Cell Reports
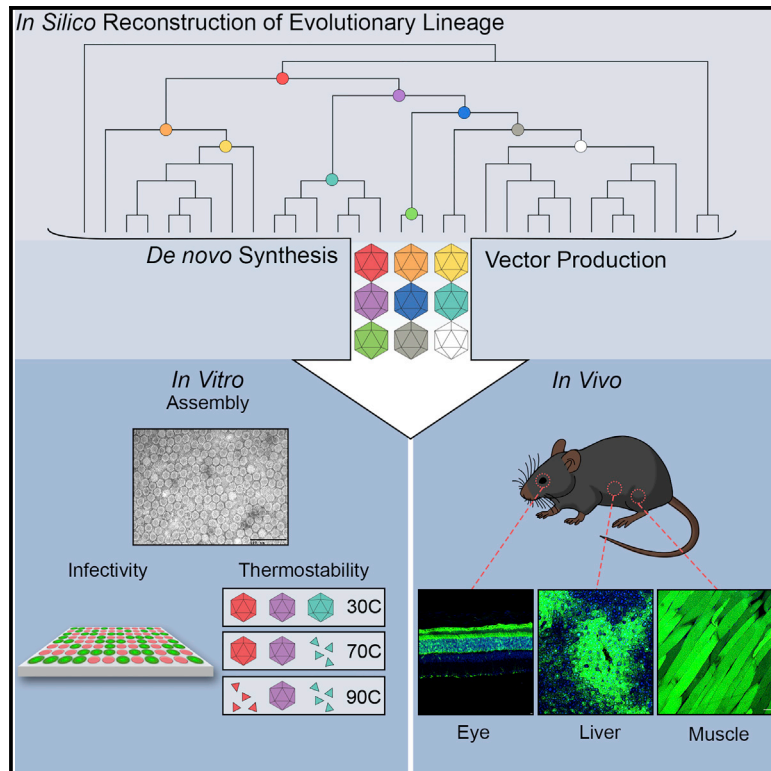
In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector

Graphical Abstract

Highlights

- In silico ancestral sequence reconstruction leads to infectious viral particles

- Anc80, an ancestor of AAV1, 2, 8, and 9, is a potent in vivo gene therapy vector

- The putative evolutionary lineage of AAV is functionally restored

- Ancestral sequence reconstruction elucidates complex structure-function relations

Authors

Eric Zinn, Simon Pacouret, Vadim Khaychuk, ..., Eva Plovie, Ru Xiao, Luk H. Vandenberghe

Correspondence

luk_vandenberghe@meei.harvard.edu

In Brief

Adeno-associated virus (AAV) vectors have enabled preclinical and, more recently, clinical therapeutic gene transfer. Zinn et al. reconstruct in silico approximations of evolutionary capsid intermediates yielding functional particles. Predicting the ancestor of the commonly used AAV serotypes 1, 2, 3, 7, 8, and 9 results in a potent gene therapy vector with broad applications.

CellPress

Resource

In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector

Eric Zinn,[1,2] Simon Pacouret,[1,2] Vadim Khaychuk,[1,2] Heikki T. Turunen,[1,2] Livia S. Carvalho,[1,2] Eva Andres-Mateos,[1,2] Samiksha Shah,[1,2] Rajani Shelke,[1,2] Anna C. Maurer,[1,2] Eva Plovie,[1,2] Ru Xiao,[1,2] and Luk H. Vandenberghe[1,2,3,*]

[1]Grousbeck Gene Therapy Center, Schepens Eye Research Institute and Massachusetts Eye and Ear Infirmary, 20 Staniford Street, Boston, MA 02114, USA
[2]Ocular Genomics Institute, Department of Ophthalmology, Harvard Medical School, 243 Charles Street, Boston, MA 02114, USA
[3]Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138, USA
*Correspondence: luk_vandenberghe@meei.harvard.edu
http://dx.doi.org/10.1016/j.celrep.2015.07.019

SUMMARY

Adeno-associated virus (AAV) vectors have emerged as a gene-delivery platform with demonstrated safety and efficacy in a handful of clinical trials for monogenic disorders. However, limitations of the current generation vectors often prevent broader application of AAV gene therapy. Efforts to engineer AAV vectors have been hampered by a limited understanding of the structure-function relationship of the complex multimeric icosahedral architecture of the particle. To develop additional reagents pertinent to further our insight into AAVs, we inferred evolutionary intermediates of the viral capsid using ancestral sequence reconstruction. In-silico-derived sequences were synthesized de novo and characterized for biological properties relevant to clinical applications. This effort led to the generation of nine functional putative ancestral AAVs and the identification of Anc80, the predicted ancestor of the widely studied AAV serotypes 1, 2, 8, and 9, as a highly potent in vivo gene therapy vector for targeting liver, muscle, and retina.

INTRODUCTION

Extensive preclinical studies have established a favorable safety profile for adeno-associated virus (AAV) vectors. In addition, AAV vectors have enabled demonstration of in vivo gene therapy efficacy in animal models of disease with etiologies ranging from inherited to infectious to common complex (Hastie and Samulski, 2015; Schnepp and Johnson, 2014). Furthermore, recent early-stage AAV clinical trials have led to the first demonstrations of clinical benefit in two forms of inherited blindness with AAV2 (Bainbridge et al., 2008; Jacobson et al., 2012; MacLaren et al., 2014; Maguire et al., 2008) and hemophilia B with AAV8 (Nathwani et al., 2011). One treatment based on AAV1 has been awarded a drug license by European regulators (Bryant et al., 2013). Based on these data, AAVs have been proposed as a platform technology for therapeutic in vivo gene delivery.

AAV is a 25-nm non-enveloped icosahedral capsid virus carrying a 4.7-kb single-stranded DNA genome flanked by inverted terminal repeats (ITRs). AAV classifies as a *Dependoparvovirus* genus within the *Parvoviridae* family. Its genome comprises genes encoding for replication (Rep), structural capsid (Cap), and assembly (AAP) proteins. AAV is a helper-dependent virus, requiring the heterologous cofactors to complete a replicative cycle that an adenovirus or herpesvirus can provide in the context of a co-infection. Replication-deficient AAV can be generated by eliminating all viral coding sequence in *cis* and providing those in *trans* during vector production. Particles generated in this manner can encode any type of transgene cassette that does not exceed the size of the genome of the wild-type virus (~4.7 kb) and enable gene transfer in vitro and in vivo to multiple cell and tissue targets (Knipe and Howley, 2013).

The AAV particle is composed of three C-terminally overlapping Cap proteins named VP1, 2, and 3. VP3, the smallest member of these structural proteins, is necessary and sufficient for full capsid assembly through multimerization. VP1, required for particle infectivity, and VP2, reported to be redundant structurally or functionally, are embedded in the wild-type T = 1 viral architecture at a VP1:2:3 ratio of ~1:1:10. A total of 60 VP monomers assemble into an icosahedral capsomer along 2-, 3-, and 5-fold axes of symmetry (Knipe and Howley, 2013). Every monomer within the 60-mer structure interfaces with seven neighboring capsid monomers. High-resolution crystallography studies identify a conserved core structure composed of an eight-stranded β-barrel motif and the αA helix as well as nine surface-exposed domains (VR-I to -IX), which can vary between the known primate AAV serotypes (DiMattia et al., 2012; Xie et al., 2002).

The structural diversity of natural AAVs is therefore largely contained within the nine surface regions of the capsid, which functionally results in divergent receptor-binding properties, post-entry trafficking, host response, and gene transfer efficiency to various cell and tissue targets. Structural determinants for many of these properties have largely remained elusive, with a few notable exceptions: receptor-binding motifs for certain serotypes (Bell et al., 2012; Kern et al., 2003; Wu et al., 2000), phospholipase activity and nuclear localization activity on VP1,2 unique domains (Girod et al., 2002; Grieger et al., 2006), a limited

number of docking sites for monoclonal antibodies (DiMattia et al., 2012), a motif enhancing T cell immunogenicity (Vandenberghe et al., 2006), and a select number major histocompatibility complex I (MHCI)-restricted T cell epitopes (Mingozzi et al., 2007) have previously been mapped.

In an era of active clinical translation in AAV gene therapy, there is a desire to better understand the structure-function relationship of AAVs within the constraints of the particle architecture in order to further model and modulate the pharmacology of this new class of drugs to improve transduction efficiency and specificity, alter tropism, and reduce immunogenicity. To enable mechanistic studies, map structural determinants of these phenotypes, and facilitate future rational AAV design aimed to mitigate the clinical limitations of the current technology, both reverse and forward genetic studies on AAVs have been pursued. A wealth of naturally occurring AAV variants have been studied extensively and yielded large descriptive datasets (Gao et al., 2003, 2004, 2005). Forward genetic studies have resulted in important findings (Adachi et al., 2014; Asuri et al., 2012; Bartel et al., 2012; Dalkara et al., 2013; Huttner et al., 2003; Koerber et al., 2008; Li et al., 2008; Lisowski et al., 2014; Maheshri et al., 2006; Mays et al., 2013; Perabo et al., 2003; Shen et al., 2007; Wu et al., 2000) but often remain constrained by the limited tolerance for structural change of the rigid architecture of AAVs (Adachi et al., 2014; Huttner et al., 2003; Lochrie et al., 2006; Mays et al., 2013; Shen et al., 2007; Vandenberghe et al., 2009; Wu et al., 2000). Rational AAV design has therefore been difficult in structurally isolating and modulating vector phenotypes while retaining integrity and the desirable aspects of its innate function and biology of the particle.

Here, we hypothesized that the divergent AAV phenotypes and structural determinants underlying its biology and pharmacology could be mapped by recreating the evolutionary lineage of this virus. Under selective pressure, evolutionary intermediates along the lineage undergo phenotypic modulation while balancing these changes with the retention of structural and functional integrity that are expected to highlight evolutionary couplings and epistatic interactions. We used ancestral sequence reconstruction (ASR) methods to predict the amino acid sequence of putative ancestral AAV capsid monomers using maximum likelihood (ML) methods (Finnigan et al., 2012). This led to the reconstruction of nine nodes along the phylogeny of common ancestry of most of the AAV serotypes under clinical testing or consideration (AAV1–3 and 6–9). Through de novo gene synthesis of these reverse-translated ancestral capsid proteins, we demonstrate assembly and infectivity of members across this AAV lineage. The most distal presumed ancestor of the lineage, Anc80, was extensively characterized for its performance as a gene therapy vector and was demonstrated to be a potent broadly applicable vehicle for gene therapy with unique properties.

RESULTS

In Silico Ancestral Sequence Reconstruction of AAV Capsid Protein

In lieu of attempting to isolate an intact ancestral viral sequence from proviral DNA or archeological samples, contemporary AAV sequence data were integrated through phylogenetic analysis and ML-ASR in order to infer the putative ancestral amino acid sequence for the AAV Cap. A total of 75 sequences of AAV serotype isolates and variants from previous biomining efforts (Gao et al., 2003, 2004) led to a robust AAV Cap phylogeny generated with PHYML (Guindon et al., 2010) with AAV5 as an outgroup. Only full-length AAV capsids were included in this analysis that were (1) naturally occurring in primate populations, (2) previously demonstrated to assemble and infect efficiently, and (3) not known to have arisen through recombination events in its natural history, as traditional phylogenic analysis and ASR do not account for horizontal evolutionary events. The dendrogram in Figure 1 models the evolutionary path of AAVs with early speciation of AAV4 and 5 serotypes, parallel to a single node, named Anc80, from which most known contemporary AAVs evolved. These serotypes include AAV1, 2, 8, and 9, currently in human gene therapy trials. Nodes in this phylogeny were named Anc and numbered sequentially. To validate our approach, the Anc80 node was developed into a recombinant virus for possible use as a gene therapy vector (Figure 1).

Anc80 was chosen in part because the reconstruction of this node was highly informed by the abundance of naturally occurring and clinically relevant AAV descendants. Furthermore, Anc80 is embedded in the phylogeny of the *Dependoparvoviridae* with known helper-dependent primate AAVs that arose prior to Anc80's speciation (Figure 1), making it more likely that the ancestrally reconstructed particle retains the basic properties shared within this family. Using ML methods, a protein sequence prediction was derived for Anc80 based on calculated posterior probabilities for each residue in a particular position. In order to account for the uncertainty in selecting the appropriate amino acid in each position, we aimed at generating all possible sequence permutations for positions with individual-amino-acid posterior probabilities with p $\geq$ 0.3. A representation of this library, Anc80Lib, is illustrated in Figure 2A in a part-structural alignment with an AAV2 and AAV8 reference capsid sequence. Practically, this led to a probabilistic sequence space composed out of sequences for which most of the 736 Anc80 capsid amino acid positions were fixed due to the high certainty in those positions by the ML-ASR, while for 11 positions, two amino acid options were provided, resulting in a sequence space encompassing 2^{11} = 2,048 permutations.

Structural and sequence alignment of Anc80Lib with extant AAVs and their X-ray crystallography data highlight significant divergence from currently known circulating AAVs. The closest homolog as determined via BLAST search is rh.10, a rhesus macaque isolate within clade E of the primate *Dependoparvoviridae*, which differs from Anc80Lib by minimally 7.8%, which accounts for 58 divergent amino acid positions (Figure 2B). AAV8 and AAV2 differ 8.7% and 12.0%, respectively, and those 65–89 variable sites are distributed unevenly over the entire VP1 protein, including the VP1 and 2 unique domains (Figures 2A and 2B). Divergence is highest in the hypervariable domains I, IV, VII, and VIII, both in terms of sequence as well as based on structural modeling of Anc80Lib clones in overlay with AAV2 and eight monomeric structures (Figures 2A and 2C). Mapping of the variable Anc80 residues onto trimeric X-ray crystallography models of AAV2 and AAV8 in Figure 2D highlight most changes to be
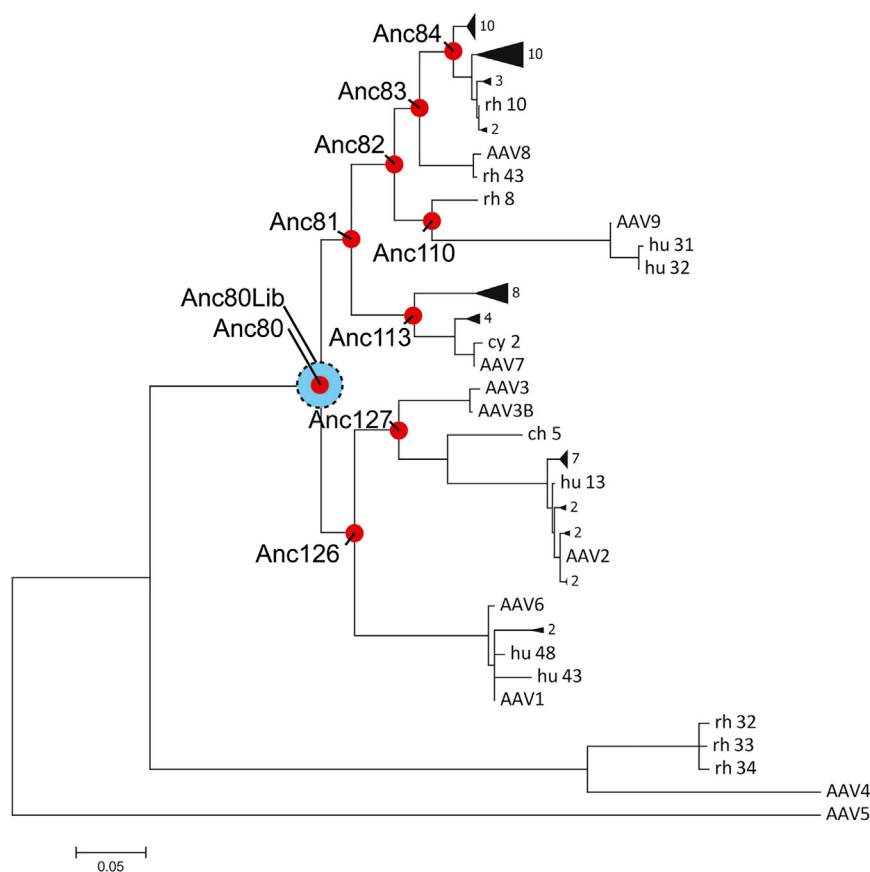
able signal over background in this rudimentary screening assay. Several lead candidates with highest luciferase signal progressed to sequencing confirmation and titration for DNase-I-resistant genome-containing particles (GCs) and infectivity on HEK293 cells. Based on these results Anc80L65, the 65th Anc80Lib clone that was evaluated, was selected for further characterization. Anc80L65 vector yields from cell lysate are between 82% and 167% of AAV2 yields, yet they were depressed compared to the high-yielding AAV8 (3%–5% relative AAV8 yields). In vitro infectivity on HEK293 is inferior to AAV2 but superior to AAV8 on a particle-per-cell basis.

Anc80L65 vector preparations were produced and purified on an iodixanol gradient at scale following traditional protocols and subjected to a variety of biochemical, biophysical, and structural analyses. Particles within a purified preparation of Anc80L65 were visualized under negative staining by electron microscopy (EM) (Figure 3A). Anc80L65 virions present as relatively uniform hexagonally shaped particles with a diameter of ~20–25 nm, not unlike other AAV capsomers. Denatured particles resolved under SDS electrophoresis into three bands of 60, 72, and 90 kDa, in a ratio of ~1:1:10 corresponding to the VP1–3 proteins from AAV2 and AAV8 particles (Figure 3B). Analytical ultracentrifugation (AUC) allowed us to determine the sedimentation coefficient of genome containing Anc80L65 at 88.9 S, slightly increased from AAV8's (85.9 S) (Figure 3C). This analysis permitted us further to determine the relative abundance of empty or lower-density assembled particles, presumed to be lacking a vector genome, as well as overall purity. One concern was that inaccurate modeling of the ancestral capsid sequence may have resulted in a structure deficient in its ability to package genomes and would result in a skewed empty versus full ratio in Anc80L65 preparations. Results indicated ~16% empty versus 85% full particles in our preparation, in line with observations with AAV8 (Figure 3C). Additionally, we hypothesized particle stability may be reduced due to suboptimal modeling of the ancestral capsid composition. We subjected the particle to heat-stability assays, which determined (against our expectations) that Anc80L65 was 15°C –30°C more heat stable than AAV2 and AAV8 (Figure 3D).

concentrated on the external surface of the virion, particularly on peak and flanks of the protrusions around the 3-fold axis of symmetry. However, a significant number of variable residues were also noted on the surface-exposed domains outside of the 3-fold axis in addition to a smaller number of variations on the internal surface of the particle and on regions of Cap that are not resolved in the X-ray structures.

Anc80 Synthesis and Basic Characterization

Anc80Lib protein sequences were subsequently reverse translated and generated by gene synthesis in pooled library format. Capsid genes were cloned into an AAV packaging plasmid encoding AAV2 Rep into pAnc80Lib following which the library was deconvoluted clonally. Individual clones (named pAnc80LX, with X a consecutive number) were evaluated in isolation to avoid potentially interfering competitive interactions in a minimally divergent library population. A portion of individual Anc80 clones were Sanger sequenced, verifying integrity and complexity requirements. Clonal Anc80 plasmids were co-transfected with a ΔF6 adenoviral helper plasmid, an expression construct for AAP derived from AAV2 (AAP2), and ITR flanked expression construct encoding luciferase. A total of 776 library clones were produced and inoculated at equal volume of producer cell lysate on HEK293 cells in a semi-high-throughput assay aiming to assess combined particle assembly and transduction efficiency. Approximately 50% of the Anc80 clones led to detect-
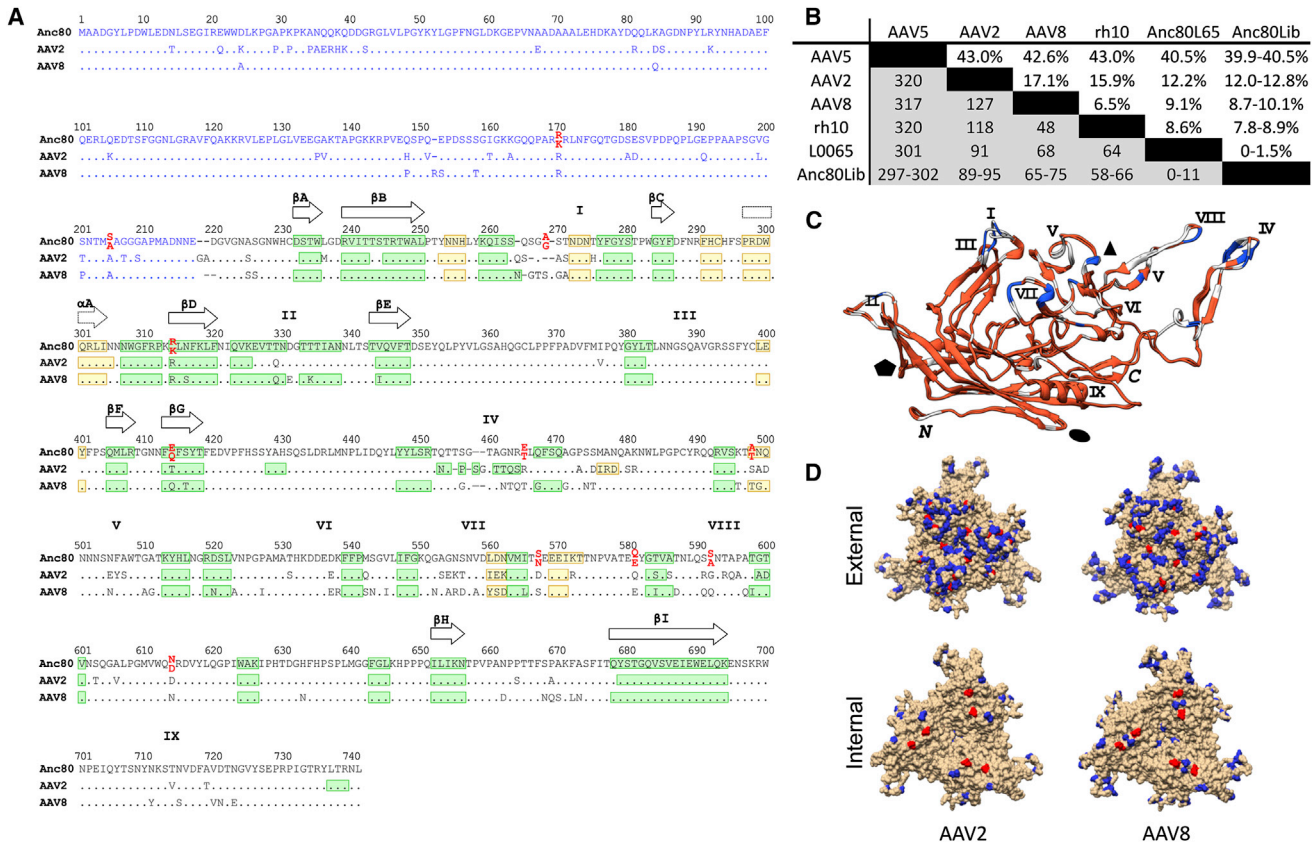
Figure 2. Sequence and Structural Analysis of Anc80 Vectors

(A) Sequence alignment of Anc80, AAV2, and AAV8 VP3 proteins. A structural alignment derived from the crystal structures of AAV2 (PDB: 1LP3) and AAV8 (PDB: 2QA0) VP3 and the predicted structure of Anc80L65 VP3 was generated with UCSF Chimera (Pettersen et al., 2004) and is represented in black print. The blue region is a non-structural alignment of the VP1/VP2 domains of AAV2, AAV8, and Anc80 (Notredame et al., 2000). Ambiguous residues in Anc80Lib are in red print with the lower position corresponding to Anc80L65 residues. β strands and α helices are represented in green and yellow, respectively. The positions of the nine β strands forming the AAV antiparallel β-barrel are depicted with plain arrows whereas the position of the conserved core α-helix is depicted with a dotted arrow. The approximate positions of variable regions (VR) I–IX are represented by the roman numerals above the sequence alignment.

(B) AAV Cap sequence divergence matrix. Above the diagonal, the table shows the percent sequence divergence from selected AAV serotypes, as well as rh.10, the most homologous VP1 sequence as determined by BLAST. Below the diagonal, the number of amino acid differences per position is presented.

(C) Superimposition of AAV2 and AAV8 VP3 crystal structures with Anc80L65 VP3 predicted structure. The color code depicts the amino acid conservation among the three aligned sequences of (A) (red, highest conservation; blue, lowest conservation). Variables regions I–IX and C/M termini are indicated in black. The approximate positions of the 2-, 3- and 5-fold axes are represented by the black ellipse, triangle, and pentagon, respectively.

(D) Structural mapping of amino acid changes as compared to AAV2 (left) and AAV8 (right) on VP1 trimer visualizing the external (top) and internal (bottom) of the virion. Colored residues are divergent in Anc80. Red residues are ambiguous via ASR and therefore dimorphic in Anc80Lib.

In Vivo Gene Transfer and Transduction of Anc80L65 in a Murine Model

Next, we evaluated the ability of Anc80L65 packaged transgenes to be delivered and expressed from three clinically relevant target tissues and routes of administration (ROA) in the C57Bl/6 mouse: (1) liver following a systemic injection, (2) skeletal muscle following direct intramuscular injection, and (3) a subretinal injection for outer retina targeting. Large-scale preparations of Anc80L65 were produced along with AAV2 and AAV8 controls with reporter genes and were injected at equal doses for liver-, muscle-, and retina-directed gene transfer in adult male C57Bl/6 mice. Expression (presented in Figure 4) was monitored qualitatively (EGFP and/or LacZ) for all three target tissues and quantitatively via serum ELISA measurement of the secreted

hA1AT (liver) at various time points. Liver-directed gene transfer was observed to be robust via two routes of administration and transgenes (Figures 4A–4C). Analogously to AAV8, hepatocytes were targeted efficiently as observed by LacZ and GFP staining surpassing the limited permissivity described for AAV2 (Nakai et al., 2002, 2005). Quantitatively, Anc80 demonstrated similar efficiency of transduction to AAV8 by intracellular reporter and a secreted serum protein transgene product. Dose-ranging studies demonstrated a linearity of gene transfer with dose above 10^{10} GCs/mouse but a threshold below which linearity was not maintained for hA1AT (and less obvious by EGFP) (Figures 4B and 4C). A biodistribution study at the high dose of 5×10^{11} GCs/mouse was conducted at days 7 and 28 postinjection to evaluate tissue distribution of vector genomes in liver, heart,
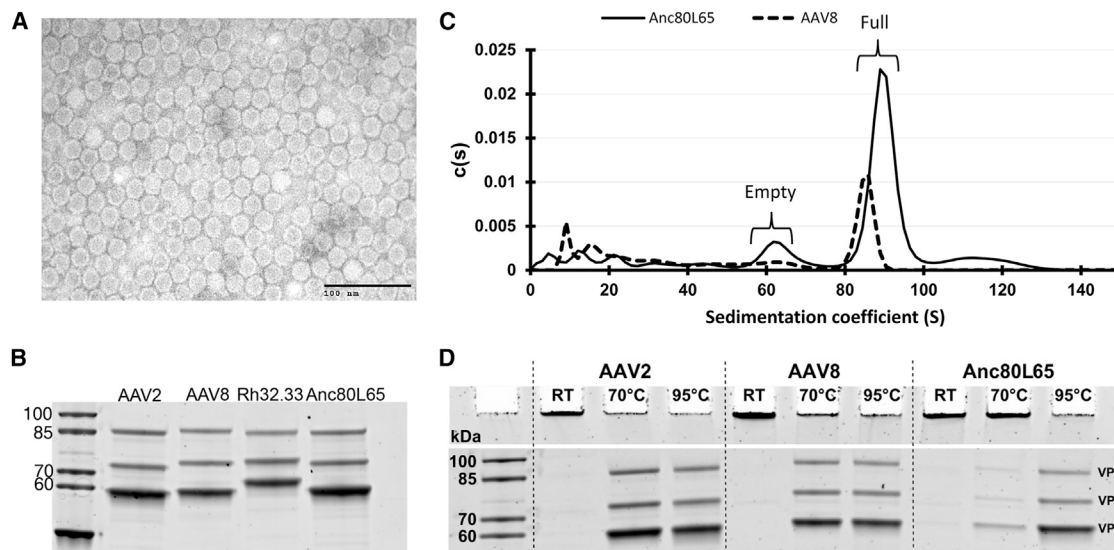
Figure 3. Biophysical and Biochemical Characterization of Anc80L65

(A) Negative staining transmission electron microscopy (TEM) of Anc80L65 particles.

(B) Anc80L65 viral protein composition by SDS-PAGE. Purified AAV preparations were analyzed by SDS-PAGE, demonstrating similar incorporation levels of the monomers VP1–3.

(C) Anc80L65 sedimentation and empty:full particle profile. Sedimentation coefficient distributions were derived from the refractive index optical measurement systems during analytical ultracentrifugation of AAV8 and Anc80L65 iodixanol purified preparations.

(D) Anc80L65 thermostability. SDS page thermostability assessment of an AAV particle illustrates dissociation into monomers for AAV2 and 8 at temperatures below 70°C for AAV2 and 8 but requiring 95°C melting temperatures for Anc80L65.

spleen, kidney, and lung of Anc80L65 alongside AAV8 as a control (Table S1). Results show similar ranges of gene transfer of Anc80 to AAV8 in the tissues tested, with moderate increases for Anc80L65 in spleen, heart, and lung. Via direct skeletal intramuscular injection, Anc80 efficiently targeted myofibers proximal to the injection site and longitudinally extending across the fiber (Figures 4A and S1). Retinal transduction after subretinal injection is efficient in targeting the retina pigment epithelium (RPE), as was the case in AAV2 and AAV8 as previously noted (Auricchio et al., 2001). Photoreceptor targeting, a more difficult cell target, as is documented for AAV2, was observed with AAV8 and Anc80L65 (Allocca et al., 2007). While both AAV8 and Anc80L65 targeted the majority of photoreceptor cells, transduction with Anc80L65 leads consistently to higher expression levels per cell. A limited number of cells in the inner retina were also observed to be GFP positive by Anc80L65 transduction (Figure 4A).

Anc80L65 Gene Transfer and Expression in Non-human Primate Liver

Given the robust hepatotropism of Anc80L65 in mice, we aimed at evaluating gene transfer of Anc80L65 in a large animal model. Six female rhesus macaques that were previously enrolled in prior studies unrelated to AAV were injected via saphenous vein with either AAV8 or Anc80L65 at a clinically relevant dose of 10^{12} GCs/kg (Table S2). AAVs expressing the rhesus cDNA for the β subunit of the chorionic gonadotropin (rhCG), a transgene product that the animals are tolarized for in order to avoid a non-self transgene immune response, was used as transgene.

Animals were selected based on neutralizing antibody (NAB) to AAV8 and Anc80L65 with serum levels prior to injection below 1/4 titer. Gene transfer was assessed by TaqMan qPCR for vg of total liver DNA (caudal lobe) 70–71 days following injection (Figure 4D). Surprisingly, two out of three control AAV8-injected animals had underwhelming gene transfer (<0.1 vg/dg), likely due to low-level NAB at the time of injection undetectable by standard NAB assays as reported in previous studies. One AAV8 animal, presumably with no or minimal NAB to AAV8, demonstrated gene transfer levels for liver within the expected range of 0.81 vg/dg. Anc80L65 gene transfer was efficient, with three animals yielding hepatic transgene copy numbers ranging from 0.73 to 3.56 vg/dg. Liver expression was monitored via qRT-PCR (Figure 4E); Anc80L65 led to expression superior to the AAV8 and achieved rhCG transcript levels between 13% and 26% of total GAPDH mRNA amounts in all liver lobes.

Safety, Immunology, and Toxicology of Anc80L65

The consideration to use any efficient gene-delivery vector system for therapeutic application requires extensive evaluation of its safety for clinical use. In addition, the use of a novel agent that may approximate an ancestral state of a *Dependoparvovirus* may further raise those concerns. Here, in a non-formal preclinical setting, we examined several important aspects that may limit Anc80L65 from a safety perspective. Animal expression studies (Figure 4) were monitored for obvious signs of toxicity during the in-life phase of the study and for target-tissue-specific toxicity. No notable adversity was found to be associated with the vector injection. In brief, vector administration following
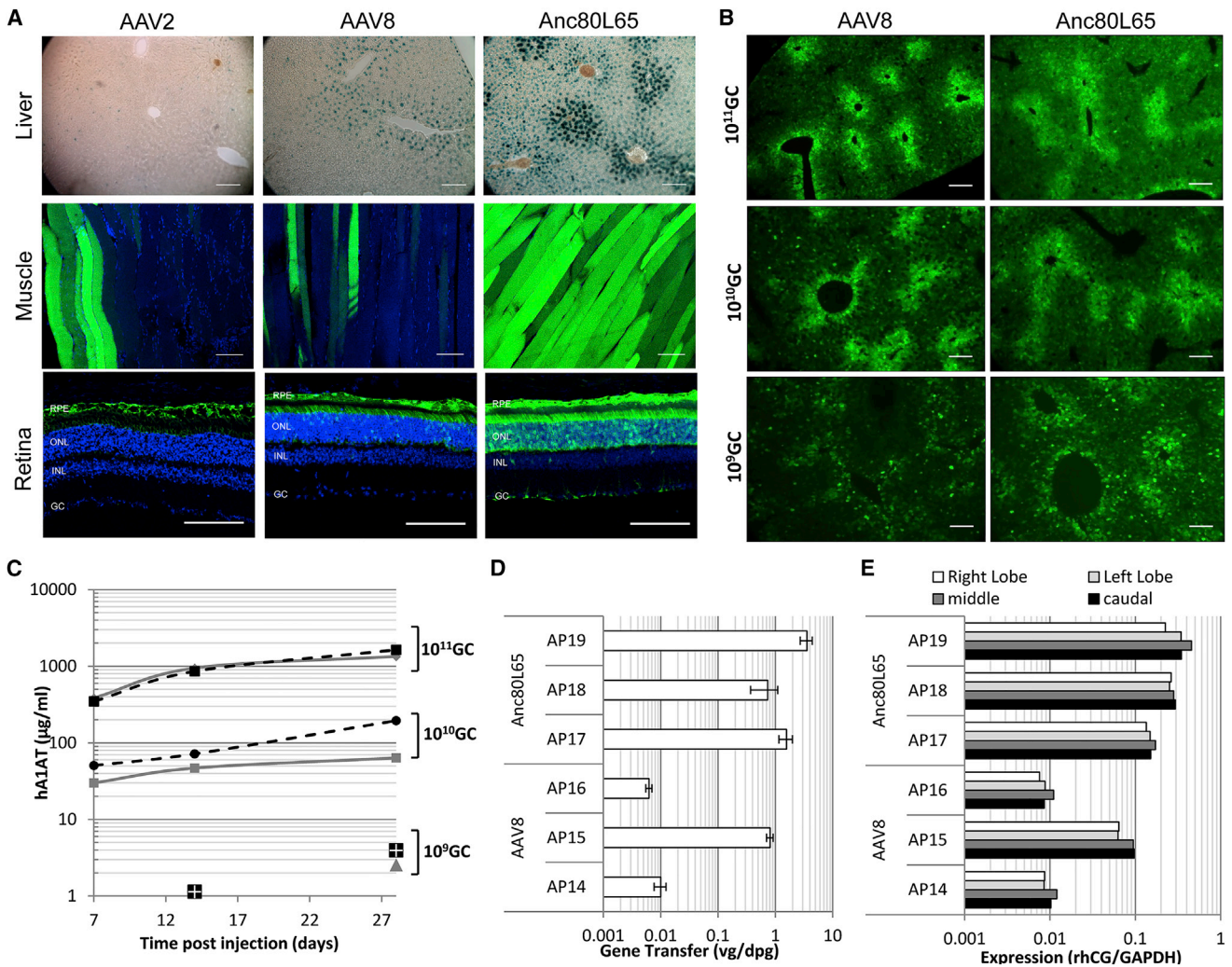
Figure 4. Anc80L65 In Vivo Transduction Biology

(A) Mouse liver transduction and lacZ transgene expression comparison of AAV2, AAV8, and Anc80L65.TBG.nLacZ in liver 28 days after intraperitoneal delivery at a dose of 3.9×10^{10} GCs (C57Bl/6, n = 3) (top). AAV2, AAV8, and Anc80L65 muscle tropism in mouse 28 days following an intramuscular delivery at a dose of 10^{10} GCs to the rear-right thigh (gastrocnemius) (n = 5) (middle). See also Figure S1. Comparison of EGFP transgene expression among AAV2, AAV8, and Anc80L65 in the murine retina after subretinal delivery at a dose of 2×10^{9} GCs (bottom). AAV2 shows high affinity for RPE cells, while both RPE and photoreceptors are targeted using AAV8 and Anc80L65 vectors, with Anc80L65 showing higher transduction efficiency compared to AAV2 and 8 (C57Bl/6, n = 4 eyes).

(B) Qualitative dose-response hepatic EGFP-expression analysis following dosing of 10^{11} (top), 10^{10} (middle), and 10^{9} (bottom) GCs comparing AAV8 and Anc80L65 by retro-orbital intravenous injection in mouse. Both AAV8 and Anc80L65 lead to comparable EGFP expression at equal dose throughout the dose ranging (C57Bl/6, n = 3).

(C) Quantitative AAV dose-response analysis measuring mouse serum levels of recombinant human alpha 1-antitrypsin (hA1AT) transgene expression from AAV8 (black symbols: square, 10^{11} GCs; circle, 10^{10} GCs; and four-square, 10^{9} GCs) and Anc80L65 (gray symbols: diamond, 10^{11} GCs; square, 10^{10} GCs; and triangle, 10^{9} GCs) (C57Bl/6, n = 5). See also Tables S1 and S3–S5.

(D) Rhesus macaque liver gene transfer of AAV8 and Anc80L65 expressing Rhesus chorionic-gonadotropin (rhCG) following saphenous vein injection of a dose of 1×10^{12} GCs/kg. Genomic DNA was harvested from macaque liver-lobes and viral genome (vg) per diploid genome (dpg) was measured by qPCR assay. One AAV8 animal and all three Anc80L65 animals successfully received ~1–3 vg per diploid cell of the caudal liver lobe, while 2 AAV8 animals likely had low-level NAB resulting in vector neutralization and limited liver gene transfer. Error bars represent SD. See also Tables S2, S6, and S7.

(E) Transgene mRNA expression of AAV8 and Anc80L65 in NHP caudal, right, left, and middle liver lobes by TaqMan probe-specific qRT-PCR. Protein levels are not available due to lacking antibody and rhCG standards. Quantitation of rhCG transcript is normalized with endogenous GAPDH mRNA levels.

intraperitoneal (maximum dose tested [mdt]: 3.9×10^{10} GCs/mouse), retro-orbital vein injection (mdt: 5×10^{11} GCs/mouse), subretinal (mdt: 2×10^{9} GCs/eye), intravitreal (mdt: 2×10^{9} GCs/eye), and direct intramuscular (mdt: 10^{10} GCs/mouse)

were not observed to have overt toxicity. A more direct assessment was performed in a high-dose intravenous injection of 5×10^{11} GCs/mouse (equivalent to ~2×10^{13} GC/kg) of Anc80L65.TBG.eGFP alongside the following controls: (1) AAV8
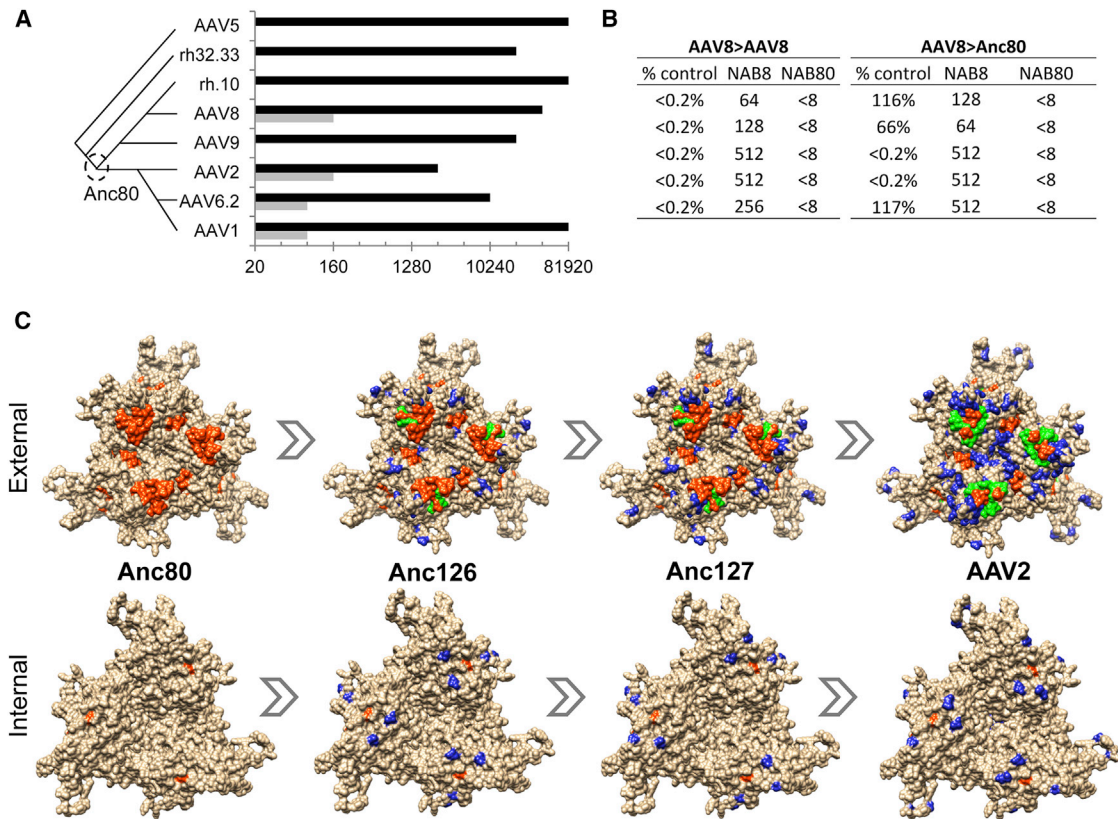
Figure 5. Immunological Characterization of Anc80L65

(A) Rabbit anti-AAV serum cross-reactivity. Rabbit antiserum raised against AAV serotypes (y axis) was tested for NAB to Anc80L65 (gray bars) versus the homologous AAV serotype (black bars) in order to assess relative sero-cross-reactivity. Values (x axis) represent highest dilution at which >50% neutralization is achieved. Phylogenetic relationship between immunizing serotypes is depicted schematically on the left.

(B) Mouse in vivo gene transfer cross-neutralization. C57Bl/6 mice received an intravenous injection of AAV8 or Anc80L65.CASI.EGFP.2A.A1AT 25 days following an intramuscular injection with either saline or AAV8.TBG.nLacZ. 14 days following the second injections serum was titrated by ELISA for hA1AT expression. The table presents the relative hA1AT levels of the pre-immunized mice versus the non-immunized for each vector (% control) and the NAB titer dilutions for AAV8 (NAB8) and Anc80L65 (NAB80) 24 hr prior to the second injection in the immunized group (n = 5).

(C) A non-structural multiple sequence alignment among Anc80, Anc126, Anc127, and AAV2 VP3 sequences was generated using the T-coffee alignment package. AAV2 trimer structure was generated using UCSF Chimera. The blue residues represent residues different from Anc80. The orange residues are defined T and B cell epitopes on AAV2 (Gurda et al., 2013; Mingozzi et al., 2007). Green residues show the overlap between orange and blue residues to highlight mapped epitopes altered in the putative evolutionary intermediates. Human T cell epitopes with major histocompatibility complex haplotype: *VP*QYGYLTL (B*0702), *SAD*NNNS*EY* (A*0101), YHLNGRDSL (B*1501), and TTSTRTWAL (B*0801). (Mingozzi et al., 2007) Mouse B cell epitopes of defined AAV2 antibody *SAD*NNNS plus *RGNRQ* for C37B Fab (Gurda et al., 2013). In *italics* are the residues within each epitope that are distinct between Anc80L65 and AAV2.

with the same transgene cassette and (2) an equal volume saline injection. Mice were phlebotomized pre-injection, 2 hr, and 1, 3, 7, 14, and 28 days postinjection, and blood was analyzed for cell blood counts (CBC) and serum chemistry (Chem) (Tables S3 and S4). CBC/Chem values for Anc80L65 were within normal range or comparable to controls. Serum from the 2-hr, 24-hr, 3-day, and 7-day time points were further evaluated for cytokines as a measure of innate immune response to the vector antigens by multiplex 23 cytokine analysis (Table S5). Cytokines for Anc80L65 were overall concordant with those for saline and AAV8 control serum, and no major cytokine elevations or decreases were observed; however, in some instances, they were moderately outside the ranges set by the saline control values in a manner that was more apparent for Anc80L65 than AAV8. Similar analyses were performed on the blood from the

rhesus studies described in Figures 4D and 4E. Analogous to the mouse studies, CBC and Chem values did not demonstrate signs of toxicity related to the AAV8 or Anc80L65 test article (Tables S6 and S7)

Pre-existing immunity to AAV serotypes is known to block gene transfer and may put the patient at risk for adversity due to recall of memory T cells toward vector antigens shared with the naturally occurring *wild-type* virus involved in the primary infection. We used high-titer rabbit antiserum raised against AAV serotypes 1, 2, 5, 6.2, 8, 9, and rh.32.33. We furthermore included rh.10, as its sequence is most closely homologous to Anc80L65, differing in 8.6% of residues. In Figure 5A, sera were tested for their ability to neutralize Anc80L65 versus the homologous vector capsid it was raised against. Results demonstrate no cross-reactivity to the structurally highly divergent

AAV5 and rh32.33, while AAV2, 6.2, and 8, presumed descendants of Anc80L65, demonstrated low-level cross-reactivity, with between 16- and 1,024-fold reduced neutralization compared to the homologous virus. Among Anc80 lineage members, no cross-reactivity was observed above the limit of sensitivity for AAV9 and rh.10. Next, we aimed at validating these results in an in vivo model for neutralization by pre-immunizing animals for AAV8 via intramuscular route and, 25 days following the immunization, assessing the neutralization of Anc80L65 following intravenous injection in comparison to AAV8 (Figure 5B). Neutralization was complete for AAV8 in the AAV8 pre-immunized animals. Anc80L65 was neutralized in 2/5 animals, yet demonstrated between 60%–117% of transduction in three out of five animals notwithstanding demonstrated AAV8 NAB in those animals. These results demonstrate partial cross-reactivity of Anc80L65 with AAV8 in rabbit and mouse.

AAV Lineage Analysis and Reconstruction

Strengthened by the successful synthesis of Anc80L65 based on ASR and its demonstration as a producible, stable, and highly infectious agent for gene therapy, we aimed at providing additional validation of our approach and modeling methodology by reconstructing the lineage of AAV further. Our ambition generating this additional set of reagents was to provide structural intermediates of phenotypically distinct AAVs permitting empirical evaluation of the structure-function relationship within this viral family and highlighting important epistatic couplings informative to future AAV rational design approaches. A total of eight additional evolutionary intermediates of AAV were reconstructed by ASR and synthesized in the laboratory (Figure 1): Anc81, Anc82, Anc83, Anc84, Anc110, and Anc113 were resolved in the branching leading toward AAV7, 8, and/or 9, while Anc126 and Anc127 are positioned in the natural history of AAV1, 2, and/or 3. For each of these, the sequence was determined by selecting the amino acid with highest posterior probability per position (Figure S2). First, we determined GC viral vector yields following HEK293 standard triple transfection by TaqMan qPCR. Results (shown in Figure 6A) demonstrate increased productivity from Anc80 as the putative ancestor in the AAV7–9 lineage, in line with the higher production yields of those serotypes such as AAV8. The AAV1–3 branch did not present yield increases, and a very poor particle yield was observed for Anc126. It is possible that Anc126 yields can be improved upon through leveraging the statistical space, as was the case for Anc80. However, it is equally likely that Anc126 ASR is less informed due to undersampling of this branch of the AAV phylogeny. We further tested infectivity of the produced particles at equal particle doses in vitro on HEK293 by GFP and luciferase. All newly synthesized Anc vectors demonstrated infectivity, although at varying degrees (Figure 6B). In the AAV7, 8, and 9 lineage, infectious titers were overall depressed and more similar to the AAV8 phenotype than that of Anc80. Anc127, the only intermediate in the Anc80 to AAV2 lineage that could be tested at equal dose, demonstrated declined transduction efficiency as compared to both Anc80 and AAV2. We further tested the heat-stability profile of selected evolutionary intermediates in both branches of this lineage (Figure 6C). Interestingly, Anc81 and Anc82 demonstrated high yet moderately decreased melting temperature in a thermostability

assay compared to Anc80L65, suggesting over time a gradual reduction of particle denaturation temperature in this branch. In contrast, Anc127 demonstrated an even further increase from the already highly thermostable Anc80L65 vector.

Lastly, we explored the ability of ASR to disrupt known epitopes to AAV2. Only few B or T cell epitopes have been mapped on AAV2 to date, all of which were mapped onto Anc80L65, Anc126, Anc127, and AAV2, representing the AAV2 lineage. The introduction of the sequential mutations between these putative evolutionary intermediates highlights in Figure 5C the overlap between the mutations and two out of four human T cell epitopes (Mingozzi et al., 2007) and two out of two mouse B cell epitopes (Gurda et al., 2012). These data highlight the potential of ASR to be used as a method to eliminate or modulate antigenic regions onto the AAV capsid and may suggest immunity was a major selective pressure in the natural history of AAVs.

DISCUSSION

Forward genetics studies on multimeric complex protein assemblies are intricately difficult due to the structural constraints imposed by secondary, tertiary, and quaternary interactions within and between monomers. Many viruses form icosahedral capsids as an economical way to build larger complex shells from smaller repeat units without expending excessive genetic bandwidth. These 20-faceted structures are composed of 60 subunits (or multiples thereof), each of which can integrate one or more structural viral proteins. As a function of its triangulation number and monomer size, *Parvoviridae*, with AAV as a member, are therefore one of the smallest known viruses, with a capsid diameter in the 18-nm to 26-nm range (Knipe and Howley, 2013). In order for the virus to evolutionary maintain the benefits from this icosahedral structure, sequence permutations are iteratively evaluated, not unlike traditional forward genetic structure-function experimentation in a laboratory setting. Only those monomer conformations that efficiently integrate in the icosahedral structure *and* can provide a selective advantage are able to persist, ultimately driving speciation. We sought to investigate, reconstruct, and learn from this natural experimentation using the simplest of all icosahedral viruses, AAV, as a model.

Here, we aimed at reconstructing the viral lineage of AAV, a commonly used virus for gene transfer through methodologies similar to Gullberg et al.'s reconstruction of Coxsackievirus putative ancestor (Gullberg et al., 2010). First, putative ancestral sequences of the AAV capsid protein was inferred using in silico phylogenetic and statistical modeling. Next, these sequences were synthesized de novo, and using a traditional virological technique, the virions derived were evaluated for assembly and infectivity. We anticipated the solution space containing monomers that structurally integrate into a viral particle and result in infectious virions to be minimally small, given that the limited knowledge base on intra- and intermolecular epistasis or the minimal structural components of AAV that constitute infectivity or function. To hedge for this anticipated low probability of success, ML statistics were used to inform us; we generated a probabilistic sequence space providing margins to account for the ambiguity of the heuristics used. Screening of the vector library that emerged from this sequence space yielded Anc80L65, a
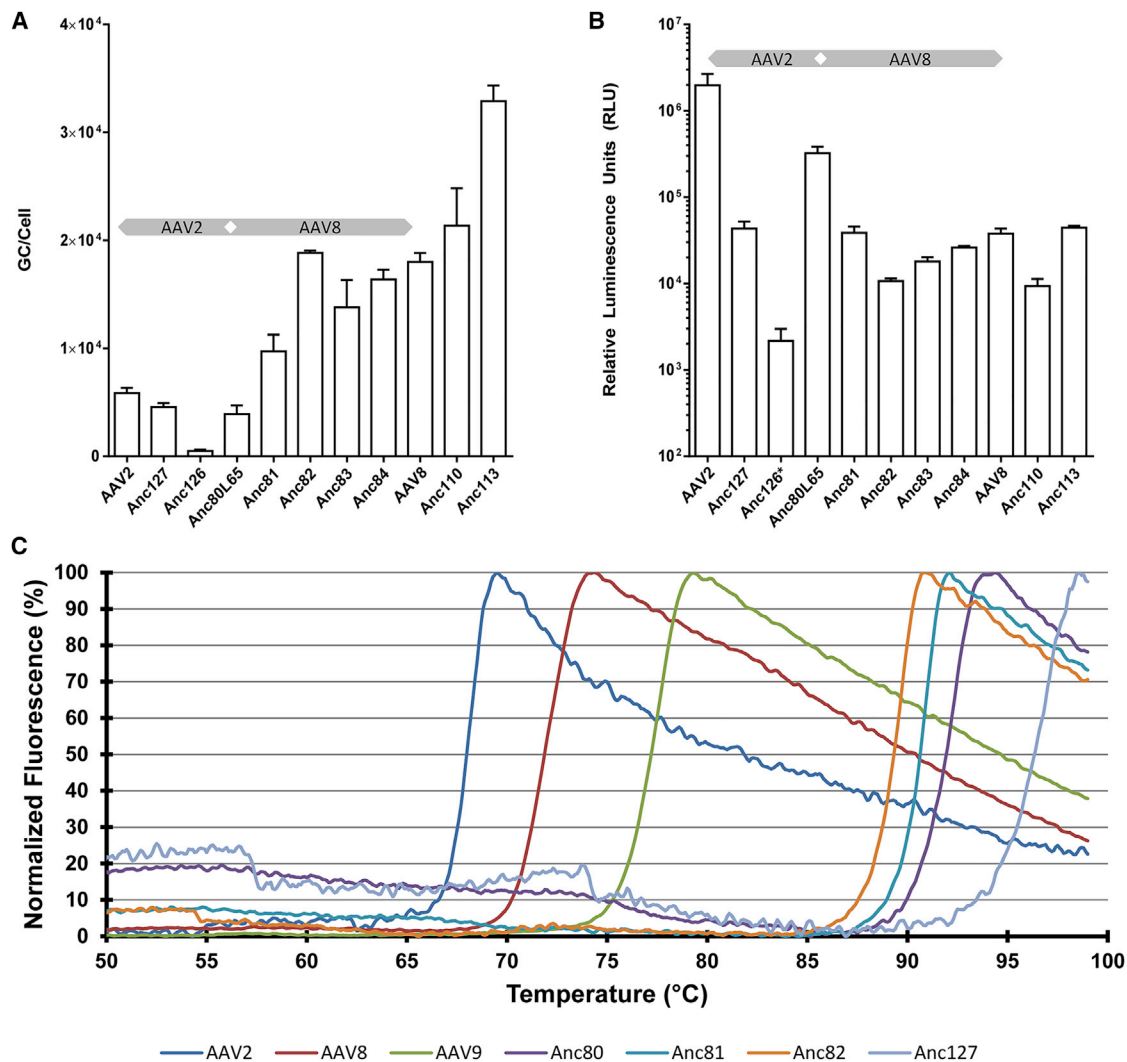
Figure 6. AAV Lineage Reconstruction Modulates Production, Infectivity, and Thermostability

(A) Production of nine ancestral and two extant viral vectors containing a luciferase reporter gene driven by a CMV promoter determined by qPCR. Error bars represent SD of three biological replicates.

(B) Ancestral and extant viral vectors were used to transduce HEK293 cells at a particle to cell ratio of 1.9×10^3. Error bars represent SD of three distinct lots of vector. *Anc126 was added at ratios between 2.1×10^2 and 3.5×10^2 GCs/cell due to low vector yield. Gray diverging arrows in (A) and (B) schematically illustrate AAV2 and AAV8 lineage phenotypic evolution.

(C) Sypro orange thermostability assay indicating denaturation temperatures of selected ancestral and extant AAV vectors. See also Figure S2.

viral-like and AAV-like particle, via structural and biochemical measures (Figures 2 and 3) and 8.6% or 64 amino acids divergent from the closest known AAV sequence. Anc80L65 furthermore demonstrates high stability, unlike any of the tested extant descendants. Anc80L65 is infectious (Figure 6B), and it is efficient as a gene transfer vehicle in murine models for retina, liver, and muscle targeting and in non-human primates for liver-directed gene transfer (Figure 4). Its utility as a gene transfer vector was further evaluated in a set of safety studies that did not illustrate vector-related toxicity concerns (Tables S1–S7).

An important property of a gene transfer vector for clinical utility is its ability to be produced efficiently. Vector yields are a function of the production system and innate structural assembly

efficiencies. Anc80L65, exceeding our expectation, yields GCs in approximately equivalent ranges to AAV2, the most studied clinical AAV technology to date. AAV2 is reduced compared to other high-yielding AAVs (Figure 6A). In an attempt to understand the variation in particle yield, we tested the hypothesis that this is due to reduced stability. In contrast, Anc80L65 was thermostable up to temperatures of 92°C, compared to 68° and 72°C for AAV2 and 8. It is possible that this high thermostability suggests that a higher activation energy threshold is needed for Anc80L65 to dissociate and, possibly inversely, assemble. Further work is required to investigate this fascinating relationship. Another possibility relates to what can be referred to as the *"Back to the Future"* concern; here, we present data on an

attempt to take the ancestral state of a single virion component and bring it to life in an otherwise fully contemporary context. This process conceptually ignores the possibility that co-evolution took place with other AAV components (e.g., AAP, ITR), helper functions provided by adenovirus or other helper virus functions, and the host cell. If indeed such co-evolution occurred, then those heterologous ancestral components may be required to bring full functionality to Anc80L65 or any ancestral AAV capsomer. To date, no data are available to suggest such a missing ancestral link is prohibiting further potentiation of Anc80L65 in its assembly, packaging, or transduction biology.

Building on the validation that Anc80L65 brought to our methodology, we expanded the set of evolutionary intermediates for experimentation to eight additional Anc particles along the lineage toward AAV1–3 and 7–9 (Figure 1). We demonstrate assembly, packaging, and in vitro infectivity of these particles. Particularly along the AAV8 branching from Anc80L65, our data suggest an increased productivity of viral yields illustrating that close putative descendants of Anc80L65 were able to overcome the production limitation discussed above (Figure 6).

In addition to ASR providing a novel methodology to synthetically derive viral vectors with novel biology, as demonstrated for Anc80L65, the availability of functional intermediates of divergent extant viruses may enable the elucidation of relevant questions to the unique biology of the distinct AAV serotypes. In turn, these data may ultimately make possible structure-guided design of AAV and tailor it to the specific requirement of a clinical application. Our approach may furthermore shed light on the evolutionary pressures of AAV along its modeled natural history. One such presumed evolutionary pressure is of highest concern for gene therapy applications. Given that AAV is widely circulating in humans, long-lived memory T and B cell responses pose a problem for gene therapy across broad populations. Indeed, individuals with pre-existing immunity (PEI) to AAV (over 50% in some populations) are currently excluded from participating in many AAV clinical trials, as no adequate alternative mitigation strategies exist to address the problem (Louis Jeune et al., 2013). Here, we demonstrate that ASR is able to modulate known epitopes that govern B and T cell PEI to novel antigenic sites and thereby possibly reduce the affinity of MHCI-TCR and virus-NAB interactions (Figure 5). Unlike rabbit vaccination experiments as in Figure 5A, human PEI is highly cross-reactive across many serotypes, making it difficult for any AAV-like particle escape this wide breadth PEI in humans (Brantly et al., 2009; Calcedo et al., 2009). To evaluate seroprevalence of Anc vectors and how these data inform a gene therapy outcome, a more robust and predictive assay is required as are extensive seroprevalence studies. Indeed, our data highlight traditional NAB assays poorly correlating with in vivo neutralization (Figures 4D and 5B), as was previously noted (Wang et al., 2010).

In conclusion, ASR is shown to be a powerful methodology to generate functional intermediates of complex and structurally constrained biological assemblies, here exemplified for an icosahedral virus, AAV. The ancestral reconstruction of Anc80L65, the common putative ancestor of AAV1–3 and 7–9, yielded a highly potent vector particle with potential use in gene therapy applications via in silico and synthetic biology methods. The resolution of the lineage from Anc80 toward these

extant serotypes provides a toolset of gene transfer reagents that further can help elucidate complex structure-function relationships within AAVs and eventually may facilitate structure-based design of this potential new class of genetic drugs.

EXPERIMENTAL PROCEDURES

Ancestral Sequence Reconstruction of AAV Capsids

These studies were conducted following approval of the Harvard Medical School Institutional Biosafety Committee (Committee on Microbiological Safety), which included considerations of dual use of the vector reagents described here. Ancestral capsid sequences were reconstructed using maximum-likelihood methods as in Finnigan et al. (2012). An alignment of 75 AAV capsids (GenBank accession numbers in the Supplemental Experimental Procedures) was generated using PRANK v.121002 using the −F option (Löytynoja and Goldman, 2005, 2008) and the JTT+F+G model was determined to be the phylogenetic model of best fit through the Aikake Information Criterion as implemented in ProtTest3 (Darriba et al., 2011). The alignment and best-fit model were then used to infer a phylogeny through PhyML 3.0 (Guindon et al., 2010), which was evaluated through the approximate likelihood-ratio test (aLRT) (Anisimova and Gascuel, 2006) as implemented in PhyML. Ancestral capsid sequences were then inferred with PAML 4.6 (Yang, 2007) through the Lazarus package developed by the Thornton group.

In order to compensate for the uncertainty inherent to the reconstruction, a script was written to assess the computed posterior probabilities to identify ambiguously reconstructed sites. Positions were considered ambiguous along the reconstructed capsid sequence if more than one amino acid had a posterior probability greater than 0.3. Eleven such sites were identified on Anc80, each with two probable amino acids. These 11 dimorphic sites were then incorporated into a DNA library using the codons from a modern virus (rh.10). Because the reconstruction did not consider the coevolution of AAP and the capsid, the AAP open-reading frame (ORF) was ablated by changing the non-canonical CTG start codon to CAG during library design. In addition, another downstream ATG also in the AAP ORF was ablated by changing the codon to AAG. These modifications did not alter the amino acids in the *cap* ORF. The DNA library was then synthesized by DNA2.0 and subsequently subcloned into expression vectors via restriction enzyme digest and ligation.

In Vitro Characterization of AAV Ancestral Lineage Vectors

To identify and characterize functional AAV capsids within Anc80Lib, individual clones from the subcloned DNA library were isolated and used to produce luciferase-containing vector in 6 or 96 wells with AAP2 provided in *trans*. Crude vector was isolated by filtering cell lysate through a 0.4-μm filter after 48 hr had elapsed since transfection. Next, equal volumes of this crude vector preparation were added to 96-well plates confluent with HEK293 cells, which were evaluated for their luciferase activity an additional 48 hr later. In total, 776 clones were evaluated. Crude preparations of vector containing a cytomegalovirus-driven luciferase were produced by triple transfection in a six-well format, supplementing AAP in *trans* to ancestral AAV vectors. In total, three different independent biological replicates were produced per vector. DNase-I-resistant transgenes were quantified as above. These crude preparations of virus were each then evaluated for their ability to transduce HEK293 cells in technical triplicates at an MOI of 1.9E3 GCs/cell with the exception of Anc126, which was added at MOIs between 2.1E2 and 3.5E2 GCs/cell. After 48 hr had elapsed, the transduced cells were assessed for luciferase via luminescence assay.

AAV Vector Preparation

Large-scale polyethylenimine transfections of AAV *cis*, AAV *trans*, and adenovirus helper plasmid were performed in a ten-layer hyperflask (Corning) with near-confluent monolayers of HEK293 cells. Plasmids were transfected at a ratio of 2:1:1 (260 μg of adenovirus helper plasmid/130 μg of *cis* plasmid/130 μg of *trans* plasmid). Transfections for production of Anc vectors were supplemented with pAAP2 in amounts equivalent to the AAV *cis* plasmid. PEI Max (Polysciences)/DNA ratio was maintained at 1.375:1 (w/w). The transfection and downstream purification process were performed as previously

described (Lock et al., 2010). DNase-I-resistant vector genomes copies were used to titrate AAV preparations by TaqMan qPCR amplification (Applied Biosystems 7500, Life Technologies) with primers and probes detecting promoter, transgene, or poly-adenylation signal coding regions of the transgene cassette. The purity of the large-scale preparations was evaluated by SDS-PAGE gel electrophoresis.

In Vivo Gene Transfer
Mouse
C57BL/6 male mice (6–8 weeks old) were purchased from Charles River Laboratories and kept at the Schepens Eye Research Institute (SERI) Animal Facility. All animal procedures were performed in accordance with protocols approved by the institutional animal care and use committees at SERI.

For liver-directed studies, male C57BL/6 mice were each injected retro-orbitally (150 μl) or intraperitoneally (100 μl). Blood was collected via submandibular bleeds using GoldenRod animal lancets (MEDIpoint). Animals were euthanized, and livers were collected. Expression of lacZ was assessed using a β-galactosidase staining kit (Life Technologies, catalog no. K1465-01). In skeletal muscle studies, male C57BL/6 mice were injected into the rear-right gastrocnemius muscle (50 μl). Retinal studies were performed by subretinal injection of male C57BL/6 mice (2 μl).

Non-human Primate
Experiments with rhesus monkeys were performed at New England Primate Research Center (NEPRC). All experimental procedures were approved by the Office for Research Subject Protection, Harvard Medical Area Standing Committee on Animals, the Harvard Medical School Institutional Animal Care and Use Committee. Animals were sedated with ketamine or telazol in combination with dexdomitor. AAV was administered intravenously via the saphenous vein in a 20 ml volume at an approximate rate of 1 ml/min. After recovering from the injection, the animals were monitored clinically for general well-being and followed for 2 months. During this time, phlebotomies were performed at regular intervals to evaluate immune response to AAV and toxicity. After 70 days, monkeys were euthanized and liver samples were harvested. Tissue was harvested in RNALater (Life Technologies).

Additional details are provided in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The Anc AAV capsids reported in this paper have been deposited to the Genbank and are available under accession number Genbank: KT235804–KT235812.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.07.019.

AUTHOR CONTRIBUTIONS

E.Z. conceived, designed, and conducted in silico and in vitro experiments and helped with the drafting of the manuscript. S.P. and A.C.M. conceived and performed structural modeling and thermostability studies. V.K., H.T.T., R.S., S.S., R.X., E.P., and L.S.C. conducted various aspects of the immunological and/or in vivo experiments. E.A.-M. and E.P. contributed to the manuscript and scientific management of the program. L.H.V. conceived and supervised the scientific program and wrote the manuscript.

ACKNOWLEDGMENTS

REFERENCES

Adachi, K., Enoki, T., Kawano, Y., Veraz, M., and Nakai, H. (2014). Drawing a high-resolution functional map of adeno-associated virus capsid by massively parallel sequencing. Nat. Commun. 5, 3075.

Allocca, M., Mussolino, C., Garcia-Hoyos, M., Sanges, D., Iodice, C., Petrillo, M., Vandenberghe, L.H., Wilson, J.M., Marigo, V., Surace, E.M., and Auricchio, A. (2007). Novel adeno-associated virus serotypes efficiently transduce murine photoreceptors. J. Virol. 81, 11372–11380.

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst. Biol. 55, 539–552.

Asuri, P., Bartel, M.A., Vazin, T., Jang, J.H., Wong, T.B., and Schaffer, D.V. (2012). Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. Mol. Ther. 20, 329–338.

Auricchio, A., Kobinger, G., Anand, V., Hildinger, M., O'Connor, E., Maguire, A.M., Wilson, J.M., and Bennett, J. (2001). Exchange of surface proteins impacts on viral vector cellular specificity and transduction characteristics: the retina as a model. Hum. Mol. Genet. 10, 3075–3081.

Bainbridge, J.W., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N., et al. (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. N. Engl. J. Med. 358, 2231–2239.

Bartel, M.A., Weinstein, J.R., and Schaffer, D.V. (2012). Directed evolution of novel adeno-associated viruses for therapeutic gene delivery. Gene Ther. 19, 694–700.

Bell, C.L., Gurda, B.L., Van Vliet, K., Agbandje-McKenna, M., and Wilson, J.M. (2012). Identification of the galactose binding domain of the adeno-associated virus serotype 9 capsid. J. Virol. 86, 7326–7333.

Brantly, M.L., Chulay, J.D., Wang, L., Mueller, C., Humphries, M., Spencer, L.T., Rouhani, F., Conlon, T.J., Calcedo, R., Betts, M.R., et al. (2009). Sustained transgene expression despite T lymphocyte responses in a clinical trial of rAAV1-AAT gene therapy. Proc. Natl. Acad. Sci. USA 106, 16363–16368.

Bryant, L.M., Christopher, D.M., Giles, A.R., Hinderer, C., Rodriguez, J.L., Smith, J.B., Traxler, E.A., Tycko, J., Wojno, A.P., and Wilson, J.M. (2013). Lessons learned from the clinical development and market authorization of Glybera. Hum. Gene Ther. Clin. Dev. 24, 55–64.

Calcedo, R., Vandenberghe, L.H., Gao, G., Lin, J., and Wilson, J.M. (2009). Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses. J. Infect. Dis. 199, 381–390.

Dalkara, D., Byrne, L.C., Klimczak, R.R., Visel, M., Yin, L., Merigan, W.H., Flannery, J.G., and Schaffer, D.V. (2013). In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous. Sci. Transl. Med. 5, 189ra76.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27, 1164–1165.

DiMattia, M.A., Nam, H.J., Van Vliet, K., Mitchell, M., Bennett, A., Gurda, B.L., McKenna, R., Olson, N.H., Sinkovits, R.S., Potter, M., et al. (2012). Structural insight into the unique properties of adeno-associated virus serotype 9. J. Virol. 86, 6947–6958.

Finnigan, G.C., Hanson-Smith, V., Stevens, T.H., and Thornton, J.W. (2012). Evolution of increased complexity in a molecular machine. Nature 481, 360–364.

Gao, G., Alvira, M.R., Somanathan, S., Lu, Y., Vandenberghe, L.H., Rux, J.J., Calcedo, R., Sanmiguel, J., Abbas, Z., and Wilson, J.M. (2003). Adeno-associated viruses undergo substantial evolution in primates during natural infections. Proc. Natl. Acad. Sci. USA 100, 6081–6086.

Gao, G., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X., and Wilson, J.M. (2004). Clades of Adeno-associated viruses are widely disseminated in human tissues. J. Virol. 78, 6381–6388.

Gao, G., Vandenberghe, L.H., and Wilson, J.M. (2005). New recombinant serotypes of AAV vectors. Curr. Gene Ther. 5, 285–297.

Girod, A., Wobus, C.E., Zádori, Z., Ried, M., Leike, K., Tijssen, P., Kleinschmidt, J.A., and Hallek, M. (2002). The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. J. Gen. Virol. 83, 973–978.

Grieger, J.C., Snowdy, S., and Samulski, R.J. (2006). Separate basic region motifs within the adeno-associated virus capsid proteins are essential for infectivity and assembly. J. Virol. 80, 5199–5210.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Gullberg, M., Tolf, C., Jonsson, N., Mulders, M.N., Savolainen-Kopra, C., Hovi, T., Van Ranst, M., Lemey, P., Hafenstein, S., and Lindberg, A.M. (2010). Characterization of a putative ancestor of coxsackievirus B5. J. Virol. 84, 9695–9708.

Gurda, B.L., Raupp, C., Popa-Wagner, R., Naumer, M., Olson, N.H., Ng, R., McKenna, R., Baker, T.S., Kleinschmidt, J.A., and Agbandje-McKenna, M. (2012). Mapping a neutralizing epitope onto the capsid of adeno-associated virus serotype 8. J. Virol. 86, 7739–7751.

Gurda, B.L., DiMattia, M.A., Miller, E.B., Bennett, A., McKenna, R., Weichert, W.S., Nelson, C.D., Chen, W.J., Muzyczka, N., Olson, N.H., et al. (2013). Capsid antibodies to different adeno-associated virus serotypes bind common regions. J. Virol. 87, 9111–9124.

Hastie, E., and Samulski, R.J. (2015). Adeno-associated virus at 50: a golden anniversary of discovery, research, and gene therapy success–a personal perspective. Hum. Gene Ther. 26, 257–265.

Huttner, N.A., Girod, A., Perabo, L., Edbauer, D., Kleinschmidt, J.A., Büning, H., and Hallek, M. (2003). Genetic modifications of the adeno-associated virus type 2 capsid reduce the affinity and the neutralizing effects of human serum antibodies. Gene Ther. 10, 2139–2147.

Jacobson, S.G., Cideciyan, A.V., Ratnakaram, R., Heon, E., Schwartz, S.B., Roman, A.J., Peden, M.C., Aleman, T.S., Boye, S.L., Sumaroka, A., et al. (2012). Gene therapy for leber congenital amaurosis caused by RPE65 mutations: safety and efficacy in 15 children and adults followed up to 3 years. Arch. Ophthalmol. 130, 9–24.

Kern, A., Schmidt, K., Leder, C., Müller, O.J., Wobus, C.E., Bettinger, K., Von der Lieth, C.W., King, J.A., and Kleinschmidt, J.A. (2003). Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. J. Virol. 77, 11072–11081.

Knipe, D.M., and Howley, P.M. (2013). Fields Virology, Sixth Edition (Wolters Kluwer/Lippincott Williams & Wilkins Health).

Koerber, J.T., Jang, J.H., and Schaffer, D.V. (2008). DNA shuffling of adeno-associated virus yields functionally diverse viral progeny. Mol. Ther. 16, 1703–1709.

Li, W., Asokan, A., Wu, Z., Van Dyke, T., DiPrimio, N., Johnson, J.S., Govindaswamy, L., Agbandje-McKenna, M., Leichtle, S., Redmond, D.E., Jr., et al. (2008). Engineering and selection of shuffled AAV genomes: a new strategy for producing targeted biological nanoparticles. Mol. Ther. 16, 1252–1260.

Lisowski, L., Dane, A.P., Chu, K., Zhang, Y., Cunningham, S.C., Wilson, E.M., Nygaard, S., Grompe, M., Alexander, I.E., and Kay, M.A. (2014). Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. Nature 506, 382–386.

Lochrie, M.A., Tatsuno, G.P., Christie, B., McDonnell, J.W., Zhou, S., Surosky, R., Pierce, G.F., and Colosi, P. (2006). Mutations on the external surfaces of adeno-associated virus type 2 capsids that affect transduction and neutralization. J. Virol. 80, 821–834.

Lock, M., Alvira, M., Vandenberghe, L.H., Samanta, A., Toelen, J., Debyser, Z., and Wilson, J.M. (2010). Rapid, simple, and versatile manufacturing of recombinant adeno-associated viral vectors at scale. Hum. Gene Ther. 21, 1259–1271.

Louis Jeune, V., Joergensen, J.A., Hajjar, R.J., and Weber, T. (2013). Pre-existing anti-adeno-associated virus antibodies as a challenge in AAV gene therapy. Hum. Gene Ther. Methods 24, 59–67.

Löytynoja, A., and Goldman, N. (2005). An algorithm for progressive multiple alignment of sequences with insertions. Proc. Natl. Acad. Sci. USA 102, 10557–10562.

Löytynoja, A., and Goldman, N. (2008). Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science 320, 1632–1635.

MacLaren, R.E., Groppe, M., Barnard, A.R., Cottriall, C.L., Tolmachova, T., Seymour, L., Clark, K.R., During, M.J., Cremers, F.P., Black, G.C., et al. (2014). Retinal gene therapy in patients with choroideremia: initial findings from a phase 1/2 clinical trial. Lancet 383, 1129–1137.

Maguire, A.M., Simonelli, F., Pierce, E.A., Pugh, E.N., Jr., Mingozzi, F., Bennicelli, J., Banfi, S., Marshall, K.A., Testa, F., Surace, E.M., et al. (2008). Safety and efficacy of gene transfer for Leber's congenital amaurosis. N. Engl. J. Med. 358, 2240–2248.

Maheshri, N., Koerber, J.T., Kaspar, B.K., and Schaffer, D.V. (2006). Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. Nat. Biotechnol. 24, 198–204.

Mays, L.E., Wang, L., Tenney, R., Bell, P., Nam, H.J., Lin, J., Gurda, B., Van Vliet, K., Mikals, K., Agbandje-McKenna, M., and Wilson, J.M. (2013). Mapping the structural determinants responsible for enhanced T cell activation to the immunogenic adeno-associated virus capsid from isolate rhesus 32.33. J. Virol. 87, 9473–9485.

Mingozzi, F., Maus, M.V., Hui, D.J., Sabatino, D.E., Murphy, S.L., Rasko, J.E., Ragni, M.V., Manno, C.S., Sommer, J., Jiang, H., et al. (2007). CD8(+) T-cell responses to adeno-associated virus capsid in humans. Nat. Med. 13, 419–422.

Nakai, H., Thomas, C.E., Storm, T.A., Fuess, S., Powell, S., Wright, J.F., and Kay, M.A. (2002). A limited number of transducible hepatocytes restricts a wide-range linear vector dose response in recombinant adeno-associated virus-mediated liver transduction. J. Virol. 76, 11343–11349.

Nakai, H., Fuess, S., Storm, T.A., Muramatsu, S., Nara, Y., and Kay, M.A. (2005). Unrestricted hepatocyte transduction with adeno-associated virus serotype 8 vectors in mice. J. Virol. 79, 214–224.

Nathwani, A.C., Tuddenham, E.G., Rangarajan, S., Rosales, C., McIntosh, J., Linch, D.C., Chowdary, P., Riddell, A., Pie, A.J., Harrington, C., et al. (2011). Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. N. Engl. J. Med. 365, 2357–2365.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205–217.

Perabo, L., Büning, H., Kofler, D.M., Ried, M.U., Girod, A., Wendtner, C.M., Enssle, J., and Hallek, M. (2003). In vitro selection of viral vectors with modified tropism: the adeno-associated virus display. Mol. Ther. 8, 151–157.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Schnepp, B.C., and Johnson, P.R. (2014). Adeno-associated virus delivery of broadly neutralizing antibodies. Curr. Opin. HIV AIDS 9, 250–256.

Shen, X., Storm, T., and Kay, M.A. (2007). Characterization of the relationship of AAV capsid domain swapping to liver transduction efficiency. Mol. Ther. 15, 1955–1962.

Vandenberghe, L.H., Wang, L., Somanathan, S., Zhi, Y., Figueredo, J., Calcedo, R., Sanmiguel, J., Desai, R.A., Chen, C.S., Johnston, J., et al. (2006). Heparin binding directs activation of T cells against adeno-associated virus serotype 2 capsid. Nat. Med. 12, 967–971.

Vandenberghe, L.H., Breous, E., Nam, H.J., Gao, G., Xiao, R., Sandhu, A., Johnston, J., Debyser, Z., Agbandje-McKenna, M., and Wilson, J.M. (2009). Naturally occurring singleton residues in AAV capsid impact vector performance and illustrate structural constraints. Gene Ther. 16, 1416–1428.

Wang, L., Calcedo, R., Wang, H., Bell, P., Grant, R., Vandenberghe, L.H., Sanmiguel, J., Morizono, H., Batshaw, M.L., and Wilson, J.M. (2010). The pleiotropic effects of natural AAV infections on liver-directed gene transfer in macaques. Mol. Ther. 18, 126–134.

Wu, P., Xiao, W., Conlon, T., Hughes, J., Agbandje-McKenna, M., Ferkol, T., Flotte, T., and Muzyczka, N. (2000). Mutational analysis of the adeno-associated virus type 2 (AAV2) capsid gene and construction of AAV2 vectors with altered tropism. J. Virol. 74, 8635–8647.

Xie, Q., Bu, W., Bhatia, S., Hare, J., Somasundaram, T., Azzi, A., and Chapman, M.S. (2002). The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. Proc. Natl. Acad. Sci. USA 99, 10405–10410.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.