AAV DNA Barcode-Seq to evaluate vector biodistribution to the liver. HN1 liver detargeting was attenuated compared to the parental AAV9-N272A capsid, yet consistently demonstrated ~4-fold reduction compared to AAV9 across mouse strains, rhesus, and human hepatocytes (xenograft model).

The clinical utility of HN1 remains to be determined. Although it is tempting to argue that the observed 4-fold greater transduction efficiency of motor neurons, coupled with a 4-fold reduction in liver biodistribution would combine to effectively increase the therapeutic index of an AAV-HN1-SMN therapy and reduce the risk of hepatotoxicity (Feldman et al. 2020), these observations must be confirmed in a larger cohort of animals.

Finally, the identification of any variant that outperforms the gold standard AAV9 for CNS transduction is significant. To our knowledge, no such variant has been published in a peer reviewed journal. However, the Gradinaru group has published a pre-print demonstrating that their AAV.CAP-B10 is superior to AAV9 in marmoset brain (Flytzanis et al. 2020) and a very recent presentation by Voyager Therapeutics at the American Society for Gene & Cell Therapy claimed that the transcription-dependent TRACER platform identified an AAV variant with 1000-fold better performance than AAV9 following intravenous administration (Nonnenmacher 2021).

4.2   Conclusion

The field of AAV-mediated gene therapy is rapidly developing, yet continues to utilize vectors that are far from optimized. In some cases, such unoptimized use has led to safety concerns and this is likely the underlying cause of death for three patients in the last year. On the other hand, AAV-mediated gene therapy has demonstrated clear success and allowed the potentially curative treatment of diseases that were previously unaddressed. In order to realize the true potential of AAV-mediated gene therapy, there is a critical need to develop enhanced AAV vectors that can be utilized in safe and efficacious therapies.

Altogether, this work has resulted in the development and successful application of a novel, transcription-dependent AAV capsid directed evolution platform termed TRADE. TRADE adds to the rapidly growing field of capsid engineering and promotes a future of customizable AAV vectors that can be leveraged to address a wide variety of diseases with different gene delivery requirements. Furthermore, we identify the HN1 capsid variant, which exhibits enhanced neuronal transduction efficiency and specificity in the non-human primate and may be useful as clinical vector for CNS gene therapy.

We believe this work contributes to an imminent future when the potential of AAV vectors can fully come to bear on unmet needs in both basic science medicine, leading to scientific breakthroughs and novel therapies.

REFERENCES

Adachi, Kei, Tatsuji Enoki, Yasuhiro Kawano, Michael Veraz, and Hiroyuki Nakai. 2014. "Drawing a High-Resolution Functional Map of Adeno-Associated Virus Capsid by Massively Parallel Sequencing." *Nature Communications* 5: 3075. https://doi.org/10.1038/ncomms4075.

Akache, Bassel, Dirk Grimm, Kusum Pandey, Stephen R. Yant, Hui Xu, and Mark A. Kay. 2006. "The 37/67-Kilodalton Laminin Receptor Is a Receptor for Adeno-Associated Virus Serotypes 8, 2, 3, and 9." *Journal of Virology* 80 (19): 9831–36. https://doi.org/10.1128/jvi.00878-06.

Alencastro, Gustavo De, Katja Pekrun, Paul Valdmanis, Matthew Tiffany, Jianpeng Xu, and Mark A. Kay. 2020. "Tracking Adeno-Associated Virus Capsid Evolution by High-Throughput Sequencing." *Human Gene Therapy* 31 (9–10): 553–64. https://doi.org/10.1089/hum.2019.339.

Arnett, A. L.H., L. R. Beutler, A. Quintana, J. Allen, E. Finn, R. D. Palmiter, and J. S. Chamberlain. 2013. "Heparin-Binding Correlates with Increased Efficiency of AAV1-and AAV6-Mediated Transduction of Striated Muscle, but Negatively Impacts CNS Transduction." *Gene Therapy* 20 (5): 497–503. https://doi.org/10.1038/gt.2012.60.

Arnold, Frances H. 2018. "Directed Evolution: Bringing New Chemistry to Life." *Angewandte Chemie - International Edition* 57 (16): 4143–48. https://doi.org/10.1002/anie.201708408.

Asokan, Aravind, Julie B. Hamra, Lakshmanan Govindasamy, Mavis Agbandje-McKenna, and Richard J. Samulski. 2006. "Adeno-Associated Virus Type 2 Contains an Integrin A5β1 Binding Domain Essential for Viral Cell Entry." *Journal of Virology* 80 (18): 8961–69. https://doi.org/10.1128/jvi.00843-06.

Atchison, RW, Bruce C Casto, and William McD Hammon. 1965. "Adenovirus-Associated Defective Virus Particles." *Science* 149 (3685): 754–56.

Audentes. 2020. "Audentes Therapeutics Provides Update on the ASPIRO Clinical Trial Evaluating AT132 in Patients with X-Linked Myotubular Myopathy." Audentes Press Release. 2020. https://www.audentestx.com/press_release/audentes-therapeutics-provides-update-on-the-aspiro-clinical-trial-evaluating-at132-in-patients-with-x-linked-myotubular-myopathy/.

Aydemir, Fikret, Maxim Salganik, Justyna Resztak, Jasbir Singh, Antonette Bennett, Mavis Agbandje-McKenna, and Nicholas Muzyczka. 2016. "Mutants at the 2-Fold Interface of Adeno-Associated Virus Type 2 (AAV2) Structural Proteins Suggest a Role in Viral Transcription for AAV Capsids." *Journal of Virology* 90 (16): 7196–7204. https://doi.org/10.1128/jvi.00493-16.

Azuma, Hisaya, Nicole Paulk, Aarati Ranade, Craig Dorrell, Muhsen Al-Dhalimy, Ewa Ellis, Stephen Strom, Mark A. Kay, Milton Finegold, and Markus Grompe. 2007. "Robust Expansion of Human Hepatocytes in Fah-/-/Rag2 -/-/Il2rg-/- Mice." *Nature Biotechnology* 25 (8): 903–10. https://doi.org/10.1038/nbt1326.

Bainbridge, James W.B., Alexander J. Smith, Susie S. Barker, Scott Robbie, Robert Henderson, Kamaljit Balaggan, Ananth Viswanathan, et al. 2008. "Effect of Gene Therapy on Visual Function in Leber's Congenital Amaurosis." *New England Journal of Medicine* 358 (21): 2231–39. https://doi.org/10.1056/NEJMoa0802268.

Bantel-Schaal, Ursula, Birgit Hub, and Juergen Kartenbeck. 2002. "Endocytosis of Adeno-Associated Virus Type 5 Leads to Accumulation of Virus Particles in the Golgi Compartment." *Journal of Virology* 76 (5): 2340–49. https://doi.org/10.1128/jvi.76.5.2340-2349.2002.

Bartlett, Jeffrey S., Rose Wilcher, and R. Jude Samulski. 2000. "Infectious Entry Pathway of Adeno-Associated Virus and Adeno-Associated Virus Vectors." *Journal of Virology* 74 (6): 2777–85. https://doi.org/10.1128/jvi.74.6.2777-2785.2000.

Batista, Ana Rita, Oliver D. King, Christopher P. Reardon, Crystal Davis, Shankaracharya, Vivek Philip, Heather Gray-Edwards, et al. 2020. "Ly6a Differential Expression in Blood-Brain Barrier Is Responsible for Strain Specific Central Nervous System Transduction Profile of AAV-PHP.B." *Human Gene Therapy* 31 (1–2): 90–102. https://doi.org/10.1089/hum.2019.186.

Becerra, S P, F Koczot, P Fabisch, and J A Rose. 1988. "Synthesis of Adeno-Associated Virus Structural Proteins Requires Both Alternative MRNA Splicing and Alternative Initiations from a Single Transcript." *Journal of Virology* 62 (8): 2745–54. https://doi.org/10.1128/jvi.62.8.2745-2754.1988.

Bell, C. L., B. L. Gurda, K. Van Vliet, M. Agbandje-McKenna, and J. M. Wilson. 2012. "Identification of the Galactose Binding Domain of the Adeno-Associated Virus Serotype 9 Capsid." *Journal of Virology* 86 (13): 7326–33. https://doi.org/10.1128/jvi.00448-12.

Bennett, Jean, Manzar Ashtari, Jennifer Wellman, Kathleen A Marshall, Laura L. Cyckowski, Daniel C Chung, Sarah McCague, et al. 2012. "Gene Therapy: AAV2 Gene Therapy Readministration in Three Adults with Congenital Blindness." *Science Translational Medicine* 4 (120): 1–24. https://doi.org/10.1126/scitranslmed.3002865.

Bennicelli, Jeannette, John Fraser Wright, Andras Komaromy, Jonathan B. Jacobs, Bernd Hauck, Olga Zelenaia, Federico Mingozzi, et al. 2008. "Reversal of Blindness in Animal Models of Leber Congenital Amaurosis Using Optimized AAV2-Mediated Gene Transfer." *Molecular Therapy* 16 (3): 458–65. https://doi.org/10.1038/sj.mt.6300389.

Besse, Aurore, Stephanie Astord, Thibaut Marais, Marianne Roda, Benoit Giroux, François Xavier Lejeune, Frederic Relaix, Piera Smeriglio, Martine Barkats, and Maria Grazia Biferi. 2020. "AAV9-Mediated Expression of SMN Restricted to Neurons Does Not Rescue the Spinal Muscular Atrophy Phenotype in Mice." *Molecular Therapy* 28 (8): 1887–1901. https://doi.org/10.1016/j.ymthe.2020.05.011.

Bey, Karim, Johan Deniaud, Laurence Dubreil, Béatrice Joussemet, Joseph Cristini, Carine Ciron, Juliette Hordeaux, et al. 2020. "Intra-CSF AAV9 and AAVrh10 Administration in Nonhuman Primates: Promising Routes and Vectors for Which Neurological Diseases?" *Molecular Therapy - Methods and Clinical Development*

17 (June): 771–84. https://doi.org/10.1016/j.omtm.2020.04.001.

Biswas, Moanaro, Damien Marsic, Ning Li, Chenhui Zou, Gloria Gonzalez-Aseguinolaza, Irene Zolotukhin, Sandeep R.P. Kumar, et al. 2020. "Engineering and In Vitro Selection of a Novel AAV3B Variant with High Hepatocyte Tropism and Reduced Seroreactivity." *Molecular Therapy - Methods and Clinical Development* 19 (December): 347–61. https://doi.org/10.1016/j.omtm.2020.09.019.

Bleker, Svenja, Florian Sonntag, and Jürgen A. Kleinschmidt. 2005. "Mutational Analysis of Narrow Pores at the Fivefold Symmetry Axes of Adeno-Associated Virus Type 2 Capsids Reveals a Dual Role in Genome Packaging and Activation of Phospholipase A2 Activity." *Journal of Virology* 79 (4): 2528–40. https://doi.org/10.1128/jvi.79.4.2528-2540.2005.

Bohlen, Martin O., Thomas J. McCown, Sara K. Powell, Hala G. El-Nahal, Tierney Daw, Michele A. Basso, Marc A. Sommer, and R. Jude Samulski. 2020. "Adeno-Associated Virus Capsid-Promoter Interactions in the Brain Translate from Rat to the Nonhuman Primate." *Human Gene Therapy* 31 (21–22): 1155–68. https://doi.org/10.1089/hum.2020.196.

Bossis, Ioannis, and John A. Chiorini. 2003. "Cloning of an Avian Adeno-Associated Virus (AAAV) and Generation of Recombinant AAAV Particles." *Journal of Virology* 77 (12): 6799–6810. https://doi.org/10.1128/jvi.77.12.6799-6810.2003.

Boutin, Sylvie, Virginie Monteilhet, Philippe Veron, Christian Leborgne, Olivier Benveniste, Marie Françoise Montus, and Carole Masurier. 2010. "Prevalence of Serum IgG and Neutralizing Factors against Adeno-Associated Virus (AAV) Types 1, 2, 5, 6, 8, and 9 in the Healthy Population: Implications for Gene Therapy Using AAV Vectors." *Human Gene Therapy* 21 (6): 704–12. https://doi.org/10.1089/hum.2009.182.

Boye, Sanford L., Antonette Bennett, Miranda L. Scalabrino, K. Tyler McCullough, Kim Van Vliet, Shreyasi Choudhury, Qing Ruan, James Peterson, Mavis Agbandje-McKenna, and Shannon E. Boye. 2016. "Impact of Heparan Sulfate Binding on Transduction of Retina by Recombinant Adeno-Associated Virus Vectors." *Journal of Virology* 90 (8): 4215–31. https://doi.org/10.1128/jvi.00200-16.

Burghes, Arthur H.M., and Christine E. Beattie. 2009. "Spinal Muscular Atrophy: Why Do Low Levels of Survival Motor Neuron Protein Make Motor Neurons Sick?" *Nature Reviews Neuroscience* 10 (8): 597–609. https://doi.org/10.1038/nrn2670.

Byrne, Leah C., Timothy P. Day, Meike Visel, Jennifer A. Strazzeri, Cécile Fortuny, Deniz Dalkara, William H. Merigan, David V. Schaffer, and John G. Flannery. 2020. "In Vivo-Directed Evolution of Adeno-Associated Virus in the Primate Retina." *JCI Insight* 5 (10). https://doi.org/10.1172/JCI.INSIGHT.135112.

Calcedo, Roberto, Luk H. Vandenberghe, Guangping Gao, Jianping Lin, and James M. Wilson. 2009. "Worldwide Epidemiology of Neutralizing Antibodies to Adeno-Associated Viruses." *Journal of Infectious Diseases* 199 (3): 381–90. https://doi.org/10.1086/595830.

Cao, Maohua, Hong You, and Paul L. Hermonat. 2014. "The X Gene of Adeno-

Associated Virus 2 (AAV2) Is Involved in Viral DNA Replication." *PLoS ONE* 9 (8). https://doi.org/10.1371/journal.pone.0104596.

Carpentier, Andre C., Frederique Frisch, Sebastien M. Labbe, Rene Gagnon, Janneke De Wal, Stephen Greentree, Harald Petry, Jaap Twisk, Diane Brisson, and Daniel Gaudet. 2012. "Effect of Alipogene Tiparvovec (Aav1-Lpls447x) on Postprandial Chylomicron Metabolism in Lipoprotein Lipase-Deficient Patients." *Journal of Clinical Endocrinology and Metabolism* 97 (5): 1635–44. https://doi.org/10.1210/jc.2011-3002.

Challis, Rosemary C., Sripriya Ravindra Kumar, Ken Y. Chan, Collin Challis, Keith Beadle, Min J. Jang, Hyun Min Kim, et al. 2019. "Systemic AAV Vectors for Widespread and Targeted Gene Delivery in Rodents." *Nature Protocols* 14 (2): 379–414. https://doi.org/10.1038/s41596-018-0097-3.

Chan, Ken Y., Min J. Jang, Bryan B. Yoo, Alon Greenbaum, Namita Ravi, Wei Li Wu, Luis Sánchez-Guardado, et al. 2017. "Engineered AAVs for Efficient Noninvasive Gene Delivery to the Central and Peripheral Nervous Systems." *Nature Neuroscience* 20 (8): 1172–79. https://doi.org/10.1038/nn.4593.

Chandler, Randy J., Matthew C. La Fave, Gaurav K. Varshney, Niraj S. Trivedi, Nuria Carrillo-Carrasco, Julien S. Senac, Weiwei Wu, et al. 2015. "Vector Design Influences Hepatic Genotoxicity after Adeno-Associated Virus Gene Therapy." *Journal of Clinical Investigation* 125 (2): 870–80. https://doi.org/10.1172/JCI79213.

Chandler, Randy J., Matthew C. LaFave, Gaurav K. Varshney, Shawn M. Burgess, and Charles P. Venditti. 2016. "Genotoxicity in Mice Following AAV Gene Delivery: A Safety Concern for Human Gene Therapy?" *Molecular Therapy* 24 (2): 198–201. https://doi.org/10.1038/mt.2016.17.

Chao, Hengjun, Yuanbo Liu, Joseph Rabinowitz, Chengwen Li, Richard Jude Samulski, and Christopher E. Walsh. 2000. "Several Log Increase in Therapeutic Transgene Delivery by Distinct Adeno-Associated Viral Serotype Vectors." *Molecular Therapy* 2 (6): 619–23. https://doi.org/10.1006/mthe.2000.0219.

Chatterjee, Diptaman, David J. Marmion, Jodi L. McBride, Fredric P. Manfredsson, David Butler, Anne Messer, and Jeffrey H. Kordower. 2021. "Enhanced CNS Transduction from AAV.PHP.EB Infusion into the Cisterna Magna of Older Adult Rats Compared to AAV9." *Gene Therapy*. https://doi.org/10.1038/s41434-021-00244-y.

Ciesielska, Agnieszka, Piotr Hadaczek, Gabriele Mittermeyer, Shangzhen Zhou, J. Fraser Wright, Krystof S. Bankiewicz, and John Forsayeth. 2013. "Cerebral Infusion of AAV9 Vector-Encoding Non-Self Proteins Can Elicit Cell-Mediated Immune Responses." *Molecular Therapy* 21 (1): 158–66. https://doi.org/10.1038/mt.2012.167.

Dalkara, Deniz, Leah C. Byrne, Ryan R. Klimczak, Meike Visel, Lu Yin, William H. Merigan, John G. Flannery, and David V. Schaffer. 2013. "In Vivo-Directed Evolution of a New Adeno-Associated Virus for Therapeutic Outer Retinal Gene Delivery from the Vitreous." *Science Translational Medicine* 5 (189): 1–12. https://doi.org/10.1126/scitranslmed.3005708.

Dalwadi, Dhwanil A., Laura Torrens, Jordi Abril-Fornaguera, Roser Pinyol, Catherine Willoughby, Jeffrey Posey, Josep M. Llovet, et al. 2021. "Liver Injury Increases the Incidence of HCC Following AAV Gene Therapy in Mice." *Molecular Therapy* 29 (2): 680–90. https://doi.org/10.1016/j.ymthe.2020.10.018.

Davidsson, Marcus, Gang Wang, Patrick Aldrin-Kirk, Tiago Cardoso, Sara Nolbrant, Morgan Hartnor, Malin Parmar, and Tomas Bjorklund. 2018. "Barcoded Rational AAV Vector Evolution Enables Systematic in Vivo Mapping of Peptide Binding Motifs." *BioRxiv*, 335372. https://doi.org/10.1101/335372.

Depla, Josse A., Marina Sogorb-Gonzalez, Lance A. Mulder, Vivi M. Heine, Pavlina Konstantinova, Sander J. van Deventer, Katja C. Wolthers, Dasja Pajkrt, Adithya Sridhar, and Melvin M. Evers. 2020. "Cerebral Organoids: A Human Model for AAV Capsid Selection and Therapeutic Transgene Efficacy in the Brain." *Molecular Therapy - Methods and Clinical Development* 18 (September): 167–75. https://doi.org/10.1016/j.omtm.2020.05.028.

Desmet, François Olivier, Dalil Hamroun, Marine Lalande, Gwenaëlle Collod-Bëroud, Mireille Claustres, and Christophe Béroud. 2009. "Human Splicing Finder: An Online Bioinformatics Tool to Predict Splicing Signals." *Nucleic Acids Research* 37 (9): 1–14. https://doi.org/10.1093/nar/gkp215.

Deverman, Benjamin E, Piers L Pravdo, Bryan P Simpson, Sripriya Ravindra Kumar, Ken Y Chan, Abhik Banerjee, Wei-Li Wu, et al. 2016. "Cre-Dependent Selection Yields AAV Variants for Widespread Gene Transfer to the Adult Brain." *Nature Biotechnology* advance on (January 2015): 1–7. https://doi.org/10.1038/nbt.3440.

Donsante, Anthony, Daniel G. Miller, Yi Li, Carole Vogler, Elizabeth M. Brunt, David W. Russell, and Mark S. Sands. 2007. "AAV Vector Integration Sites in Mouse Hepatocellular Carcinoma." *Science* 317 (5837): 477. https://doi.org/10.1126/science.1142658.

Dudek, Amanda M, Nerea Zabaleta, Eric Zinn, Sirika Pillay, James Zengel, Jennifer Santos Franceschini, Reynette Estelien, Jan E Carette, Guo Ling, and Luk H Vandenberghe. 2019. "GPR108 Is a Highly Conserved AAV Entry Factor." *Molecular Therapy* 28 (2): 1–15. https://doi.org/10.1016/j.ymthe.2019.11.005.

Duque, Sandra, Béatrice Joussemet, Christel Riviere, Thibaut Marais, Laurence Dubreil, Anne Marie Douar, John Fyfe, Philippe Moullier, Marie Anne Colle, and Martine Barkats. 2009. "Intravenous Administration of Self-Complementary AAV9 Enables Transgene Delivery to Adult Motor Neurons." *Molecular Therapy* 17 (7): 1187–96. https://doi.org/10.1038/mt.2009.71.

Excoffon, Katherine J.D.A., James T. Koerber, David D. Dickey, Matthew Murtha, Shaf Keshavjee, Brian K. Kaspar, Joseph Zabner, and David V. Schaffer. 2009. "Directed Evolution of Adeno-Associated Virus to an Infectious Respiratory Virus." *Proceedings of the National Academy of Sciences of the United States of America* 106 (10): 3865–70. https://doi.org/10.1073/pnas.0813365106.

Farkas, Szilvia L., Zoltán Zádori, Mária Benko, Sandra Essbauer, Balázs Harrach, and Peter Tijssen. 2004. "A Parvovirus Isolated from Royal Python (Phyton Regius) Is a Member of the Genus Dependovirus." *Journal of General Virology* 85 (3): 555–61.

https://doi.org/10.1099/vir.0.19616-0.

Farris, K. David, and David J. Pintel. 2008. "Improved Splicing of Adeno-Associated Viral (AAV) Capsid Protein-Supplying Pre-MRNAs Leads to Increased Recombinant AAV Vector Production." *Human Gene Therapy* 19 (12): 1421–27. https://doi.org/10.1089/hum.2008.118.

Feldman, Amy G., Julie A. Parsons, Cullen M. Dutmer, Aravindhan Veerapandiyan, Einar Hafberg, Nolan Maloney, and Cara L. Mack. 2020. "Subacute Liver Failure Following Gene Replacement Therapy for Spinal Muscular Atrophy Type 1." *Journal of Pediatrics* 225: 252-258.e1. https://doi.org/10.1016/j.jpeds.2020.05.044.

Flotte, Terence R. 2020. "Revisiting the 'New' Inflammatory Toxicities of Adeno-Associated Virus Vectors." *Human Gene Therapy* 31 (7–8): 398–99. https://doi.org/10.1089/hum.2020.29117.trf.

Flotte, Terence R., Kevin Sullivan, Randall Wetzel, George Taylor, Barrie J. Carter, William B. Guggino, Pamela L. Zeitlin, et al. 2003. "Phase I Trial of Intranasal and Endobronchial Administration of a Recombinant Adeno-Associated Virus Serotype 2 (Raav2)-Cftr Vector in Adult Cystic Fibrosis Patients: A Two-Part Clinical Study." *Human Gene Therapy* 14 (11): 1079–88. https://doi.org/10.1089/104303403322124792.

Flotte, Terence R, Sandra A Afione, and Pamela L Zeitlin. 1994. "Adeno-Associated Virus Vector Gene Expression Occurs in Nondividing Cells in the Absence of Vector DNA Integration." *American Journal of Respiratory Cell and Molecular Biology* 11: 517–21.

Flytzanis, Nicholas, Nick Goeden, David Goertsen, Alexander Cummins, James Pickel, and Viviana Gradinaru. 2020. "Broad Gene Expression throughout the Mouse and Marmoset Brain after Intravenous Delivery of Engineered AAV Capsids," no. Ic: 1–21. https://doi.org/10.1101/2020.06.16.152975.

Foust, Kevin D., Emily Nurre, Chrystal L. Montgomery, Anna Hernandez, Curtis M. Chan, and Brian K. Kaspar. 2009. "Intravascular AAV9 Preferentially Targets Neonatal Neurons and Adult Astrocytes." *Nature Biotechnology* 27 (1): 59–65. https://doi.org/10.1038/nbt.1515.

Foust, Kevin D., Xueyong Wang, Vicki L. McGovern, Lyndsey Braun, Adam K. Bevan, Amanda M. Haidet, Thanh T. Le, et al. 2010. "Rescue of the Spinal Muscular Atrophy Phenotype in a Mouse Model by Early Postnatal Delivery of SMN." *Nature Biotechnology* 28 (3): 271–74. https://doi.org/10.1038/nbt.1610.

Franchini, Massimo, and Pier Mannuccio Mannucci. 2012. "Past, Present and Future of Hemophilia: A Narrative Review." *Orphanet Journal of Rare Diseases* 7 (1): 1–8. https://doi.org/10.1186/1750-1172-7-24.

Gao, Guangping, Mauricio R Alvira, Lili Wang, Roberto Calcedo, Julie Johnston, and James M Wilson. 2002. "Novel Adeno-Associated Viruses from Rhesus Monkeys" 99 (18): 11854–59. https://doi.org/10.1073/pnas.182412299.

Gao, Guangping, Luk H Vandenberghe, Mauricio R Alvira, You Lu, Roberto Calcedo, Xiangyang Zhou, and James M Wilson. 2004. "Clades of Adeno-Associated Viruses

Are Widely Disseminated in Human Tissues." *Journal of Virology* 78 (12): 6381–88. https://doi.org/10.1128/JVI.78.12.6381.

Garita-Hernandez, Marcela, Fiona Routet, Laure Guibbal, Hanen Khabou, Lyes Toualbi, Luisa Riancho, Sacha Reichman, et al. 2020. "AAV-Mediated Gene Delivery to 3D Retinal Organoids Derived from Human Induced Pluripotent Stem Cells." *International Journal of Molecular Sciences* 21 (3): 1–16. https://doi.org/10.3390/ijms21030994.

Gaudet, D., J. Méthot, S. Déry, D. Brisson, C. Essiembre, G. Tremblay, K. Tremblay, et al. 2013. "Efficacy and Long-Term Safety of Alipogene Tiparvovec (AAV1-LPL S447X) Gene Therapy for Lipoprotein Lipase Deficiency: An Open-Label Trial." *Gene Therapy* 20 (4): 361–69. https://doi.org/10.1038/gt.2012.43.

Girod, Anne, Christiane E. Wobus, Zoltán Zádori, Martin Ried, Kristin Leike, Peter Tijssen, Jürgen A. Kleinschmidt, and Michael Hallek. 2002. "The VP1 Capsid Protein of Adeno-Associated Virus Type 2 Is Carrying a Phospholipase A2 Domain Required for Virus Infectivity." *Journal of General Virology* 83 (5): 973–78. https://doi.org/10.1099/0022-1317-83-5-973.

Gottlieb, Scott. 2019. "Statement from FDA Commissioner Scott Gottlieb, M.D. and Peter Marks, M.D., Ph.D., Director of the Center for Biologics Evaluation and Research on New Policies to Advance Development of Safe and Effective Cell and Gene Therapies." FDA Statement. 2019.

Govindasamy, Lakshmanan, Eric Padron, Robert McKenna, Nicholas Muzyczka, Nikola Kaludov, John A. Chiorini, and Mavis Agbandje-McKenna. 2006. "Structurally Mapping the Diverse Phenotype of Adeno-Associated Virus Serotype 4." *Journal of Virology* 80 (23): 11556–70. https://doi.org/10.1128/jvi.01536-06.

Gray, Steven J, Bonita L Blake, Hugh E Criswell, Sarah C Nicolson, R Jude Samulski, and Thomas J Mccown. 2009. "Directed Evolution of a Novel Adeno-Associated Virus (AAV) Vector That Crosses the Seizure-Compromised Blood–Brain Barrier (BBB)." *Molecular Therapy* 18 (3): 570–78. https://doi.org/10.1038/mt.2009.292.

Gray, Steven J, Valerie Matagne, Lavanya Bachaboina, Swati Yadav, Sergio R Ojeda, and R Jude Samulski. 2011. "Preclinical Differences of Intravascular AAV9 Delivery to Neurons and Glia : A Comparative Study of Adult Mice and Nonhuman Primates." *Molecular Therapy* 19 (6): 1058–69. https://doi.org/10.1038/mt.2011.72.

Gray, Steven J, S. Nagabhushan Kalburgi, Thomas J McCown, and R. Jude Samulski. 2013. "Global CNS Gene Delivery and Evasion of Anti-AAV-Neutralizing Antibodies by Intrathecal AAV Administration in Non-Human Primates." *Gene Therapy* 20 (4): 450–59. https://doi.org/10.1038/gt.2012.101.

Grimm, Dirk, Joyce S. Lee, Lora Wang, Tushar Desai, Bassel Akache, Theresa A. Storm, and Mark A. Kay. 2008. "In Vitro and In Vivo Gene Therapy Vector Evolution via Multispecies Interbreeding and Retargeting of Adeno-Associated Viruses." *Journal of Virology* 82 (12): 5887–5911. https://doi.org/10.1128/jvi.00254-08.

Grimm, Dirk, Kusum Pandey, Hiroyuki Nakai, Theresa A. Storm, and Mark A. Kay. 2006. "Liver Transduction with Recombinant Adeno-Associated Virus Is Primarily

Restricted by Capsid Serotype Not Vector Genotype." *Journal of Virology* 80 (1): 426–39. https://doi.org/10.1128/jvi.80.1.426-439.2006.

Grimm, Dirk, Shangzhon Zhou, Hiroyuki Nakai, Clare E. Thomas, Theresa A. Storm, Sally Fuess, Takashi Matsushita, et al. 2003. "Preclinical in Vivo Evaluation of Pseudotyped Adeno-Associated Virus Vectors for Liver Gene Therapy." *Blood* 102 (7): 2412–19. https://doi.org/10.1182/blood-2003-02-0495.

Grimm, Dirk, and Sergei Zolotukhin. 2015. "E Pluribus Unum: 50 Years of Research, Millions of Viruses, and One Goal-Tailored Acceleration of AAV Evolution." *Molecular Therapy* 23 (12): 1819–31. https://doi.org/10.1038/mt.2015.173.

GuhaSarkar, Dwijit, James Neiswender, Qin Su, Guangping Gao, and Miguel Sena-Esteves. 2017. "Intracranial AAV-IFN-β Gene Therapy Eliminates Invasive Xenograft Glioblastoma and Improves Survival in Orthotopic Syngeneic Murine Model." *Molecular Oncology* 11 (2): 180–93. https://doi.org/10.1002/1878-0261.12020.

Halbert, Christine L., James M. Allen, and A. Dusty Miller. 2001. "Adeno-Associated Virus Type 6 (AAV6) Vectors Mediate Efficient Transduction of Airway Epithelial Cells in Mouse Lungs Compared to That of AAV2 Vectors." *Journal of Virology* 75 (14): 6615–24. https://doi.org/10.1128/jvi.75.14.6615-6624.2001.

Hanlon, Killian S., Jonah C. Meltzer, Tetyana Buzhdygan, Ming J. Cheng, Miguel Sena-Esteves, Rachel E. Bennett, Timothy P. Sullivan, et al. 2019. "Selection of an Efficient AAV Vector for Robust CNS Transgene Expression." *Molecular Therapy - Methods and Clinical Development* 15 (December): 320–32. https://doi.org/10.1016/j.omtm.2019.10.007.

Hauck, Bernd, and Weidong Xiao. 2003. "Characterization of Tissue Tropism Determinants of Adeno-Associated Virus Type 1." *Journal of Virology* 77 (4): 2768–74. https://doi.org/10.1128/jvi.77.4.2768-2774.2003.

Hauswirth, William W., Tomas S. Aleman, Shalesh Kaushal, Artur V. Cideciyan, Sharon B. Schwartz, Lili Wang, Thomas J. Conlon, et al. 2008. "Treatment of Leber Congenital Amaurosis Due to *RPE65* Mutations by Ocular Subretinal Injection of Adeno-Associated Virus Gene Vector: Short-Term Results of a Phase I Trial." *Human Gene Therapy* 19 (10): 979–90. https://doi.org/10.1089/hum.2008.107.

Havlik, L. Patrick, Katherine E. Simon, J. Kennon Smith, Kelli A. Klinc, Longping V. Tse, Daniel K. Oh, Marco M. Fanous, et al. 2020. "Coevolution of Adeno-Associated Virus Capsid Antigenicity and Tropism through a Structure-Guided Approach." *Journal of Virology* 94 (19): 1–17. https://doi.org/10.1128/jvi.00976-20.

Herzog, Roland W., Edmund Y. Yang, Linda B. Couto, J. Nathan Hagstrom, Dan Elwell, Paul A. Fields, Melissa Burton, et al. 1999. "Long-Term Correction of Canine Hemophilia B by Gene Transfer of Blood Coagulation Factor IX Mediated by Adeno-Associated Viral Vector." *Nature Medicine* 5 (1): 56–63. https://doi.org/10.1038/4743.

Hinderer, Christian, Peter Bell, Charles H. Vite, Jean Pierre Louboutin, Rebecca Grant, Erin Bote, Hongwei Yu, Bryan Pukenas, Robert Hurst, and James M. Wilson. 2014.

"Widespread Gene Transfer in the Central Nervous System of Cynomolgus Macaques Following Delivery of AAV9 into the Cisterna Magna." *Molecular Therapy - Methods and Clinical Development* 1 (September): 14051. https://doi.org/10.1038/mtm.2014.51.

Hinderer, Christian, Nathan Katz, Elizabeth Lynne Buza, Cecilia Dyer, Tamara Goode, Peter Bell, Laura Richman, and James M. Wilson. 2018. "Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose Intravenous Administration of an AAV Vector Expressing Human SMN." *Human Gene Therapy* X (X): hum.2018.015. https://doi.org/10.1089/hum.2018.015.

Hioki, H., H. Kameda, H. Nakamura, T. Okunomiya, K. Ohira, K. Nakamura, M. Kuroda, T. Furuta, and T. Kaneko. 2007. "Efficient Gene Transduction of Neurons by Lentivirus with Enhanced Neuron-Specific Promoters." *Gene Therapy* 14 (11): 872–82. https://doi.org/10.1038/sj.gt.3302924.

Hordeaux, Juliette, Elizabeth L. Buza, Cecilia Dyer, Tamara Goode, Thomas W. Mitchell, Laura Richman, Nathan Denton, et al. 2020. "Adeno-Associated Virus-Induced Dorsal Root Ganglion Pathology." *Human Gene Therapy* 31 (15–16): 808–18. https://doi.org/10.1089/hum.2020.167.

Hordeaux, Juliette, Qiang Wang, Nathan Katz, Elizabeth L Buza, Peter Bell, and James M. Wilson. 2018. "The Neurotropic Properties of AAV-PHP.B Are Limited to C57BL/6J Mice." *Molecular Therapy* 26 (3): 664–68. https://doi.org/10.1016/j.ymthe.2018.01.018.

Hordeaux, Juliette, Yuan Yuan, Peter M. Clark, Qiang Wang, R. Alexander Martino, Joshua J. Sims, Peter Bell, Angela Raymond, William L. Stanford, and James M. Wilson. 2019. "The GPI-Linked Protein LY6A Drives AAV-PHP.B Transport across the Blood-Brain Barrier." *Molecular Therapy* 27 (5). https://doi.org/10.1016/j.ymthe.2019.02.013.

Huang, Qin, Ken Y. Chan, Isabelle G. Tobey, Yujia Alina Chan, Tim Poterba, Christine L. Boutros, Alejandro B. Balazs, et al. 2019. "Delivering Genes across the Blood-Brain Barrier: LY6A, a Novel Cellular Receptor for AAV-PHP.B Capsids." *PLoS ONE* 14 (11): 1–17. https://doi.org/10.1371/journal.pone.0225206.

Huang, Samuel J, Kei Adachi, Zhen Song, Sergio Ojeda, Greg Dissen, and Hiroyuki Nakai. 2017. "Highly Efficient Brain Neuron Targeting and Hepatocyte Detargeting with a Novel AAV9-PHP.B Capsid Harboring a Liver-Detargeting Mutation." *Molecular Therapy* 25 (5S1): 257–58.

Hudry, Eloise, Eva Andres-Mateos, Eli P. Lerner, Adrienn Volak, Olivia Cohen, Bradley T. Hyman, Casey A. Maguire, and Luk H. Vandenberghe. 2018. "Efficient Gene Transfer to the Central Nervous System by Single-Stranded Anc80L65." *Molecular Therapy - Methods and Clinical Development* 10 (September): 197–209. https://doi.org/10.1016/j.omtm.2018.07.006.

Im, Dong Soo, and Nicholas Muzyczka. 1990. "The AAV Origin Binding Protein Rep68 Is an ATP-Dependent Site-Specific Endonuclease with DNA Helicase Activity." *Cell* 61 (3): 447–57. https://doi.org/10.1016/0092-8674(90)90526-K.

Jacobson, S. G., T. S. Aleman, A. V. Cideciyan, A. Sumaroka, S. B. Schwartz, E. A. M. Windsor, E. I. Traboulsi, et al. 2005. "Identifying Photoreceptors in Blind Eyes Caused by RPE65 Mutations: Prerequisite for Human Gene Therapy Success." *Proceedings of the National Academy of Sciences* 102 (17): 6177–82. https://doi.org/10.1073/pnas.0500646102.

Johnson, Jarrod S., and R. Jude Samulski. 2009. "Enhancement of Adeno-Associated Virus Infection by Mobilizing Capsids into and Out of the Nucleolus." *Journal of Virology* 83 (6): 2632–44. https://doi.org/10.1128/jvi.02309-08.

Jüttner, Josephine, Arnold Szabo, Brigitte Gross-Scherf, Rei K. Morikawa, Santiago B. Rompani, Peter Hantz, Tamas Szikra, et al. 2019. "Targeting Neuronal and Glial Cell Types with Synthetic Promoter AAVs in Mice, Non-Human Primates and Humans." *Nature Neuroscience* 22 (8): 1345–56. https://doi.org/10.1038/s41593-019-0431-2.

Karlin, David G. 2020. "Sequence Properties of the MAAP Protein and of the VP1 Capsid Protein of Adeno-Associated Viruses." *Preprints* 3 (February): 1–20. https://doi.org/10.20944/preprints202002.0234.v1.

Kashiwakura, Yuji, Kenji Tamayose, Kazuhisa Iwabuchi, Yukihiko Hirai, Takashi Shimada, Kunio Matsumoto, Toshikazu Nakamura, Masami Watanabe, Kazuo Oshimi, and Hiroyuki Daida. 2005. "Hepatocyte Growth Factor Receptor Is a Coreceptor for Adeno-Associated Virus Type 2 Infection." *Journal of Virology* 79 (1): 609–14. https://doi.org/10.1128/jvi.79.1.609-614.2005.

Kay, Mark A., Catherine S. Manno, Margaret V. Ragni, Peter J. Larson, Linda B. Couto, Alan McClelland, Bertil Glader, et al. 2000. "Evidence for Gene Transfer and Expression of Factor IX in Haemophilia B Patients Treated with an AAV Vector." *Nature Genetics* 24 (3): 257–61. https://doi.org/10.1038/73464.

Kelkar, Samir, Bishnu P. De, Guangping Gao, James M. Wilson, Ronald G. Crystal, and Philip L. Leopold. 2006. "A Common Mechanism for Cytoplasmic Dynein-Dependent Microtubule Binding Shared among Adeno-Associated Virus and Adenovirus Serotypes." *Journal of Virology* 80 (15): 7781–85. https://doi.org/10.1128/jvi.00481-06.

Kern, A., K. Schmidt, C. Leder, O. J. Müller, C. E. Wobus, K. Bettinger, C. W. Von der Lieth, J. A. King, and J. A. Kleinschmidt. 2003. "Identification of a Heparin-Binding Motif on Adeno-Associated Virus Type 2 Capsids." *Journal of Virology* 77 (20): 11072–81. https://doi.org/10.1128/jvi.77.20.11072-11081.2003.

Keswani, Sundeep G., Swathi Balaji, Louis Le, Alice Leung, Anna B. Katz, Foong Yen Lim, Mounira Habli, Helen N. Jones, James M. Wilson, and Timothy M. Crombleholme. 2012. "Pseudotyped AAV Vector-Mediated Gene Transfer in a Human Fetal Trachea Xenograft Model: Implications for in Utero Gene Therapy for Cystic Fibrosis." *PLoS ONE* 7 (8): 1–10. https://doi.org/10.1371/journal.pone.0043633.

Kim, Jihoon, Bon Kyoung Koo, and Juergen A. Knoblich. 2020. "Human Organoids: Model Systems for Human Biology and Medicine." *Nature Reviews Molecular Cell Biology* 21 (10): 571–84. https://doi.org/10.1038/s41580-020-0259-3.

King, Jason A, Ralf Dubielzig, and Dirk Grimm. 2001. "DNA Helicase-Mediated Packaging of Adeno-Associated Virus Type 2 Genomes into Preformed Capsids." *European Molecular Biology Organization* 20 (12).

Klein, Ronald L., Robert D. Dayton, Jason B. Tatom, Karen M. Henderson, and Phillip P. Henning. 2008. "AAV8, 9, Rh10, Rh43 Vector Gene Transfer in the Rat Brain: Effects of Serotype, Promoter and Purification Method." *Molecular Therapy* 16 (1): 89–96. https://doi.org/10.1038/sj.mt.6300331.

Koerber, James T., Jae Hyung Jang, and David V. Schaffer. 2008. "DNA Shuffling of Adeno-Associated Virus Yields Functionally Diverse Viral Progeny." *Molecular Therapy* 16 (10): 1703–9. https://doi.org/10.1038/mt.2008.167.

Kojima, Karin, Takeshi Nakajima, Naoyuki Taga, Akihiko Miyauchi, Mitsuhiro Kato, Ayumi Matsumoto, Takahiro Ikeda, et al. 2019. "Gene Therapy Improves Motor and Mental Function of Aromatic L-Amino Acid Decarboxylase Deficiency." *Brain* 142 (2): 322–33. https://doi.org/10.1093/brain/awy331.

Körbelin, Jakob, and Martin Trepel. 2017. "How to Successfully Screen Random Adeno-Associated Virus Display Peptide Libraries in Vivo." *Human Gene Therapy Methods* 28 (3): 109–23. https://doi.org/10.1089/hgtb.2016.177.

Kotchey, Nicole M., Kei Adachi, Maliha Zahid, Katsuya Inagaki, Rakshita Charan, Robert S. Parker, and Hiroyuki Nakai. 2011. "A Potential Role of Distinctively Delayed Blood Clearance of Recombinant Adeno-Associated Virus Serotype 9 in Robust Cardiac Transduction." *Molecular Therapy* 19 (6): 1079–89. https://doi.org/10.1038/mt.2011.3.

Kumar, Sripriya Ravindra, Timothy F. Miles, Xinhong Chen, David Brown, Tatyana Dobreva, Qin Huang, Xiaozhe Ding, et al. 2020. "Multiplexed Cre-Dependent Selection Yields Systemic AAVs for Targeting Distinct Brain Cell Types." *Nature Methods* 17 (5): 541–50. https://doi.org/10.1038/s41592-020-0799-7.

Kuzmin, Dmitry A., Maria V. Shutova, Natalie R. Johnston, Owen P. Smith, Vasily V. Fedorin, Yury S. Kukushkin, Johannes C.M. van der Loo, and Elaine C. Johnstone. 2021. "The Clinical Landscape for AAV Gene Therapies." *Nature Reviews. Drug Discovery* 20 (3): 173–74. https://doi.org/10.1038/d41573-021-00017-7.

Lang, Jonathan F., Sushila A. Toulmin, Kasey L. Brida, Laurence C. Eisenlohr, and Beverly L. Davidson. 2019. "Standard Screening Methods Underreport AAV-Mediated Transduction and Gene Editing." *Nature Communications* 10 (1): 1–10. https://doi.org/10.1038/s41467-019-11321-7.

Laughlin, Catherine A., Jon Duri Tratschin, Helen Coon, and Barrie J. Carter. 1983. "Cloning of Infectious Adeno-Associated Virus Genomes in Bacterial Plasmids." *Gene* 23 (1): 65–73. https://doi.org/10.1016/0378-1119(83)90217-2.

Lerch, Thomas F., and Michael S. Chapman. 2012. "Identification of the Heparin Binding Site on Adeno-Associated Virus Serotype 3B (AAV-3B)." *Virology* 423 (1): 6–13. https://doi.org/10.1016/j.virol.2011.10.007.

Li, Wuping, Aravind Asokan, Zhijian Wu, Terry Van Dyke, Nina DiPrimio, Johnson S. Jarrod, Lakshmanan Govindaswamy, et al. 2008. "Engineering and Selection of

Shuffled AAV Genomes: A New Strategy for Producing Targeted Biological Nanoparticles." *Molecular Therapy* 16 (7): 1252–60. https://doi.org/10.1038/mt.2008.100.

Li, Yan, Xingyi Ge, Chung Chau Hon, Huajun Zhang, Peng Zhou, Yunzhi Zhang, Yi Wu, Lin Fa Wang, and Zhengli Shi. 2010. "Prevalence and Genetic Diversity of Adeno-Associated Viruses in Bats from China." *Journal of General Virology* 91 (10): 2601–9. https://doi.org/10.1099/vir.0.020032-0.

Liguore, William A., Jacqueline S. Domire, Dana Button, Yun Wang, Brett D. Dufour, Sathya Srinivasan, and Jodi L. McBride. 2019. "AAV-PHP.B Administration Results in a Differential Pattern of CNS Biodistribution in Non-Human Primates Compared with Mice." *Molecular Therapy* 27 (11): 2018–37. https://doi.org/10.1016/j.ymthe.2019.07.017.

Lisowski, Leszek, Allison P. Dane, Kirk Chu, Yue Zhang, Sharon C. Cunningham, Elizabeth M. Wilson, Sean Nygaard, Markus Grompe, Ian E. Alexander, and Mark A. Kay. 2014. "Selection and Evaluation of Clinically Relevant AAV Variants in a Xenograft Liver Model." *Nature* 506 (7488): 382–86. https://doi.org/10.1038/nature12875.

Lochrie, Michael A., Gwen P. Tatsuno, Brian Christie, Jennifer Wellman McDonnell, Shangzhen Zhou, Richard Surosky, Glenn F. Pierce, and Peter Colosi. 2006. "Mutations on the External Surfaces of Adeno-AssociatedVirus Type 2 Capsids That Affect Transduction AndNeutralization." *Journal of Virology* 80 (2): 821–34. https://doi.org/10.1128/jvi.80.2.821-834.2006.

Lusby, E, K H Fife, and K I Berns. 1980. "Nucleotide Sequence of the Inverted Terminal Repetition in Adeno-Associated Virus DNA." *Journal of Virology* 34 (2): 402–9. https://doi.org/10.1128/jvi.34.2.402-409.1980.

Maes, Margaret, Gloria Colombo, Rouven Schulz, and Sandra Siegert. 2019. "Targeting Microglia with Lentivirus and AAV: Recent Advances and Remaining Challenges." *Neuroscience Letters* 707 (May): 134310. https://doi.org/10.1016/j.neulet.2019.134310.

Maguire, Albert M., Francesca Simonelli, Eric A. Pierce, Edward N. Pugh, Federico Mingozzi, Jeannette Bennicelli, Sandro Banfi, et al. 2008. "Safety and Efficacy of Gene Transfer for Leber's Congenital Amaurosis." *New England Journal of Medicine* 358 (21): 2240–48. https://doi.org/10.1056/NEJMoa0802315.

Maheshri, Narendra, James T. Koerber, Brian K. Kaspar, and David V. Schaffer. 2006. "Directed Evolution of Adeno-Associated Virus Yields Enhanced Gene Delivery Vectors." *Nature Biotechnology* 24 (2): 198–204. https://doi.org/10.1038/nbt1182.

Manno, Catherine S., Valder R. Arruda, Glenn F. Pierce, Bertil Glader, Margaret Ragni, John Rasko, Margareth C. Ozelo, et al. 2006. "Successful Transduction of Liver in Hemophilia by AAV-Factor IX and Limitations Imposed by the Host Immune Response." *Nature Medicine* 12 (3): 342–47. https://doi.org/10.1038/nm1358.

Mannucci, Pier M, and Edward GD Tuddenham. 2001. "The Hemophilias — From Royal Genes to Gene Therapy." *New England Journal of Medicine* 344 (23): 1773–79.

https://doi.org/10.1056/nejm200106073442307.

Marsic, Damien, Héctor R. Méndez-Gómez, and Sergei Zolotukhin. 2015. "High-Accuracy Biodistribution Analysis of Adeno-Associated Virus Variants by Double Barcode Sequencing." *Molecular Therapy - Methods and Clinical Development* 2 (September): 15041. https://doi.org/10.1038/mtm.2015.41.

Martino, Ashley T., Masataka Suzuki, David M. Markusic, Irene Zolotukhin, Renee C. Ryals, Babak Moghimi, Hildegund C.J. Ertl, Daniel A. Muruve, Brendan Lee, and Roland W. Herzog. 2011. "The Genome of Self-Complementary Adeno-Associated Viral Vectors Increases Toll-like Receptor 9-Dependent Innate Immune Responses in the Liver." *Blood* 117 (24): 6459–68. https://doi.org/10.1182/blood-2010-10-314518.

Mathiesen, Sophie N., Jasmine L. Lock, Lucia Schoderboeck, Wickliffe C. Abraham, and Stephanie M. Hughes. 2020. "CNS Transduction Benefits of AAV-PHP.EB over AAV9 Are Dependent on Administration Route and Mouse Strain." *Molecular Therapy - Methods & Clinical Development* 19 (December): 447–58. https://doi.org/10.1016/j.omtm.2020.10.011.

Matsushita, T., S. Elliger, C. Elliger, G. Podsakoff, L. Villarreal, G. J. Kurtzman, Y. Iwaki, and P. Colosi. 1998. "Adeno-Associated Virus Vectors Can Be Efficiently Produced without Helper Virus." *Gene Therapy* 5 (7): 938–45. https://doi.org/10.1038/sj.gt.3300680.

Matsuzaki, Yasunori, Ayumu Konno, Ryuta Mochizuki, Yoichiro Shinohara, Keisuke Nitta, Yukihiro Okada, and Hirokazu Hirai. 2018. "Intravenous Administration of the Adeno-Associated Virus-PHP.B Capsid Fails to Upregulate Transduction Efficiency in the Marmoset Brain." *Neuroscience Letters* 665 (June 2017): 182–88. https://doi.org/10.1016/j.neulet.2017.11.049.

Matsuzaki, Yasunori, Masami Tanaka, Sachiko Hakoda, Tatsuki Masuda, Ryota Miyata, Ayumu Konno, and Hirokazu Hirai. 2019. "Neurotropic Properties of AAV-PHP.B Are Shared among Diverse Inbred Strains of Mice." *Molecular Therapy* 27 (4): 700–704. https://doi.org/10.1016/j.ymthe.2019.02.016.

McCarty, DM, H. Fu, P. E. Monahan, C. E. Toulson, P. Naik, and R. J. Samulski. 2003. "Adeno-Associated Virus Terminal Repeat (TR) Mutant Generates Self-Complementary Vectors to Overcome the Rate-Limiting Step to Transduction in Vivo." *Gene Therapy* 10 (26): 2112–18. https://doi.org/10.1038/sj.gt.3302134.

McCarty, Douglas M., Samuel M. Young, and R. Jude Samulski. 2004. "Integration of Adeno-Associated Virus (AAV) and Recombinant AAV Vectors." *Annual Review of Genetics* 38: 819–45. https://doi.org/10.1146/annurev.genet.37.110801.143717.

McLaughlin, S K, P Collis, P L Hermonat, and N Muzyczka. 1988. "Adeno-Associated Virus General Transduction Vectors: Analysis of Proviral Structures." *Journal of Virology* 62 (6): 1963–73. https://doi.org/10.1128/jvi.62.6.1963-1973.1988.

Mendell, Jerry R., Samiah Al-Zaidy, Richard Shell, W. Dave Arnold, Louise R. Rodino-Klapac, Thomas W. Prior, Linda Lowes, et al. 2017. "Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy." *New England Journal of*

Medicine 377 (18): 1713–22. https://doi.org/10.1056/NEJMoa1706198.

Michelfelder, Stefan, Johannes Kohlschütter, Alexandra Skorupa, Sabrina Pfennings, Oliver Müller, Jürgen A. Kleinschmidt, and Martin Trepel. 2009. "Successful Expansion but Not Complete Restriction of Tropism of Adeno-Associated Virus by in Vivo Biopanning of Random Virus Display Peptide Libraries." *PLoS ONE* 4 (4). https://doi.org/10.1371/journal.pone.0005122.

Mietzsch, M., F. Broecker, A. Reinhardt, P. H. Seeberger, R. Heilbronn, and M. J. Imperiale. 2014. "Differential Adeno-Associated Virus Serotype-Specific Interaction Patterns with Synthetic Heparins and Other Glycans." *Journal of Virology* 88 (5): 2991–3003. https://doi.org/10.1128/jvi.03371-13.

Mingozzi, Federico, Marcela V. Maus, Daniel J. Hui, Denise E. Sabatino, Samuel L. Murphy, John E.J. Rasko, Margaret V. Ragni, et al. 2007. "CD18+ T-Cell Responses to Adeno-Associated Virus Capsid in Humans." *Nature Medicine* 13 (4): 419–22. https://doi.org/10.1038/nm1549.

Mittermeyer, Gabriele, Chadwick W. Christine, Kathryn H. Rosenbluth, Suzanne L. Baker, Philip Starr, Paul Larson, Paul L. Kaplan, John Forsayeth, Michael J. Aminoff, and Krystof S. Bankiewicz. 2011. "Long-Term Evaluation of a Phase 1 Study of AADC Gene Therapy for Parkinson's Disease." *Human Gene Therapy* 23 (4): 377–81. https://doi.org/10.1089/hum.2011.220.

Mount, Jane D., Roland W. Herzog, D. Michael Tillson, Susan A. Goodman, Nancy Robinson, Mark L. McCleland, Dwight Bellinger, et al. 2002. "Sustained Phenotypic Correction of Hemophilia B Dogs with a Factor IX Null Mutation by Liver-Directed Gene Therapy." *Blood* 99 (8): 2670–76. https://doi.org/10.1182/blood.V99.8.2670.

Mouw, Matthew B., and David J. Pintel. 2000. "Adeno-Associated Virus RNAs Appear in a Temporal Order and Their Splicing Is Stimulated during Coinfection with Adenovirus." *Journal of Virology* 74 (21): 9878–88. https://doi.org/10.1128/jvi.74.21.9878-9888.2000.

Müller, Oliver J., Felix Kaul, Matthew D. Weitzman, Renata Pasqualini, Wadih Arap, Jürgen A. Kleinschmidt, and Martin Trepel. 2003. "Random Peptide Libraries Displayed on Adeno-Associated Virus to Select for Targeted Gene Therapy Vectors." *Nature Biotechnology* 21 (9): 1040–46. https://doi.org/10.1038/nbt856.

Nakai, Hiroyuki, Stephen R. Yant, Theresa A. Storm, Sally Fuess, Leonard Meuse, and Mark A. Kay. 2001. "Extrachromosomal Recombinant Adeno-Associated Virus Vector Genomes Are Primarily Responsible for Stable Liver Transduction In Vivo." *Journal of Virology* 75 (15): 6969–76. https://doi.org/10.1128/jvi.75.15.6969-6976.2001.

Nam, Hyun-Joo, Michael Douglas Lane, Eric Padron, Brittney Gurda, Robert McKenna, Erik Kohlbrenner, George Aslanidi, et al. 2007. "Structure of Adeno-Associated Virus Serotype 8, a Gene Therapy Vector." *Journal of Virology* 81 (22): 12260–71. https://doi.org/10.1128/jvi.01304-07.

Nash, Kevin, Weijun Chen, and Nicholas Muzyczka. 2008. "Complete In Vitro

Reconstitution of Adeno-Associated Virus DNA Replication Requires the Minichromosome Maintenance Complex Proteins." *Journal of Virology* 82 (3): 1458–64. https://doi.org/10.1128/jvi.01968-07.

Nathwani, Amit C., Ulreke M. Reiss, Edward G.D. Tuddenham, Cecilia Rosales, Pratima Chowdary, Jenny McIntosh, Marco Della Peruta, et al. 2014. "Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B." *New England Journal of Medicine* 371 (21): 1994–2004. https://doi.org/10.1056/nejmoa1407309.

Nathwani, Amit C., Edward G.D. Tuddenham, Savita Rangarajan, Cecilia Rosales, Jenny McIntosh, David C. Linch, Pratima Chowdary, et al. 2011. "Adenovirus-Associated Virus Vector–Mediated Gene Transfer in Hemophilia B." *New England Journal of Medicine* 365 (25): 2357–65. https://doi.org/10.1056/nejmoa1108046.

Nault, Jean Charles, Shalini Datta, Sandrine Imbeaud, Andrea Franconi, Maxime Mallet, Gabrielle Couchy, Eric Letouzé, et al. 2015. "Recurrent AAV2-Related Insertional Mutagenesis in Human Hepatocellular Carcinomas." *Nature Genetics* 47 (10): 1187–93. https://doi.org/10.1038/ng.3389.

Nonnenmacher, Mathieu. 2021. "RNA-Driven Evolution of AAV Capsid Libraries Identifies Variants with High Transduction Efficiency in Non-Human Primate Central Nervous System." In *24th Annual Meeting of the American Society of Gene & Cell Therapy*.

Nonnenmacher, Mathieu, H van Bakel, R J Hajjar, and T Weber. 2015. "High Capsid-Genome Correlation Facilitates Creation of AAV Libraries for Directed Evolution." *Mol Ther* 23 (4): 675–82. https://doi.org/10.1038/mt.2015.3.

Nonnenmacher, Mathieu, Wei Wang, Matthew A. Child, Xiao-Qin Ren, Carol Huang, Amy Zhen Ren, Jenna Tocci, et al. 2020. "Rapid Evolution of Blood-Brain Barrier-Penetrating AAV Capsids by RNA-Driven Biopanning." *Molecular Therapy - Methods & Clinical Development* 20 (March): 366–78. https://doi.org/10.1016/j.omtm.2020.12.006.

Ogden, Pierce J., Eric D. Kelsic, Sam Sinai, and George M. Church. 2019. "Comprehensive AAV Capsid Fitness Landscape Reveals a Viral Gene and Enables Machine-Guided Design." *Science* 366 (6469): 1139–43. https://doi.org/10.1126/science.aaw2900.

Ojala, David S, Sabrina Sun, Jorge L Santiago-ortiz, Mikhail G Shapiro, Philip A Romero, and David V Schaffer. 2018. "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* 26 (1): 304–19. https://doi.org/10.1016/j.ymthe.2017.09.006.

Pekrun, Katja, Gustavo De Alencastro, Qing Jun Luo, Jun Liu, Youngjin Kim, Sean Nygaard, Feorillo Galivo, et al. 2019. "Using a Barcoded AAV Capsid Library to Select for Novel Clinically Relevant Gene Therapy Vectors." *JCI Insight* 4 (22). https://doi.org/10.1101/683672.

Perabo, Luca, Hildegard Bu, David M Kofler, Martin U Ried, Anne Girod, Clemens M Wendtner, and Michael Hallek. 2003. "In Vitro Selection of Viral Vectors with

Modified Tropism: The Adeno-Associated Virus Display." *Molecular Therapy* 8 (1): 151–57. https://doi.org/10.1016/S1525-0016(03)00123-0.

Perabo, Luca, Jan Endell, Susan King, Kerstin Lux, Daniela Goldnau, Michael Hallek, and Hildegard Büning. 2006. "Combinatorial Engineering of a Gene Therapy Vector: Directed Evolution of Adeno-Associated Virus." *Journal of Gene Medicine* 8 (2): 155–62. https://doi.org/10.1002/jgm.849.

Pereira, D J, DM McCarty, and N Muzyczka. 1997. "The Adeno-Associated Virus (AAV) Rep Protein Acts as Both a Repressor and an Activator to Regulate AAV Transcription during a Productive Infection." *Journal of Virology* 71 (2): 1079–88. https://doi.org/10.1128/jvi.71.2.1079-1088.1997.

Pierce, E.A., and J. Bennett. 2015. "Safety and Efficacy of RPE65 Gene Therapy Trials: Safety and Efficacy." *Cold Spring Harbor Perspectives in Medicine*, 380–82.

Pincus, Alexandra B., Sweta Adhikary, Katherine M. Lebold, Allison D. Fryer, and David B. Jacoby. 2020. "Optogenetic Control of Airway Cholinergic Neurons in Vivo." *American Journal of Respiratory Cell and Molecular Biology* 62 (4): 423–29. https://doi.org/10.1165/rcmb.2019-0378MA.

Portales-Casamar, Elodie, Douglas J. Swanson, Li Liu, Charles N. De Leeuw, Kathleen G. Banks, Shannan J. Ho Sui, Debra L. Fulton, et al. 2010. "A Regulatory Toolbox of MiniPromoters to Drive Selective Expression in the Brain." *Proceedings of the National Academy of Sciences of the United States of America* 107 (38): 16589–94. https://doi.org/10.1073/pnas.1009158107.

Powell, Sara K., R. Jude Samulski, and Thomas J. McCown. 2020. "AAV Capsid-Promoter Interactions Determine CNS Cell-Selective Gene Expression In Vivo." *Molecular Therapy* 28 (5): 1–8. https://doi.org/10.1016/j.ymthe.2020.03.007.

Powers, John M., Xiao Lan Chang, Zhen Song, and Hiroyuki Nakai. 2018. "A Quantitative Dot Blot Assay for AAV Titration and Its Use for Functional Assessment of the Adeno-Associated Virus Assembly-Activating Proteins." *Journal of Visualized Experiments* 2018 (136): 1–14. https://doi.org/10.3791/56766.

Prasad, Konkal Matt Renuka, Changhong Zhou, and James P. Trempe. 1997. "Characterization of the Rep78/Adeno-Associated Virus Complex." *Virology* 229 (1): 183–92. https://doi.org/10.1006/viro.1996.8431.

Pulicherla, Nagesh, Shen Shen, Swati Yadav, Kari Debbink, Lakshmanan Govindasamy, Mavis Agbandje-Mckenna, and Aravind Asokan. 2011. "Engineering Liver-Detargeted AAV9 Vectors for Cardiac and Musculoskeletal Gene Transfer." *Molecular Therapy* 19 (6): 1070–78. https://doi.org/10.1038/mt.2011.22.

Qing, Keyun, Cathryn Mah, Jonathan Hansen, Shangzhen Zhou, Varavani Dwarki, and Arun Srivastava. 1999. "Human Fibroblast Growth Factor Receptor 1 Is a Co-Receptor for Infection by Adeno-Associated Virus 2." *Nature Medicine* 5 (1): 71–77. https://doi.org/10.1038/4758.

Rincon, Melvin Y., Filip De Vin, Sandra I. Duqué, Shelly Fripont, Stephanie A. Castaldo, Jessica Bouhuijzen-Wenger, and Matthew G. Holt. 2018. "Widespread Transduction of Astrocytes and Neurons in the Mouse Central Nervous System after Systemic

Delivery of a Self-Complementary AAV-PHP.B Vector." *Gene Therapy* 25 (2): 83–92. https://doi.org/10.1038/s41434-018-0005-z.

Ross, Colin J.D., Jaap Twisk, Andrew C. Bakker, Fudan Miao, Dennis Verbart, Jaap Rip, Tamara Godbey, et al. 2006. "Correction of Feline Lipoprotein Lipase Deficiency with Adeno-Associated Virus Serotype 1-Mediated Gene Transfer of the Lipoprotein Lipase S447X Beneficial Mutation." *Human Gene Therapy* 17 (5): 487–99. https://doi.org/10.1089/hum.2006.17.487.

Russell, Stephen, Jean Bennett, Jennifer A Wellman, Daniel C Chung, Zi-fan Yu, Amy Tillman, Janet Wittes, et al. 2017. "Efficacy and Safety of Voretigene Neparvovec (AAV2-HRPE65v2) in Patients with RPE65-Mediated Inherited Retinal Dystrophy: A Randomised, Controlled, Open-Label, Phase 3 Trial." *The Lancet* 390 (10097): 849–60. https://doi.org/10.1016/S0140-6736(17)31868-8.

Salganik, M., F. Aydemir, H.-J. Nam, R. McKenna, M. Agbandje-McKenna, and N. Muzyczka. 2014. "Adeno-Associated Virus Capsid Proteins May Play a Role in Transcription and Second-Strand Synthesis of Recombinant Genomes." *Journal of Virology* 88 (2): 1071–79. https://doi.org/10.1128/jvi.02093-13.

Samaranch, Lluis, Waldy San Sebastian, Adrian P Kells, Ernesto A Salegio, Gregory Heller, John R Bringas, Philip Pivirotto, Stephen Dearmond, John Forsayeth, and Krystof S Bankiewicz. 2014. "AAV9-Mediated Expression of a Non-Self Protein in Nonhuman Primate Central Nervous System Triggers Widespread Neuroinflammation Driven by Antigen-Presenting Cell Transduction." *Molecular Therapy* 22 (2): 329–37. https://doi.org/10.1038/mt.2013.266.

Samulski, R. J., K. I. Berns, M. Tan, and N. Muzyczka. 1982. "Cloning of Adeno-Associated Virus into PBR322: Rescue of Intact Virus from the Recombinant Plasmid in Human Cells." *Proceedings of the National Academy of Sciences* 79 (6): 2077–81. https://doi.org/10.1073/pnas.79.6.2077.

Santiago-Ortiz, J., D. S. Ojala, O. Westesson, J. R. Weinstein, S. Y. Wong, A. Steinsapir, S. Kumar, I. Holmes, and D. V. Schaffer. 2015. "AAV Ancestral Reconstruction Library Enables Selection of Broadly Infectious Viral Variants." *Gene Therapy* 22 (12): 934–46. https://doi.org/10.1038/gt.2015.74.

Schmidt, Michael, Hisako Katano, Ioannis Bossis, and John A. Chiorini. 2004. "Cloning and Characterization of a Bovine Adeno-Associated Virus." *Journal of Virology* 78 (12): 6509–16. https://doi.org/10.1128/jvi.78.12.6509-6516.2004.

Schmit, Pauline F., Simon Pacouret, Eric Zinn, Elizabeth Telford, Fotini Nicolaou, Frédéric Broucque, Eva Andres-Mateos, et al. 2020. "Cross-Packaging and Capsid Mosaic Formation in Multiplexed AAV Libraries." *Molecular Therapy - Methods and Clinical Development* 17 (June): 107–21. https://doi.org/10.1016/j.omtm.2019.11.014.

Schoch, Susanne, Giuseppe Cibelli, and Gerald Thiel. 1996. "Neuron-Specific Gene Expression of Synapsin I: Major Role of a Negative Regulatory Mechanism." *Journal of Biological Chemistry* 271 (6): 3317–23. https://doi.org/10.1074/jbc.271.6.3317.

Schuster, Daniel J., Jaclyn A. Dykstra, Maureen S. Riedl, Kelley F. Kitto, Lalitha R. Belur, R. Scott McIvor, Robert P. Elde, Carolyn A. Fairbanks, and Lucy Vulchanova. 2014. "Biodistribution of Adeno-Associated Virus Serotype 9 (AAV9) Vector after Intrathecal and Intravenous Delivery in Mouse." *Frontiers in Neuroanatomy* 8 (JUN): 1–14. https://doi.org/10.3389/fnana.2014.00042.

Sheridan, Cormac. 2019. "To Win at Gene Therapy, Companies Pick Viruses with Production Credentials." *Nature Biotechnology* 37 (1): 5–6. https://doi.org/10.1038/nbt0119-5.

Sibbald, B. 2001. "Death but One Unintended Consequence of Gene-Therapy Trial." *CMAJ : Canadian Medical Association Journal = Journal de l"Association Medicale Canadienne* 164 (11): 1612.

Sonntag, Florian, Svenja Bleker, Barbara Leuchs, Roger Fischer, and Jürgen A. Kleinschmidt. 2006. "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." *Journal of Virology* 80 (22): 11040–54. https://doi.org/10.1128/jvi.01056-06.

Sonntag, Florian, K. Kother, K. Schmidt, M. Weghofer, C. Raupp, K. Nieto, A. Kuck, et al. 2011. "The Assembly-Activating Protein Promotes Capsid Assembly of Different Adeno-Associated Virus Serotypes." *Journal of Virology* 85 (23): 12686–97. https://doi.org/10.1128/jvi.05359-11.

Soong, N. W., L. Nomura, K. Pekrun, M. Reed, L. Sheppard, G. Dawes, and W. P.C. Stemmer. 2000. "Molecular Breeding of Viruses." *Nature Genetics* 25 (4): 436–39. https://doi.org/10.1038/78132.

Srivastava, Arun. 2016. "In Vivo Tissue-Tropism of Adeno-Associated Viral Vectors." *Current Opinion in Virology* 21: 75–80. https://doi.org/10.1016/j.coviro.2016.08.003.

———. 2020. "AAV Vectors: Are They Safe?" *Human Gene Therapy* 31 (13–14): 697–99. https://doi.org/10.1089/hum.2020.187.

Stutika, Catrin, Andreas Gogol-Döring, Laura Botschen, Mario Mietzsch, Stefan Weger, Mirjam Feldkamp, Wei Chen, and Regine Heilbronn. 2016. "A Comprehensive RNA Sequencing Analysis of the Adeno-Associated Virus (AAV) Type 2 Transcriptome Reveals Novel AAV Transcripts, Splice Variants, and Derived Proteins." *Journal of Virology* 90 (3): 1278–89. https://doi.org/10.1128/jvi.02750-15.

Summerford, Candace, J Effrey S B Artlett, and Richard Jude Samulski. 1999. "AVβ5 Integrin: A Co-Receptor for Adeno-Associated Virus Type 2 Infection" 5 (1): 1–5.

Summerford, Candace, and Richard Jude Samulski. 1998. "Membrane-Associated Heparan Sulfate Proteoglycan Is a Receptor for Adeno-Associated Virus Type 2 Virions." *Journal of Virology* 72 (2): 1438–45. https://doi.org/10.1128/jvi.72.2.1438-1445.1998.

Sweigard, J. Harry, Siobhan M. Cashman, and Rajendra Kumar-Singh. 2010. "Adenovirus Vectors Targeting Distinct Cell Types in the Retina." *Investigative*

Ophthalmology and Visual Science 51 (4): 2219–28. https://doi.org/10.1167/iovs.09-4367.

UniQure. 2021. "UniQure Announces Findings from Reported Case of Hepatocellular Carcinoma (HCC) in Hemophilia B Gene Therapy Program." UniQure Press Releases. 2021. https://tools.eurolandir.com/tools/Pressreleases/GetPressRelease/?ID=3890956&lang=en-GB&companycode=nl-qure&v=.

Vandenberghe, L. H., E. Breous, H. J. Nam, G. Gao, R. Xiao, A. Sandhu, J. Johnston, Z. Debyser, M. Agbandje-Mckenna, and J. M. Wilson. 2009. "Naturally Occurring Singleton Residues in AAV Capsid Impact Vector Performance and Illustrate Structural Constraints." *Gene Therapy* 16 (12): 1416–28. https://doi.org/10.1038/gt.2009.101.

Vasconcelos, Huber M., Brandon J. Lujan, Mark E. Pennesi, Paul Yang, and Andreas K. Lauer. 2020. "Intraoperative Optical Coherence Tomographic Findings in Patients Undergoing Subretinal Gene Therapy Surgery." *International Journal of Retina and Vitreous* 6 (1): 1–10. https://doi.org/10.1186/s40942-020-00216-1.

Vercauteren, Koen, Brad E. Hoffman, Irene Zolotukhin, Geoffrey D. Keeler, Jing W. Xiao, Etiena Basner-Tschakarjan, Katherine A. High, et al. 2016. "Superior in Vivo Transduction of Human Hepatocytes Using Engineered AAV3 Capsid." *Molecular Therapy* 24 (6): 1042–49. https://doi.org/10.1038/mt.2016.61.

Veske, Andres, Sven Erik G. Nilsson, Kristina Narfström, and Andreas Gal. 1999. "Retinal Dystrophy of Swedish Briard/Briard-Beagle Dogs Is Due to a 4-Bp Deletion in RPE65." *Genomics* 57 (1): 57–61. https://doi.org/10.1006/geno.1999.5754.

Wang, Dan, Li Zhong, Mengxin Li, Jia Li, Karen Tran, Lingzhi Ren, Ran He, et al. 2018. "Adeno-Associated Virus Neutralizing Antibodies in Large Animals and Their Impact on Brain Intraparenchymal Gene Transfer." *Molecular Therapy - Methods and Clinical Development* 11 (December): 65–72. https://doi.org/10.1016/j.omtm.2018.09.003.

Wang, Sean K., Sylvain W. Lapan, Christin M. Hong, Tyler B. Krause, and Constance L. Cepko. 2020. "In Situ Detection of Adeno-Associated Viral Vector Genomes with SABER-FISH." *Molecular Therapy - Methods and Clinical Development* 19 (December): 376–86. https://doi.org/10.1016/j.omtm.2020.10.003.

Wang, X S, S Ponnazhagan, and A Srivastava. 1996. "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." *Journal of Virology* 70 (3): 1668–77. https://doi.org/10.1128/jvi.70.3.1668-1677.1996.

Wang, Xu Shan, Selvarangan Ponnazhagan, and Arun Srivastava. 1995. "Rescue and Replication Signals of the Adeno-Associated Virus 2 Genome." *Journal of Molecular Biology* 250 (5): 573–80. https://doi.org/10.1006/jmbi.1995.0398.

Weger, S, A Wistuba, D Grimm, and J A Kleinschmidt. 1997. "Control of Adeno-

Associated Virus Type 2 Cap Gene Expression: Relative Influence of Helper Virus, Terminal Repeats, and Rep Proteins." *Journal of Virology* 71 (11): 8437–47. https://doi.org/10.1128/jvi.71.11.8437-8447.1997.

Weinmann, Jonas, Sabrina Weis, Josefine Sippel, Warut Tulalamba, Anca Remes, Jihad El Andari, Anne Kathrin Herrmann, et al. 2020. "Identification of a Myotropic AAV by Massively Parallel in Vivo Evaluation of Barcoded Capsid Variants." *Nature Communications* 11 (1): 1–12. https://doi.org/10.1038/s41467-020-19230-w.

Wilson, James M., and Terence R. Flotte. 2020. "Moving Forward after Two Deaths in a Gene Therapy Trial of Myotubular Myopathy." *Human Gene Therapy* 31 (13–14): 695–96. https://doi.org/10.1089/hum.2020.182.

Wu, Zhijian, Aravind Asokan, Joshua C. Grieger, Lakshmanan Govindasamy, Mavis Agbandje-McKenna, and R. Jude Samulski. 2006. "Single Amino Acid Changes Can Influence Titer, Heparin Binding, and Tissue Tropism in Different Adeno-Associated Virus Serotypes." *Journal of Virology* 80 (22): 11393–97. https://doi.org/10.1128/jvi.01288-06.

Wu, Zhijian, Hongyan Yang, and Peter Colosi. 2010. "Effect of Genome Size on AAV Vector Packaging." *Molecular Therapy* 18 (1): 80–86. https://doi.org/10.1038/mt.2009.255.

Xiao, P.-J., and R. J. Samulski. 2012. "Cytoplasmic Trafficking, Endosomal Escape, and Perinuclear Accumulation of Adeno-Associated Virus Type 2 Particles Are Facilitated by Microtubule Network." *Journal of Virology* 86 (19): 10462–73. https://doi.org/10.1128/jvi.00935-12.

Xiao, Xiao, Juan Li, and Richard Jude Samulski. 1998. "Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus." *Journal of Virology* 72 (3): 2224–32. https://doi.org/10.1128/jvi.72.3.2224-2232.1998.

Xie, Qing, Weishu Bu, Smita Bhatia, Joan Hare, Thayumanasamy Somasundaram, Arezki Azzi, and Michael S. Chapman. 2002. "The Atomic Structure of Adeno-Associated Virus (AAV-2), a Vector for Human Gene Therapy." *Proceedings of the National Academy of Sciences of the United States of America* 99 (16): 10405–10. https://doi.org/10.1073/pnas.162250899.

Xue, K., M. Groppe, A. P. Salvetti, and R. E. MacLaren. 2017. "Technique of Retinal Gene Therapy: Delivery of Viral Vector into the Subretinal Space." *Eye (Basingstoke)* 31 (9): 1308–16. https://doi.org/10.1038/eye.2017.158.

Zhang, M. Q. 1998. "Statistical Features of Human Exons and Their Flanking Regions." *Human Molecular Genetics* 7 (5): 919–32. https://doi.org/10.1093/hmg/7.5.919.

Zhong, Li, Baozheng Li, Cathryn S Mah, Lakshmanan Govindasamy, Mavis Agbandje-McKenna, Mario Cooper, Roland W Herzog, et al. 2008. "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 of the United States of America* 105 (22): 7827–32. https://doi.org/10.1073/pnas.0802866105.

Zhou, Xiaohuai, Xinghua Zeng, Zhenghong Fan, Chengwen Li, Thomas McCown, R. Jude Samulski, and Xiao Xiao. 2008. "Adeno-Associated Virus of a Single-Polarity DNA Genome Is Capable of Transduction in Vivo." *Molecular Therapy* 16 (3): 494–99. https://doi.org/10.1038/sj.mt.6300397.

Zinn, Eric, Simon Pacouret, Vadim Khaychuk, Heikki T. Turunen, Livia S. Carvalho, Eva Andres-Mateos, Samiksha Shah, et al. 2015. "In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector." *Cell Reports* 12 (6): 1056–68. https://doi.org/10.1016/j.celrep.2015.07.019.