Cross-species evolution of a highly potent AAV variant for therapeutic gene transfer and genome editing

Recombinant adeno-associated viral (AAV) vectors are a promising gene delivery platform, but ongoing clinical trials continue to highlight a relatively narrow therapeutic window. Effective clinical translation is confounded, at least in part, by differences in AAV biology across animal species. Here, we tackle this challenge by sequentially evolving AAV capsid libraries in mice, pigs and macaques. We discover a highly potent, cross-species compatible variant (AAV.cc47) that shows improved attributes benchmarked against AAV serotype 9 as evidenced by robust reporter and therapeutic gene expression, Cre recombination and CRISPR genome editing in normal and diseased mouse models. Enhanced transduction efficiency of AAV.cc47 vectors is further corroborated in macaques and pigs, providing a strong rationale for potential clinical translation into human gene therapies. We envision that ccAAV vectors may not only improve predictive modeling in preclinical studies, but also clinical translatability by broadening the therapeutic window of AAV based gene therapies.
predicted by preclinical animal models. For instance, studies focused on dorsal root ganglion pathology following AAV vector treatments in NHP and pig models appear to directly correlate with AAV transduction efficiency and dose; however, such toxicologic aspects have generally not been observed in mouse models, or reported in clinical trials such as in case of SMA to date. Similarly, no vector-triggered toxicity was observed in dose escalation studies utilizing AAV8 or AAV8 vectors over-expressing micro/minidystrophin in the mdx mouse or canine models, respectively. Interestingly, other side effects such as renal toxicity and thrombocytopenia have been reported in clinical trials utilizing AAV9, but similar adverse events have not been reported to date with rh74 vectors. While the exact mechanisms related to these toxicities remain under investigation, these recent findings highlight significant disparities in AAV biology across species, but also the need for improved vector design that can (i) mitigate potential risks and improve the therapeutic window associated with AAV gene therapies as well as (ii) enable better predictive modeling leading to effective clinical translation.

To this end, several AAV library-based approaches have been developed in the past decade. For instance, directed evolution has been utilized to engineer novel AAV capsids with improved CNS or muscle tropism, improved tropism for human tissue or the ability to evade neutralizing antibodies. These efforts have yielded novel muscle or neurotropic AAV capsids such as the rationally engineered variants AAV28, AAV28G9,21, chimeric AAVB1, AAVRXR obtained through DNA shuffling of natural capsid isolates9,17, AAV-PHP.B, AAV.CAP.B10, AAVMYO, MyoAAV, amongst others obtained through capsid surface peptide display. A notable drawback that has been reported is that directed evolution of AAV libraries in a single animal model can yield variants specific for those species, as exemplified in case of AAV-PHP.B, which utilizes the C57/B6 mouse strain-specific receptor, Ly6A. Indeed, it remains to be seen whether evolving AAV libraries solely in NHP’s yields only primate sub-species-specific variants and ultimately, whether NHP-derived AAV capsids are the most suitable candidates for clinical evaluation. To address these challenges, we describe a novel cross-species-based approach that can be readily implemented into any directed evolution workflow independent of the engineering strategy or library-based approach. Specifically, we iteratively cycled AAV libraries administered intravenously and isolated amplified from CNS tissue in pigs, mice, and non-human primates to generate cross-species compatible AAVs (ccAAVs). Gene transfer efficiencies of different enriched capsid variants were determined in mice, which yielded a lead capsid variant, AAV.cc47. We benchmarked the transduction efficiencies in mouse liver, CNS, cardiac and skeletal muscle tissues in a mouse model using a Cre recombinase cassette packaged within AAV.cc47 or AAV9 and administered at three different doses as well as increased acid alpha-glucosidase (GAA) expression in the mouse CNS, heart, and skeletal muscle with AAV.cc47 compared to AAV9. Additionally, we demonstrated higher genome editing efficiencies in the mouse heart and skeletal muscle using optimized CRISPR/Cas9 cassettes in a reporter mouse model as well as the mdx mouse model for DMD. Further, we demonstrate the cross-species compatibility of AAV.cc47 compared to AAV9 in pigs, and nonhuman primates following intrathecal/cisterna magna dosing. Our approach corroborates the notion that iterative evolution in multiple animal models can yield potent AAV variants with cross-species compatibility that can potentially enable better predictive preclinical modeling and help improve the therapeutic window of AAV gene therapies in humans.

Results

Cross-species cycling of AAV capsid libraries

To engineer cross-species compatible AAVs (ccAAVs), we created AAV9-based capsid libraries and cycled them through pigs, mice, and non-human primates following systemic administration (Fig. 1a). Libraries were created through saturation mutagenesis of the AAV9 capsid surface loop corresponding to amino acids (AA) 452–458 (VP1 subunit numbering). This region corresponds to variable region IV (VR-IV; red) of the capsid, which is the farthest radially protruding surface loop, and plays a multifunctional role in AAV biology, including a neutralizing antibody binding site as well as transduction efficiency and cell surface receptor binding, due to similarity to the 3-fold-symmetry axis of the AAV9 capsid. Briefly, binding interactions of galactosylated glycans and AAVR with the AAV9 capsid are known to occur around the 3-fold axis. In addition to the functional attributes outlined above, previous studies by our laboratory22,23 and the Gradinaru laboratory20 have demonstrated the importance of VR-IV in evolving novel AAV capsids with improved functionality. Moreover, the VR-IV loop is less permissive towards peptide insert approaches compared to VR-VIII. Taken together, VR-IV is an important surface region for incorporation into AAV capsid libraries. This region is highlighted in 3D structures of the VP3 subunit monomer, trimer and full capsid conformations to highlight the protruding nature of this surface loop (Fig. 1b). Wild-type AAV libraries packaging mutant genomes consisting of AAV2 Rep and AAV9 mutant Cap, flanked by AAV2 ITRs were first cycled in 3-week old piglets dosed at 1e13 vg/kg via intravenous (IV) administration. Pig brains were harvested 3 days post injection and further dissected into parietal, occipital, and temporal cortices, cerebellum, hippocampus, thalamus, hypothalamus, and medulla regions. Viral genomes amplified from extracted DNA were pooled and cloned to generate the next AAV library, which was administered IV in 8-week-old C57/B6 mice at a dose of 2e13 vg/kg. The brains from these mice were harvested 3 days post injection and dissected into cortex, dentate gyrus, hippocampus, and cerebellum regions; then processed and prepared as above. The final species used for in vivo cycling were 2-year-old cynomolgus macaques. Libraries were administered via intravenous infusion at a dose of 3.3e13 vg/kg. Macaque brains were harvested 7 days post injection and dissected into temporal, parietal, occipital, and frontal lobes, thalamus, corpus callosum, hippocampus, cerebellum, brainstem, pons, and midbrain regions. Amplified AAV genomes from macaques as well as from each species in earlier rounds were then prepared for next-generation sequencing.

Parental and evolved viral libraries were sequenced with an Illumina NovaSeq platform. Sequencing reads were analyzed using a custom in-house Perl script and plotted in R (Fig. 1c). Each bubble corresponds to an individual library amino acid sequence and bubble diameter corresponds to fold enrichment between parental and evolved libraries. Enrichment analysis of individual AAV clones following evolution in each animal species revealed a 100 to 69,000-fold enrichment between evolved variants and parental representation. The dominant lead variant AAV.cc47, accounted for 67% of the overall mapped reads after the final round. AAV.cc47 percent representation and fold enrichment and total library diversity was determined following evolution in each species (Supplementary Fig. 1). Using monkey next-generation sequencing (NGS) data, we performed a consensus motif analysis of the top 100 AAV clones from the final round. This analysis revealed 452-GVSLGGG-458 as the major consensus sequence of the top clones, which was determined to be the amino acid sequence for AAV.cc47 (Fig. 1d). We further analyzed the NGS data of parental and evolved (monkey) libraries by calculating the amino acid frequency at each position in our 7-mer libraries. Additionally, we determined the differential representation of amino acids at each library position between the two data sets. A clear selection for our consensus motif is corroborated by our evolved variant heatmap (Fig. 1e). The parental library shows a broad distribution of amino acids across positions 1–7, with a bias for glycine at each position illustrated by the yellow gradient. The least and most frequent amino acids are shown in blue and yellow, respectively, while the 50th percentile is shown in white. In order to assess the manufacturability of AAV.cc47...
compared to AAV9, we collated titers obtained by qPCR analysis of yields from production runs of these two capsids packaging the same transgene cassettes produced in both adherent and suspension HEK293 systems. We found no significant difference in manufacturability of AAV.cc47 compared to AAV9 in either adherent or suspension production systems (Fig. 1f–g).

**AAV.cc47 is more potent than AAV9 vectors in normal mice**

We first compared the transduction profiles of AAV.cc47 and AAV9 vectors packaging a self-complementary CBh-mCherry cassette following systemic administration at 5e13vg/kg at 4 weeks post-dosing (Fig. 2a). Native mCherry fluorescence from tissue sections of injected mice show increased transduction efficiency in heart, tibialis anterior, and brain, while a modest, yet statistically significant increase in liver transduction was observed (Fig. 2b, c). Quantification of fluorescence intensity in the heart, tibialis anterior, and liver shows AAV.cc47 has a 21-, 16-, and 2-fold higher fluorescence intensity, respectively (Fig. 2d). Further, we observed increased transduction and spread of AAV.cc47 throughout the mouse brain compared to AAV9 (Fig. 2c). A greater number of neurons transduced in the cortex (CTX), cerebellum (CB), and hippocampus (HC) by the evolved variant was determined with 4-, 3-, and 4-fold greater transduction in this cell type, respectively,
AAV.cc47 vectors demonstrate improved genome editing efficiency

We then wanted to explore whether IV dosing of AAV.cc47 packaging CRISPR/Cas9 and guide RNAs can improve genome editing efficiency compared to AAV9 in Ai9 mice (Fig. 4a). To determine optimal conditions for genome editing, we evaluated a panel of vector design strategies based on a dual AAV vector approach, which requires Staphylococcus aureus (5a) Cas9 and two single guide RNAs (gRNAs) targeting the Rosa26 locus in Ai9 mice. Similar to the CR recombination approach described earlier, this CRISPR-based approach is designed to remove the lox-STOP-lox region, which will then enable tdTomato expression in different tissues. In Design 1, we used single-stranded AAV genomes with the CMV enhancer-beta-actin (CB) promoter driving SaCas9 and a different gRNA driven by the hU6 promoter; mixed in a 1:1 ratio (dose = 1e14 vg/kg each vector). Design 2 is similar to Design 1, except the vectors were mixed in a 1:3 Cas9-gRNA vector ratio (dose = 2e14 vg/kg each vector; Supplementary Fig. 6a–c); Design 3 is similar to Design 2, except that the guide RNAs are expressed within a self-complementary AAV genome and the Cas9 and guide vectors were mixed in a 1:1 ratio (dose = 2e14 vg/kg each vector; Fig. 4a). Following systemic administration, AAV.cc47-mediated CRISPR/Cas9 dual vector designs edited the Rosa26 locus at a higher efficiency than AAV9 in all 4 AAV-CRISPR designs evaluated. Notably, the self-complementary genome to deliver two single guide RNAs yielded the highest percentage of editing in the heart with a 3-fold increase in genome editing of the Rosa26 locus with AAV.cc47 compared to AAV9, determined by the number of tdTomato positive cardiomyocytes normalized to DAPI (Fig. 4b). We also observed a 3-fold increase in genome editing in titilis anterior with AAV.cc47 compared to AAV9 as shown by the increased number of tdTomato expressing myofibers (Fig. 4c). Consistent with reporter gene and Cre expression studies, we observed no significant fold change in genome editing efficiencies in hepatocytes with AAV.cc47 and AAV9 packaging any of the vector designs in the liver (Fig. 4d).

Vector genome biodistribution corroborated in vivo studies in Cre studies, showing no difference in heart, titilis anterior, and liver tissues between AAV.cc47 and AAV9 (Supplementary Fig. 6d). Interestingly, gRNA expression was 3- to 4-fold greater in all three tissues of AAV.cc47 treated mice compared to AAV9 (Fig. 4e). Based on these observations, it is tempting to speculate that AAV.cc47 may not have a binding or uptake advantage in tissues, but may benefit from a post-entry event compared to AAV9. SaCas9 expression profile was different and relatively low, and in some cases below the limit of detection, compared to gRNA expression (Fig. 4f).

AAV.cc47 transduces NHP brain and heart at higher efficiency than AAV9

To assess cross-species compatibility, AAV9 or AAV.cc47 vectors packaging a scCBh-mCherry cassette were dosed at 1e13 vg/kg via intracisterna magna (ICM) infusions in two-year-old cynomolgus macaques and harvested tissue 13 days post-injection (Fig. 5a). Immunohistochemical staining against mCherry in NHP brain revealed robust expression and spread in the monkey cortex and cerebellum with AAV.cc47 compared to AAV9, while penetration into the brain parenchyma was also greater in AAV.cc47-treated monkeys (Fig. 5b). Neuronal and glial transduction in the cortex was also higher with our evolved variant compared to AAV9 and a greater number of mCherry positive Purkinje neurons were observed in AAV.cc47 treated cerebellums (Fig. 5b). Expression was observed in peripheral tissues outside the access of the CSF, such as the heart, where AAV.cc47 demonstrated enhanced transduction compared to AAV9, and the liver (Fig. 5c, d). Biochemical analysis of tissues further corroborates the enhanced transduction phenotype seen in histology of NHP tissues.
with AAV.cc47 (Supplementary Fig. 7a–f). This was most notable in the premotor cortex, where mCherry protein levels quantified by ELISA show a 10-fold higher protein expression in this brain region with AAV.cc47 compared to AAV9. Equal vector genome biodistribution between both capsids was further corroborated in NHPs following ICM dosing (Supplementary Fig. 7a, d). Images of tissue from both AAV-injected NHPs have been provided (Supplementary Figs. 8–10).

Analysis of the lumbar and thoracic spinal cord regions by immunohistochemical staining of mCherry reveals greater transduction of cell types in these two spinal cord regions with AAV.cc47, while DRG expression was observed for both AAV9 and AAV.cc47 (Supplementary Fig. 10).

We wanted to evaluate the transduction of AAV.cc47 in other species as well following intraCSF dosing. Briefly, PO-1 C57/B6
neonates were injected by intracerebroventricular (ICV) administration of 2e10vg of AAV.cc47 or AAV9. Vectors packaged a self-complementary genome containing CBh-mCherry, and mice were sacrificed 4 weeks post injection (Supplementary Fig. 11a). Immunofluorescence staining of brain cross-sections was performed with an anti-rabbit mCherry primary antibody and show increased transduction and spread throughout the mouse brain of AAV.cc47 treated mice compared to AAV9, especially in the cortex, hippocampus, and cerebellum, (Supplementary Fig. 11b). We also evaluated transduction in the pigs following intraCSF dosing of AAV.cc47 or AAV9 vectors packaging a self-complementary genome containing CBh-eGFP. Briefly, we administered 3.3e13vg via intrathecal (IT) infusion into the lumbar cistern of 3-week-old piglets and harvested tissue 13 days post injection (Supplementary Fig. 11c). Pig brains were immunohistochemically stained with an anti-rabbit eGFP primary antibody and corroborate the findings from our NHP studies of increased transduction and spread throughout the cortex and cerebellum of AAV.cc47 treated mice (Supplementary Fig. 11d). A markedly greater number of neuronal and glial cells transduced in these brain regions with AAV.cc47 compared to AAV9 was observed. These results collectively illustrate that AAV.cc47 is a cross-species compatible vector capable of greater transduction in the brain of mice, pig, and NHPs.

**AAV.cc47 demonstrates increased GAA transgene expression in a mouse model of Pompe disease**

Following the evaluation of AAV.cc47 in wild-type mice and demonstrated cross-species compatibility, we investigated the ability of AAV.cc47 vectors to deliver therapeutic cargo in a mouse model of disease. To this end, AAV.cc47 and AAV9 vectors were compared in the acid alpha-glucosidase knockout (Gaa−/−) mouse model of Pompe disease. Mice were dosed IV (1.3e14vg/kg) with AAV vectors packaging a single-stranded genome encoding a functional copy of GAA driven by the CBh promoter (Fig. 6a). Analysis of GAA enzyme levels in brain tissues at 4 weeks post-dosing not only revealed that AAV.cc47-GAA is superior to AAV9-GAA vectors, but also enabled enzyme activity levels to reach 67% of WT mice compared to only 16% of WT levels in AAV9-GAA treated mice (Fig. 6b). In different segments of the spinal cord, GAA activity from AAV.cc47-GAA treated mice were nearly identical to that of WT mice, and activity from AAV9-GAA treated mice was only 26% of WT (Supplementary Figure 12a). In mouse heart and tibialis anterior, only AAV.cc47-GAA achieved significant increases in enzyme activity over all three other groups evaluated, Gaa (−/−) mock, AAV9-GAA, and WT (Fig. 6b). Consistent with earlier results, enzyme activity assays in the liver were nearly equal between AAV.cc47-GAA and AAV9-GAA vectors with both expressing supraphysiological levels (Fig. 6b). No statistical difference was observed in vector genome copy numbers (biodistribution) between AAV.cc47 and AAV9 in any of these tissues (Fig. 6c, Supplementary Fig. 12b).

**AAV.cc47 mediated gene editing demonstrates improved dystrophin restoration in the mdx mouse model**

Having determined an optimal dual AAV-CRISPR cassette design strategy consisting of a single-stranded vector expressing CMV-SaCas9 and a self-complementary gRNA vector, we further exploited this approach to edit the mouse Dmd locus. Briefly, we replaced the gRNAs targeting the ROSA26 locus with gRNAs targeting introns 22 and 23 of Dmd. CRISPR/gRNA cassettes packaged in AAV.cc47 or AAV9 were administered IV at 1.4e14 vg/kg each vector (1:1 ratio) in Pax7-nGFP mdx mice (Fig. 6d). These mice harbor a premature stop codon in exon 23 which prevents normal expression of dystrophin and upon excision of exon 23, full-length dystrophin expression is restored. In the mdx mouse heart, AAV.cc47-CRISPR treatment results in 6-fold higher dystrophin expression, in relation to laminin expression, compared to AAV9 (Fig. 6e, g). In the tibialis anterior of mdx mice, AAV.cc47-CRISPR treatments result in a 2-fold higher restoration of dystrophin than AAV9-treated animals (Fig. 6f, g). Genome editing outcomes from both vectors were determined via Taqman quantitative RT-PCR by quantifying the percent of exon 23 deleted Dmd transcripts in each tissue. AAV.cc47 mediated genome editing resulted in 3- and 5-fold more exon 23 deleted transcripts in heart and tibialis anterior, respectively, compared to AAV9 (Fig. 6h). Quantification of vector genomes and gRNA expression corroborate our previous findings that there is no measurable difference in vector genome biodistribution to tissues, while increased gRNA expression is observed in AAV.cc47 treated tissue (Supplementary Fig. 13).

**Discussion**

Multiple CNS, cardiac and skeletal muscle-focused gene therapies being evaluated in the clinic utilize AAV9 vectors. The widespread adoption of this vector to deliver therapeutic transgenes is due in part to the broad tissue tropism and abundance of transduction data in different preclinical models. Previous efforts by our lab have demonstrated the plasticity of capsid surface spike regions, particularly VR-IV and VR-VIII, through structure-guided evolution of different AAV serotypes to achieve improved transduction, tropism, and antibody evasion. In this study, we markedly improved the transduction profile of AAV9 by evolving capsid libraries generated by saturation mutagenesis of VR-IV residues in three different animal models (pigs, mice, and non-human primates). As alluded to earlier, this multi-species approach was designed to specifically address the concern that newly evolved/engineered AAV variants may not readily translate from mouse models of disease to large animal models for toxicity/biodistribution studies and ultimately to humans in a scalable and predictable fashion.

At the structural level, we observed a striking shift toward amino acids with hydrophobic or no side chains in different positions N452G, G435V, S454S, G455L, Q456G, N457G, and Q458G. The newly evolved 452-GVSLGGG-458 did not impact capsid stability based on production yields with a range of cassette. Ongoing efforts to further understand...
Fig. 3 | AAV.cc47 outperforms AAV9 in Cre recombination in the Ai9 reporter mice at all tested doses following intravenous (IV) administration. A 8–10 week old Ai9 reporter mice (n = 4) were injected intravenously at 4e11, 4e12, and 4e13 vg/kg with AAV9 (red) or AAV.cc47 (blue). Vectors delivered a single stranded AAV genome with the cytomegalovirus (CMV) promoter driving Cre recombinase. Representative immunofluorescence images and quantification for tdTomato (red) and DAPI (blue) for mice injected at each dose in heart (b), tibialis anterior (c), and liver (d). Fifty micrometers thick cross sections were obtained via a vibratome for heart and liver, while 7 μm thick cross sections were obtained for skeletal muscle via cryostat. Total number of tdTomato+ cells (heart and liver) and total number of tdTomato+ myofibers (tibialis anterior) were counted and normalized to the total number of DAPI+ nuclei in each tissue. e Vector genome copy numbers per μg DNA were calculated by normalizing Cre recombinase copy numbers to total μg DNA input for qPCR quantification and are represented as log vg/μg DNA. Each dot represents an individual mouse, fold change is listed above significance, the dash and bars represent the mean value and error bars represent the standard error mean. Quantifications in panels B-E are listed in ascending order of dosing from left to right, starting with 4e11 vg/kg and going to 4e13 vg/kg for each tissue type. Statistical significance was determined by One-Way ANOVA with Tukey’s post-test (3 treatment groups) for vector genome biodistribution analysis (High dose heart, P < 0.0279). An unpaired two-tailed Student t Test (2 treatment groups) was used for histological quantification (High dose heart, P < 0.0001; tibialis anterior, P < 0.0001; liver, P < 0.001; Mid dose heart, P < 0.0001; tibialis anterior, P < 0.0023; liver, P < 0.2666; low dose heart, P < 0.0006; tibialis anterior, P < 0.0001; liver, P < 0.0019). *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001; ns not significant. Source data are provided as a Source Data file.
other AAV.cc47 attributes also include assessment of antigenicity as determined by neutralizing antibody assays. It is important to note that the AAV9 VR-IV surface loop has been previously implicated as an immunodominant epitope in antibody recognition. Assessment of the immunogenicity of AAV.cc47 compared to AAV9 will require a large NHP cohort in a separately designed study. We have, however, evaluated the ability of AAV.cc47 to evade neutralizing antibodies (NAbs) from pig, NHP and human antisera (Supplementary Figure 14a–c). Our studies indicate that AAV.cc47 is neutralized similar to AAV9 and corroborates earlier observations that the newly evolved antigenic epitope in VR-IV alone is not sufficient for NAb evasion. However, it is important to note that, unlike the parental AAV9 surface epitope VR-IV, this newly evolved epitope is devoid of charged or aromatic surface exposed side chains, thereby presenting minimal features generally thought to be essential for receptor or immune interactions.

The cross-species evolution approach revealed several interesting trends. At a macro level, AAV library cycling leads to decreased...
Fig. 4 | Systemic administration of AAV.cc47-CRISPR increases genome editing efficiencies in heart and skeletal muscle of A9 reporter mice compared to AAV9. a 8–10 week old A9 reporter mice (n = 4) were injected with AAV.cc47 (blue) or AAV9 (red) using our lead dual AAV vector system consisting of a single stranded genome with cytomegalovirus (CMV) driven SaCas9 and a self-complementary genome with 2 gRNAs driven by U6 promoters targeting the Rosa26 locus. The total dose administered was 2.0e14 vg/kg (1.0e14 vg/kg each vector). Representative immunofluorescence images for tdTomato (red) and histological quantification in heart (b), skeletal muscle (c), and liver (d). Fifty micrometers thick cross sections were obtained via a vibratome for the heart and liver, while 7 μm thick cross sections were obtained for skeletal muscle via cryostat. Quantification of the percentage of tdTomato+ cells (heart and liver) and myofibers (tibialis anterior) was normalized to the number of DAPI+ nuclei in each tissue. For all histological quantification, 2 sections per mouse and 3 images per section (total of 6 images) were used to quantify % transduced cells and myofibers using ImageJ. Quantitative RT-PCR was performed with primers amplifying each gRNA (e) or SaCas9 (f) mRNA relative to mouse Actb. Each dot represents an individual mouse, fold change is listed above significance, the bar graphs represent the mean value and error bars represent the standard error mean. Statistical significance was determined by a two-tailed Student t Test (2 treatment groups; heart, P < 0.0009; tibialis anterior, P < 0.0227; liver, P < 0.2892) or One-Way ANOVA with Tukey’s posttest (3 treatment groups; heart gRNA, P < 0.0314, Cas9, P < 0.7966; tibialis anterior gRNA, P < 0.0348, Cas9, P < 0.4459; liver gRNA, P < 0.0126, Cas9, P < 0.8205). *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001; ns not significant. Source data are provided as a Source Data file.

Fig. 5 | AAV.cc47 transduces non-human primate (NHP) brain and heart more effectively than AAV9 following intracisternal magna (ICM) infusion. a ICM infusion of 1e13 total vg of AAV9 or AAV.cc47 vectors in 2-year-old cynomolgus macaques (n = 2). Organs were harvested at 13 days post infusion. Vectors packaging self-complementary AAV genome with a truncated chicken beta-actin (CBh) promoter driving mCherry expression were utilized for the study. Representative images of immunohistochemically stained NHP brain (b), heart (c), and liver (d) are shown. Tissue was embedded in paraffin following fixation and cut 5 μm thick and 3,3’-Diaminobenzidine (DAB) was used to visualize mCherry protein in tissues. Source data are provided as a Source Data file.
sequence diversity in each cycling step. Whether this phenomenon arises from the number of cycles alone or is also influenced by a specific animal model is unclear. Nonetheless, it is interesting to note that the percent representation of conserved amino acid residues for each position in VR-IV increases, particularly after cycling in NHPs. In particular, we observed significant enrichment of neutral amino acid residues with no to small side chains (esp. glycine, alanine, valine) in VR-IV as we cycled across each species. The final NHP cycling step markedly enriched Gly, Ser and Leu residues within VR-IV. Amongst these, the Gly residues were preferably enriched in multiple species. Novel receptor or host factor usage and other functional ramifications of these structural changes are as yet unknown. It is, however pertinent to mention that AAV.cc47 is still AAVR-dependent for cellular infection (similar to AAV9; Supplementary Fig. 1d).

High-throughput sequencing data revealed that the AAV.cc47 variant started to emerge as a promising candidate after evolving the
AAV library in pigs, stringent selective pressure on our AAV capsid libraries was also observed in the non-human primate brain. This was evidenced not only by percent representation but also by the fold enrichment of AAV.cc47 compared to the parental library following evolution in this species. We acknowledge that the current study only demonstrates improved potency of AAV.cc47 in CNS, cardiac and liver tissue in NHPs, when administered through the cisterna magna. These observations were recapitulated in the pig brain, where due to their size and amount of vector required, the animals were dosed intravenously as well. Thus, a critical aspect of AAV.cc47 biology that we continue to explore is the ability to transduce multiple tissues at high efficiency as well as toxicity studies in large animal models following different routes of administration (IV, intracoronary etc.).

AAV.cc47 demonstrated superior gene transfer efficiency in multiple organs compared to AAV9 in three different mouse models as evidenced by reporter gene expression, Cre recombinase and GAA enzyme activity. Striking differences were observed in the brain, heart, and tibialis anterior, while the liver appeared to demonstrate saturated transgene expression at the dosages tested, corroborating broad tissue tropism similar to AAV9. These results corroborate that AAV.cc47 can potentially improve the therapeutic window for certain gene therapy applications. Surprisingly, no difference in vector copy number or tissue biodistribution were observed between the two AAV vectors. Since our evolution approach did not select for capsids with improved tropism, but rather enhanced transduction efficiency, the comparable biodistribution of AAV.cc47 and AAV9 in all tissues and in different species is thought-provoking. It is also noteworthy to mention that we observed improved transgene expression, but similar biodistribution regardless of the cassette (ss vs sc) or promoter utilized. At this writing, it is tempting to speculate that AAV.cc47 has evolved to exploit a post-entry mechanism that may account for increased potency, while maintaining tissue uptake similar to AAV9. For instance, intracellular trafficking steps such as improved endosomal escape, nuclear entry, capsid uncoating-genome release and/or transcriptional efficiency could contribute to the observed capsid profile. Further studies are warranted to dissect a plausible mechanism of action.

AAV-mediated genome editing faces several significant challenges12,13. Of particular relevance, our understanding of the correlation between delivery, CRISPR/Cas9 and guide RNA (gRNA) expression and editing efficiency continues to evolve. In the current study, we developed an optimal configuration of dual AAV vector systems required, while concurrently evaluating the editing efficiencies afforded by AAV.cc47 and AAV9 vectors. In A9 mice, we observed that higher gRNA to Cas9 expressing vector ratios or the use of self-complementary AAV cassettes to express gRNAs in conjunction with an improved capsid (AAV.cc47) markedly improved gene editing efficiency. These results corroborate the notion that gRNA levels might be rate limiting in in vivo genome editing strategies14,15. Analysis of RNA transcript levels of both SaCas9 and gRNAs in these tissues revealed that AAV.cc47 vectors indeed transduced tissues more efficiently than AAV9 vectors further lending support to an improved post-entry mechanism. Of note, while the amount of SaCas9 mRNA transcripts observed at 4 weeks was low post-injection, gRNA levels were found to be relatively high. This disparity in RNA levels between SaCas9 and gRNA could arise from promoter strength differences (Pol-II vs Pol-III), CMV promoter silencing in muscle, an immune response to SaCas9 in adult mice, or AAV genome composition (single-stranded SaCas9 vs self-complementary gRNA cassettes). Further, previous studies have shown that the CMV promoter is prone to silencing in muscle16 and this has been observed in AAV-CRISPR muscle genome editing as well17. It is plausible that an immune response to Cas9 in adult mice and promoter silencing could have contributed to this effect. The U6 promoter-driven gRNAs do not appear to be subject to this at least at earlier time points reported here. It should also be noted that overall editing efficiencies in adult cardiac and skeletal muscle tissues (but not the liver) following IV dosing of dual AAV vector systems show room for improvement in general. Nevertheless, our studies unequivocally demonstrate that AAV.cc47 significantly improves genome editing efficiency with CRISPR/Cas9 systems; however, we acknowledge that additional improvements to cassette design might further enhance desired outcomes.

In summary, our study corroborates that cross-species evolution can yield highly potent AAV variants with the ability to transduce multiple tissues across different preclinical animal models. Notably, we test different mouse strains, cynomolgus macaques and/or pigs, all of which constitute critical path in the clinical development of different human gene therapies. The lead capsid, AAV.cc47 is currently being evaluated for a broad range of therapeutic gene transfer applications in preclinical models of musculoskeletal, cardiac, CNS and kidney disease. While mechanistic insights continue to be explored, we postulate that subjecting AAV capsid libraries to evolutionary pressure in different species is likely to comprehensively address both host-specific and host-agnostic factors (essential or restrictive) that impact AAV transduction. This cross-species approach can be readily adapted to any AAV library evolution/screening and, when combined with promoter-specific selection strategies, potentially improve the clinical translatability of next-generation AAV vectors.

**Methods**

**Study design**

The overall goal of this study was to determine whether cycling of AAV capsid libraries through multiple animal species would yield a cross-species compatible AAV variant with improved properties. The general approach was to select preclinically relevant animal models to cycle our AAV libraries through and directly compare gene transfer and genome editing efficiencies benchmarked against the parental AAV9 serotype. For all dosing strategies, injections were blinded for each test vector evaluated. In the case of the Pompe study, tissue
analysis from different cohorts was blinded. All mouse and pig protocols were approved by the Institutional Animal Care and Use Committee (IACUC) at Duke University (mouse: Protocol A189-21-09) and North Carolina State College of Veterinary Medicine (pigs: Protocol 20-425). NHP studies and protocols were reviewed and approved by IACUC at Southern Research (Birmingham, AL; Protocol IS863.01) and CR MWN IACUC at Charles River Labs (Kalamazoo, MI; Protocol 2728-017), both accredited by the Association for Assessment and Accreditation of Laboratory Animal Care-International (AAALAC). All studies were carried out using applicable Standard Operating Procedures at Southern Research (Birmingham, AL) and Charles River Labs (Kalamazoo, MI) and previously reported by our lab[35]. During the study, the care and use of animals was conducted in accordance with the guidelines of the USA National Research Council, the US Department of Agriculture (Animal Welfare Act: Public Law 99–198) and those of the Guide for the Care and Use of Laboratory Animals (National Academies Press, 2011). In general, sample size for mouse studies ranges from 3–6 replicates for each condition per experiment, with specific information provided in the figure legends and methods and materials section. In the case of our large animal studies, 1–2 replicates were used for each species.

AAV plasmid libraries

The AAV9 VR-IV plasmid library was constructed as previously described[27-28]. Briefly, the parental library plasmid pTR-wtAAV9-VR-IV library consisting of AAV2 ITRs flanking AAV2 Rep and AAV9 Cap genes was generated by randomizing a portion of Cap corresponding to VR-IV (amino acids 452–458 in Cap) via saturation mutagenesis. The subsequent evolved library plasmids differed from the parental library plasmid only in the generation of the library inserts. Prior to the incorporation of the randomized insert containing degenerate nucleotides at the selected region into the parental library, a region of Cap encompassing VR-IV was replaced with a partial GFP sequence to prevent any incorporation of WT VR-IV sequences in the parental library plasmid. Sequencing library sequences were ligated into our AAV library plasmid backbone prior to production of the next round AAV capsid library. Newly evolved AAV strains were further processed to remove existing endotoxins using a Pierce high-capacity endotoxin removal spin column (Thermo Fisher Scientific). Endotoxin levels were then quantified via a Chromogenic LAL endotoxin assay kit (GenScript).

Cross-species cycling of AAV capsid libraries

AAV capsid libraries were produced as described in the section above. The first round of AAV evolution occurred in the pig brain following intravenous administration. The brain was harvested and dissected further into multiple brain regions prior to genomic DNA isolation. AAV genomes were amplified from pig genomic DNA using primers targeting AAV9 Cap (Supplementary Table 1). PCR amplicons containing library sequences were ligated into our AAV library plasmid backbone and used to generate the round 2 AAV9 library. The second round of AAV evolution occurred in mice following intravenous administration. Mouse brains were harvested 3 days post injection and dissected further into multiple brain regions. AAV9 Cap containing the library region was amplified through PCR and cloned into our AAV library plasmid backbone prior to production of the next round AAV libraries. The final round of AAV evolution in non-human primates occurred following intravenous administration of the last AAV library preparation. Brains from these primates were harvested 7 days post injection and were processed as described for pigs and mice. In vivo cycling between these three species was completed before next-generation sequencing of capsid variants.

High-throughput sequencing analysis and identification of newly evolved AAV strains

Sequence diversity of the parental and evolved AAV9 capsid libraries were assessed by using the viral library preparations. The evolved library sequences were generated following PCR of DNA amplified tissue followed by plasmid cloning and viral production. Both the parental and evolved viral libraries after production were DNase-I treated prior to extraction of viral genomes from the capsids and subsequent addition of Illumina adapter sequences via PCR. The first round of PCRs adds half of the Illumina sequencing adapters using primers specific to the ampiclon and the second round of PCR adds the remaining adapter sequences and index sequence. After each round of PCR, the products were purified using the PureLink PCR Micro Kit (Invitrogen). The quality of the amplicons was verified using a Bioanalyzer (Agilent), and concentrations were quantified using a Qubit spectrometer (Thermo Fisher Scientific). Libraries were then prepared for sequencing with the Illumina NovaSeq 6000 S-Prime Reagent Kit (300 cycles), following manufacturer’s instructions, and sequenced on the Illumina NovaSeq system. Demultiplexed reads were analyzed with an updated in-house Perl script[36]. Reads are surveyed for nucleotide sequences flanking the library regions and intermediate sequences are
counted and ranked. These nucleotide sequences are then translated, and the resulting amino acid sequences are also counted and ranked. Then, the percent representation and fold enrichment between parental and evolved libraries were calculated and ranked. Amino acid sequences were plotted according to their percent representation and enrichment for each library. A second Perl script was used to calculate the frequency of each amino acid at each position in the library, considering each capsid variant. Perl scripts were executed using Strawberry Perl for Windows software v 1.32.1. Bubble plots representing capsid mutants and amino acid distribution heatmaps were generated using the R studio graphics package v3.5.2.

AAV packaging plasmid constructs

The plasmids saAAV-CBh-mCherry and scAAV-CBh-eGFP were generated or obtained as described previously. The plasmid ssAAV-CMV-Cre-B6 used for our dose escalation studies in A9 mice was generously provided by Dr. Charles Gersbach (Duke University). For AAV-CRISPR studies, constructs containing the CMV enhancer/beta-actin (CB) promoter driving Staphylococcus aureus (Sa) Cas9 and gRNAs targeting the mouse Rosa26 locus were provided by Dr. William Lagor (Baylor College of Medicine; Supplementary Table 1) through the Somatic Cell Genome Editing Consortium (SCGE). The same gRNA 1 and 2 were both used in all A9 CRISPR studies. For AAV-CRISPR studies where both gRNAs are on a single vector, both gRNAs and U6 promoters used in the CB strategy were cloned into a single-stranded AAV plasmid containing 2 gRNA scaffolds via Gibson assembly (NEB). For A9 CRISPR studies using a self-complementary gRNA vector, the dual gRNA insert from the ssAAV GRNA plasmid was cloned into the plasmid saAAV-CBh-mCherry replacing the CB-mCherry insert with the gRNA insert via Gibson assembly. For mxd CRISPR studies, gRNA sequences targeting the intronic regions around exon 23 in Dmd were used as previously reported. See Supplementary Table 1 for the Dmd targeting gRNA 1 and 2 sequences. The ssAAV-CMV-SaCas9 plasmid used in A9 and mxd studies was generously provided by Dr. Charles Gersbach (Duke University). The ssAAV-CBh-GAA plasmid used in Gaa (-/-) mice was generously provided by Dr. Mai K Elmallah (Duke University).

Mouse studies

All mouse strains used in this study were maintained at Duke University School of Medicine with the assistance of Duke's Division of Laboratory Animal Resources (DLAR). Mice were housed in a temperature-controlled (~18–23°C, 40–60% humidity) and enriched environment, with a 12 h light/dark cycle, and provided standard chow and water.

Intravenous dosing in wild-type reporter mice

All mouse IV dosing occurred in in-house bred P0–P1 male and female C57/B6 neonates (~n = 6). Pups were rapidly anesthetized on ice for 2 min followed by IV injection using a stereotaxic apparatus. AAV vectors (2 μL total volume) packaging the CBh-mCherry transgene cassette were injected into the left lateral ventricle using a Micro-injection Syringe Pump (World Precision Instruments) at a controlled rate of 0.066 μL per second with a Hamilton 700 series syringe and 26 g needle attached to a KOPF-900 small animal stereotaxic instrument (KOPF instruments). The following stereotaxic coordinates were used for injections: 0.8 mm relative to the sagittal sinus, 1.5 mm rostral to transverse sinus, and 1.5 mm deep. Dosing in these mice consisted of a total of 2e10vg administered in 2 μL. Mice were revived using a heating pad and rubbed in the bedding after AAV injections before being placed back with the dam. Mouse brains were harvested 4 weeks post infusion.

Pig evolution and dosing studies

Pigs used in this study were a cross between Landrace, Yorkshire, and Duroc strains. Pig studies consisting of AAV evolution or reporter expression were all performed in 3-week-old newly weaned piglets weighing approximately 7 kgs. For AAV evolution in pigs (n = 2), male and female piglets were injected systemically via the heart vena cava with 1e13 vg/kg of AAV libraries. Brain and spinal cord were harvested 3–4 days post injection. For intrathecal reporter studies in pig, 3-week-old male pigs (n = 1) weighing approximately 7–10 kgs were anesthetized via a ketamine/xylazine cocktail containing 20 mg/kg ketamine +2 mg/kg xylazine. AAV vectors (2 mL total volume) packaging a CBh-eGFP transgene cassette were injected into the lumbar cistern with a total of 3e13vg and were infused at a rate of 1 mL per minute. Pigs brains were harvested 13 days post infusion.

Non-human primate evolution and dosing studies

Female Cynomolgus macaques aged 2–3 years old were used for our AAV evolution studies. These primates were injected intravenously with 3.3e13vg/kg of AAV libraries. Primates were sacrificed 7 days post injection and target organs were harvested prior to library amplification. For reporter studies in non-human primates (n = 2), female 2–3 year-old Cynomolgus macaques were administered a total of 1e13vg in a total volume of 2 mLs via intracranial magna infusion. Target organs from injected primates were harvested 13 days post infusion.

Histological processing and staining

All C57/B6 and A9 tissues isolated were post fixed in 10% formalin (VWR) overnight. Livers, hearts, and brains were washed 3x in 1xPBS,
embedded in 3% agarose and sectioned 50 μm thick via a vibratome (Leica Biosystems). Tibialis anterior skeletal muscles were incubated in 30% sucrose overnight following fixation, embedded in O.C.T compound (Electron Microscopy Sciences), and frozen in liquid nitrogen-cooled isopentane. O.C.T. blocks were sectioned 7 μm thick on a Leica cryostat. For mCherry evaluation, tissue was mounted on slides with ProLong Gold Antifade Mountant with DAPI (Thermo Fisher Scientific) and imaged for native fluorescence. To quantify the percent of transduced neurons in the brain, tissue was incubated in blocking solution (5% normal goat serum, 0.1% Triton X-100 in 1x PBS) for 1 h at room temperature followed by immunofluorescence staining of brain tissue with anti-chicken mCherry (1:750; Abcam) and anti-rabbit NeuN (1:500; Abcam) primary antibodies overnight at 4 °C. Tissue was then washed 3x in 1xPBS and incubated with anti-chicken Alexa Fluor 488 (1:500; Invitrogen) and anti-rabbit Alexa Fluor 647 (1:500; Invitrogen) secondary antibodies for 1 h. For immunofluorescence staining in A9-Cre and A9-CRISPR studies, tibialis anterior skeletal muscles were incubated in citrate buffer (Abcam) at 80°C for 25 min and allowed to cool to room temperature in cool water. Heart, tibialis anterior, liver, and brain sections were then incubated in blocking solution, containing 5% normal goat serum and 0.1% Triton X-100 in 1x PBS, for 1 h at room temperature. Brains were further incubated in Mouse on Mouse (M.O.M.) blocking reagent (Vector labs) for 24 h. Heart, tibialis anterior, liver, and brain sections were then incubated with an anti-rabbit RFP primary antibody (1:500; Rockland Immunocchemicals), and in addition, brains were also incubated with anti-mouse NeuN primary antibody (1:500; Abcam), diluted in blocking buffer overnight. Tissue sections were washed 3x in 1xPBS the following day and incubated with anti-rabbit Alexa Fluor 647 (1:500; Invitrogen) and anti-mouse Alexa Fluor 488 (brain only: 1:500; Invitrogen) secondary antibodies. For A9-Cre studies, tissue was then washed 3x in 1xPBS followed by secondary antibody incubation with an anti-rabbit Alexa Fluor 488 (1:1000; Invitrogen) and anti-mouse Alexa Fluor 647 (brain: 1:500; Abcam) secondary antibodies in addition to DAPI (Thermo Fisher Scientific). For A9-CRISPR studies, tissue was washed 3x in 1xPBS followed by secondary antibody incubation with an anti-rabbit Alexa Fluor 647 (1:1000; Abcam) and DAPI (Thermo Fisher Scientific). For immunofluorescence staining of ICV injected mouse brains, brain sections were prepared for staining as described above except sections were incubated with an anti-rabbit mCherry primary antibody (1:750; Abcam) and anti-rabbit Alexa Fluor 647 secondary antibody (1:500; Abcam). Immunostained tissue sections were then mounted on slides with ProLong Gold antifade reagent without DAPI (Invitrogen).

All NHP and pig tissues isolated were post fixed in 10% formalin (VWR) overnight. Mdx hearts and tibialis anterior skeletal muscles were then incubated in 30% sucrose overnight following fixation, embedded in O.C.T compound (Electron Microscopy Sciences), and frozen in liquid nitrogen-cooled isopentane. O.C.T. blocks were sectioned 10μm (heart) and 7μm (tibialis anterior) thick on a Leica cryostat. For immunofluorescence staining in mdx CRISPR studies, tibialis anterior skeletal muscles were incubated in citrate buffer (Abcam) at 80°C for 25 min and allowed to cool to room temperature in cool water. Heart and tibialis anterior sections were then incubated in a blocking solution, containing 5% normal goat serum and 0.1% Triton X-100 in 1xPBS, for 1 h at room temperature. Next, tissues were incubated in Mouse on Mouse (M.O.M.) blocking reagent (Vector labs) overnight. Tissue was then incubated with anti-mouse dystrophin (1:500; MANDYS8; Sigma-Aldrich) and anti-rabbit laminin (1:300; Abcam) primary antibodies diluted in a blocking buffer overnight. Next day, tissue was washed 3x with 1xPBS followed by secondary antibody incubation with an anti-rabbit Alexa Fluor 488 (1:500; Invitrogen) and anti-mouse Alexa Fluor 647 (1:500; Abcam), followed by DAPI staining (Thermo Fisher Scientific). Immunostained tissue sections were then mounted on slides with ProLong Gold antifade reagent without DAPI (Invitrogen).
produced cDNA was used with probes against the Dmd exon 4-5 junction to quantify total Dmd transcripts and another probe against Dmd exon 22–24 junction to quantify the amount of Dmd exon 23-deleted transcripts. A commercial Taqman assay for 18S ribosomal RNA was used as a housekeeping control. Each plate was run in duplicate with all three probes on a single plate for each tissue. Taqman Fast Advanced Master Mix (LifeTech) was used following manufacturer’s instructions for cycling conditions. Delta-Ct values between exon 4–5 and exon 22–24 were used to quantify the percentage of Dmd exon 23 deleted transcripts compared to total Dmd transcripts.

Acid alpha-glucosidase (GAA) activity assay
GAA enzymatic activity was quantified as previously described. Briefly, snap-frozen tissues were homogenized using a FastPrep bead system (MP Biomedicals), followed by three cycles of freezing and thawing. The samples were then clarified by centrifugation. Samples were evaluated in duplicate in a black 96-well plate. 20 µL of the sample was incubated with 4 µmol/l benzyl-α-D-glucopyranoside substrate (Sigma-Aldrich) at 37 °C for 1 h in sodium acetate (Sigma-Aldrich) buffer at pH 4.3. The reaction is stopped with sodium carbonate (Sigma-Aldrich) buffer pH 10.7. Fluorescence was read with a Varioskan Lx3 (Thermo Fisher Scientific) at 360 nm excitation and 460 nm emission. The values were compared to those from a standard curve made with 4 mM methylumbelliferyl (Sigma-Aldrich). Samples were normalized to total protein content measured by DC Protein Assay (Bio-Rad), performed as described by the manufacturer.

Determination of mCherry protein levels by ELISA
mCherry protein levels were quantified with a commercially available colorimetric ELISA kit (Abcam). Briefly, a tagged capture antibody and HRP-labeled detection antibody were mixed in solution with standards, controls, or samples generating a capture antibody/mCherry/detection antibody complex which is then immobilized to the wells of a 96-well plate coated with an anti-tag antibody. 3,3′,5,5′-Tetramethylbenzidine (TMB) substrate was added, the reaction was stopped, and absorbance was measured at 450 nm.

Luciferase expression assay
HuH7 scramble and AAVR knockout cell lines were generated in-house at Duke University. Cells were seeded in 24-well plates at 50,000 cells per well in DMEM supplemented with 10% FBS and pen strep. The cells were then transduced with AAV9 or AAV.cc47 vectors packaging a single stranded CBA-luciferase reporter at a MOI of 10,000, 50,000, and 100,000 vector genomes per cell and incubated at 37 °C for 48 h. Cells were then lysed in 100 µL Passive Lysis Buffer (Promega), and transduction was assessed by measuring luciferase transgene expression with a Victor X3 microplate reader (PerkinElmer) and luciferase assay reagent (Promega) per manufacturer specifications.

AAV Neutralizing antibody assay
Antiserum (from different species, diluted 1:5) was mixed with an equal volume of recombinant AAV9-Luc or AAV.cc47-Luc vectors and then incubated at room temperature for 30 min in tissue culture–treated, black, glass-bottom, 96-well plates (Corning). Then, HEK293s were seeded at 10,000 cells/well on top of the incubated vectors. Vectors were used at a MOI of 100,000 vg/cell and transduced cells were incubated in 5% CO2 at 37 °C for 48 h. Cells were then lysed with 25 µL of 1× passive lysis buffer (Promega) for 30 min at room temperature. Luciferase activity was measured on a VICTOR3 multilabel plate reader (PerkinElmer) immediately after the addition of 25 µL of luciferin (Promega). All readouts were normalized to controls with no antibody/antisera treatment. Recombinant AAV vectors packaging CBA-Luc transgenes, antibodies, sera, were pre-diluted in DMEM and used in this assay.

Statistical analysis
Data are represented as mean ± standard error mean. For data sets with two groups, significance was determined using an unpaired student’s t-test unless otherwise noted. For data sets with three groups, significance was determined by one-way ANOVA, with Tukey’s post-test. Where indicated, ns represents not significant; *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001; ns not significant.

Reporting summary
Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The NGS datasets for capsid libraries reported in this article are available under Sequence Read Archive accession code PRJNA869670. All other data associated with this study are present in the paper or the Supplementary Materials. Source data for each relevant figure is provided in a Source Data file. The data that support the findings of this study are available from the corresponding author upon reasonable request. Correspondence and requests for materials should be addressed to AA at aravind.asokan@duke.edu. Source data with this paper.

Code availability
The Perl scripts used for NGS analysis of AAV variants have been described previously. The scripts have been deposited into Zenodo repository and can be accessed via https://doi.org/10.5281/zenodo.7075694. Correspondence should be addressed to AA at aravind.asokan@duke.edu.

References
1. Kuzmin, D. A. et al. The clinical landscape for AAV gene therapies. Nat. Rev. Drug Discov. 20, 173–174 (2021).
2. Mendell, J. R. et al. Current clinical applications of in vivo gene therapy with AAVs. Mol. Ther. 29, 464–488 (2021).
3. Mendell, J. R. et al. Single-dose gene-replacement therapy for spinal muscular atrophy. N. Engl. J. Med. 377, 1713–1722 (2017).
4. Mullard, A. Gene therapy community grapples with toxicity issues, as pipeline matures. Nat. Rev. Drug Discov. 20, 804–805 (2021).
5. Wilson, J. M. & Flotte, T. R. Moving forward after two deaths in a gene therapy trial of myotubular myopathy. Hum. Gene Ther. 31, 695–696 (2020).
6. High-dose AAV gene therapy deaths. Nat Biotechnol 38, 910 (2020).
7. Hinderer, C. et al. Severe toxicity in nonhuman primates and piglets following high-dose intravenous administration of an adeno-associated virus vector expressing human SMN. Hum. Gene Ther. 29, 285–298 (2018).
8. Hordeaux, J. et al. Adeno-associated virus-induced dorsal root ganglion pathology. Hum. Gene Ther. 31, 808–818 (2020).
9. Potter, R. A. et al. Dose-escalation study of systemically delivered rAAVrh74.MHCK7.micro-dystrophin in the mdx Mouse Model of Duchenne Muscular Dystrophy. Hum. Gene Ther. 32, 375–389 (2021).
10. Le Guiner, C. et al. Long-term microdystrophin gene therapy is effective in a canine model of Duchenne muscular dystrophy. Nat. Commun. 8, 16105 (2017).
11. Mendell, J. R. et al. Assessment of systemic delivery of rAAVrh74.MHCK7.micro-dystrophin in children with duchenne muscular dystrophy: a nonrandomized controlled trial. JAMA Neurol. 77, 1122–1131 (2020).
12. Li, C. & Samulski, R. J. Engineering adeno-associated virus vectors for gene therapy. Nat. Rev. Genet 21, 255–272 (2020).
13. Zolotukhina, S. & Vandenberghhe, L. H. AAV capad design: a Goldilocks challenge. Trends Mol. Med 28, 183–193 (2022).
14. Asokan, A. et al. Reengineering a receptor footprint of adeno-associated virus enables selective and systemic gene transfer to muscle. Nat. Biotechnol. 28, 79–82 (2010).
15. Shen, S. et al. Engraftment of a galactose receptor footprint onto adeno-associated viral capsids improves transduction efficiency. J. Biol. Chem. 288, 28814–28823 (2013).
16. Choudhury, S. R. et al. In vivo selection yields AAV-B1 capsid for central nervous system and muscle gene therapy. Mol. Ther. 24, 1247–1257 (2016).
17. Albright, B. H. et al. Mapping the structural determinants required for AAVrh.10 transport across the blood-brain barrier. Mol. Ther. 26, 510–523 (2018).
18. Deverman, B. E. et al. Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. Nat. Biotechnol. 34, 204–209 (2016).
19. Goertsen, D. et al. AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset. Nat. Neurosci. 25, 106–115 (2022).
20. Weinmann, J. et al. Identification of a myotrophic AAV by massively parallel in vivo evaluation of barcoded capsid variants. Nat. Commun. 11, 5432 (2020).
21. Tabebordbar, M. et al. Directed evolution of a family of AAV capsid variants enabling potent muscle-directed gene delivery across species. Cell 184, 4191–4193.e4922 (2021).
22. Hordeaux, J. et al. The Neurotropic Properties of AAV-PHP-B Are Limited to C57BL/6J Mice. Mol. Ther. 26, 664–668 (2018).
23. Huang, Q. et al. Delivering genes across the blood-brain barrier: LY6A, a novel cellular receptor for AAV-PHP-B capsids. PLoS One 14, e0225206 (2019).
24. DiMattia, M. A. et al. Structural insight into the unique properties of adeno-associated virus serotype 9. J. Virol. 86, 6947–6958 (2012).
25. Adachi, K., Enoki, T., Kawano, Y., Veraz, M. & Nakai, H. Drawing a high-resolution functional map of adeno-associated virus capsid by massively parallel sequencing. Nat. Commun. 5, 3075 (2014).
26. Emmanuel, S. N. et al. Structurally mapping antigenic epitopes of adeno-associated virus 9: development of antibody escape variants. J. Virol. 96, e025121 (2022).
27. Tse, L. V. et al. Structure-guided evolution of antigenically distinct adeno-associated virus variants for immune evasion. Proc. Natl Acad. Sci. USA 114, E4812–E4821 (2017).
28. Havlík, L. P. et al. Coevolution of adeno-associated virus capsid antigenicity and tropism through a structure-guided approach. J Virol 94, e00976-20 (2020).
29. Nelson, C. E. et al. Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. Nat. Med 25, 427–432 (2019).
30. Wang, D., Tai, P. W. L. & Gao, G. Adeno-associated virus vector as a platform for gene therapy delivery. Nat. Rev. Drug Discov. 18, 358–378 (2019).
31. Gough, V. & Gersbach, C. A. Immunity to Cas9 as an Obstacle to Persistent Genome Editing. Mol. Ther. 28, 1389–1391 (2020).
32. Wang, D., Zhang, F. & Gao, G. CRISPR-based therapeutic genome editing: strategies and in vivo delivery by AAV vectors. Cell 181, 136–150 (2020).
33. Hakim, C. H. et al. AAV CRISPR editing rescues cardiac and muscle function for 18 months in dystrophic mice. JCI Insight 3, e124297 (2018).
34. Min, Y. L. et al. CRISPR-Cas9 corrects Duchenne muscular dystrophy exon 44 deletion mutations in mice and human cells. Sci. Adv. 5, eaav4324 (2019).
35. Brooks, A. R. et al. Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. J. Gene Med 6, 395–404 (2004).
36. Elmore, Z. C. et al. Rescuing AAV gene transfer from neutralizing antibodies with an IgG-degrading enzyme. JCI Insight 5, 19 e139881 (2020).
37. Elmore, Z. C. et al. The membrane associated accessory protein is an adeno-associated viral egress factor. Nat. Commun. 12, 6239 (2021).
38. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13, 133–140 (2010).
39. Sambasivan, R. et al. Distinct regulatory cascades govern extracellular and pharyngeal arch muscle progenitor cell fates. Dev. Cell 16, 810–821 (2009).
40. Raben, N. et al. Targeted disruption of the acid alpha-glucosidase gene in mice causes an illness with critical features of both infantile and adult human glycogen storage disease type II. J. Biol. Chem. 273, 19086–19092 (1998).
41. Tabebordbar, M. et al. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science 351, 407–411 (2016).
42. McCall, A. L. et al. Reduction of autophagic accumulation in Pompe disease mouse model following gene therapy. Curr. Gene Ther. 19, 197–207 (2019).
43. McCall, A. L. et al. Glycogen accumulation in smooth muscle of a Pompe disease mouse model. J. Smooth Muscle Res 57, 8–18 (2021).
44. Havlík, L. P. et al. Receptor switching in newly evolved adeno-associated Viruses. J. Virol. 95, e0058721 (2021).

Acknowledgements
This study was funded by NIH grants awarded to AA (UG3AR07336, UH3AR075336) in addition to support from Translating Duke Health. We would like to acknowledge A. Rosales, D. Fiftis, S. Stuppy and M. L. Huston for their assistance with molecular biology, L. Edwards and K. Gleason for their assistance in completing in-life portions for our pig studies, N. Olby for her assistance in performing the intrathecal infusions in pigs, J. McNamara and E. Matthews for their assistance with dissection of pig brains, H. Daniels for her assistance with breeding the Pax7-nGFP-mdx mice used in these studies, and Duke DLAR for their assistance with mouse care. We acknowledge the Duke light microscopy core facility, the Duke Surgery Substrate Services Core research support staff, and Translating Duke Health initiative for their support as well as StrideBio Inc., and Charles River laboratories for funding and technical assistance with NHP studies. Figures were created in part using BioRender.

Author contributions
Conceptualization: T.J.G. and A.A. Methodology: T.J.G., T.J.S., D.K.O., L.P.H., L.B., A.L.R., G.W.D., M.K.E., and C.A.G. In vivo studies: T.J.G., L.B., K.S., M.M.F., A.L.R., M.S.M., and R.M.C.R. Investigation: T.J.G., L.B., K.S., M.M.F., A.L.R., M.S.M., and R.M.C.R. Visualization: T.J.G., L.P.H., and A.A. have led patent applications on the subject matter of this manuscript. A.A. is a co-founder at StrideBio and TorqueBio and serves on the advisory boards of Atsena Therapeutics, Isolere Bio, Mammoth Bio, Ring Therapeutics, and Kriya Therapeutics. C.A.G. is an advisor to Sarepta Therapeutics, Tune Therapeutics, Levo Therapeutics, and Iveric Bio and a co-founder of Tune Therapeutics, Element Genomics, and Locus Biosciences. R.M.C.R. and M.S.M. were employed at StrideBio Inc., at this time. The remaining authors declare no competing interests.

Competing interests
T.J.G., L.P.H., and A.A. have filed patent applications on the subject matter of this manuscript. A.A. is a co-founder at StrideBio and TorqueBio and serves on the advisory boards of Atsena Therapeutics, Isolere Bio, Mammoth Bio, Ring Therapeutics, and Kriya Therapeutics. C.A.G. is an advisor to Sarepta Therapeutics, Tune Therapeutics, Levo Therapeutics, and Iveric Bio and a co-founder of Tune Therapeutics, Element Genomics, and Locus Biosciences. R.M.C.R. and M.S.M. were employed at StrideBio Inc., at this time. The remaining authors declare no competing interests.
Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-33745-4.

Correspondence and requests for materials should be addressed to Aravind Asokan.

Peer review information Nature Communications thanks Alexandra Stanton and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022