A single amino acid variant in the variable region I of AAV capsid confers liver detargeting

Ruxiao Xing*[1], Mengyao Xu*[1], Darcy Reil[1], April Destefano[1], Mengtian Cui[1], Nan Liu[1], Jialing Liang[1], Guangchao Xu[1], Li Luo[1], Meiyu Xu[1], Fang Zhang[1], Phillip W.L. Tai[1], Yuquan Wei[2], Alisha M. Gruntman[1], Terence R. Flotte[1], Guangping Gao[#1,3], Dan Wang[#1,4]

[1]Department of Genetic and Cellular Medicine, Horae Gene Therapy Center, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA

[2]State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, China

[3]Department of Microbiology, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA

[4]RNA Therapeutics Institute, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA

*Co-first authors

[#]Co-corresponding authors

Correspondence should be addressed to:

Dan Wang, Ph.D.
Department of Genetic and Cellular Medicine
UMass Chan Medical School
Worcester, MA 01605, USA
Telephone: +1-(774) 455-4574
Fax: +1-(508) 856-1552
E-mail: Dan.Wang@umassmed.edu

Guangping Gao, Ph.D.
Department of Genetic and Cellular Medicine
UMass Chan Medical School
Worcester, MA 01605, USA
Telephone: +1-(508) 856-3563
Fax: +1-(508) 856-1552
E-mail: Guangping.Gao@umassmed.edu

ABSTRACT

AAV capsid serotypes isolated from nature have been widely used in gene delivery and gene therapy. Recently, more than 1,000 distinct AAV capsids were identified from human clinical samples by high-throughput, long-read DNA sequencing. In this study, we tap into this broad natural biodiversity of AAV capsids to develop liver-tropic AAV capsids. We initially screened a subset of variants derived from AAV8 (n=159) for packaging efficiency. The high-yielding variants were subjected to a barcoded vector library screen in mice and ferrets for their ability to mediate liver gene transfer. Although no variant surpassed AAV8 for liver targeting, several exhibited a liver detargeting phenotype. Among these, we focused on the N271D variant (AAV8 VP1 numbering), located in the variable region I (VR-1), which has been previously implicated in influencing liver tropism. The liver detargeting phenotype of AAV8.N271D was confirmed by single vector administration in mice. Additionally, we grafted the N271D variant onto AAV9 and MyoAAV capsids (N270D by AAV9 VP1 numbering). The AAV9.N270D and MyoAAV.N270D vectors showed a similar liver-detargeting phenotype, although muscle targeting was moderately reduced. Although we did not identify any capsid variants that outperform AAV8 in liver transduction, this study reinforces the important role of VR-1 in modulating liver tropism and highlights the potential of engineering VR-1 residues to reduce liver gene transfer and associated toxicity observed in several gene therapy studies.

21 **INTRODUCTION**

22 Adeno-associated virus (AAV) has a single-stranded DNA genome packaged in an icosahedral
23 capsid that consists of 60 protein monomers, VP1, VP2, and VP3[1]. The capsid initiates interactions with
24 the host, including binding to cell surface receptors and recognition by the host immune system. In
25 recombinant AAV (rAAV), which is widely used as an *in vivo* gene therapy delivery platform[2], the capsid
26 is either identical to or derived from wild-type AAV (wtAAV), and therefore is inherently critical to
27 determine therapeutically relevant properties, including tissue tropism and immunogenicity. To develop
28 AAV capsids with clinically favorable properties, one way is to isolate wtAAV capsid sequences that
29 exist in nature followed by testing in the context of rAAV. Naturally occurring wtAAV capsids evolve
30 within host tissues during infections[3] and may inherently possess desirable features for therapeutic
31 gene delivery, such as tissue-specific fitness and the ability to evade host immune surveillance. This
32 approach has led to the discovery of a series of AAV capsid serotypes[4], many of which have been used
33 in clinical gene therapy and continue to serve as the workhorse for approved AAV-based gene
34 therapies to date[5]. Additionally, these naturally occurring AAV capsid variants serve as a foundation for
35 capsid engineering through various approaches, most notably peptide insertion combined with library
36 selection and directed evolution, to further enhance desirable rAAV attributes[6].

37 From a biology perspective, these naturally occurring AAV capsid sequences offer ample
38 opportunities to study the sequence- and structure-function relationship. For example, aligning the
39 amino acid sequences of various AAV capsids results in the identification of nine variable regions
40 (VRs), which diverge to a higher degree than the remaining regions among different serotypes.
41 Structural studies revealed that VR-I to VR-IX are mostly exposed to the outer surface of the
42 assembled capsid, and play critical roles in binding to a diverse array of cell surface receptors and host
43 antibodies[7]. For example, AAV2.5 was created by incorporating five residues from AAV1 into AAV2, and
44 showed enhanced muscle tropism[8]. A follow-up study demonstrated that, among the five residues, the
45 265T insertion in VR-I determined the high muscle tropism of AAV2.5[9]. AAV-DJ is an engineered capsid

46 created by shuffling eight capsid serotypes and exhibits resistance to neutralization by human
47 intravenous immunoglobulin (IVIG)[10]. Structural studies showed that the unique structure of VR-1 in
48 AAV-DJ as compared with AAV2 abrogated binding to the A20 neutralizing antibody[11]. Indeed, VR-1
49 contributes to the 3-fold protrusions on AAV capsid, which play critical roles in binding to cell surface
50 receptors and antibodies[12]. Understanding the capsid-host interactions at the molecular level also
51 enables rational engineering by manipulating specific residues to achieve desired capsid properties[13,14].
52 Recent high-throughput, long-read DNA sequencing technologies have allowed more efficient and in-
53 depth profiling of naturally existing AAV capsid variants. For example, we previously identified more
54 than 1,000 unique AAV capsid variants from human clinical biopsies using the PacBio sequencing
55 platform[15]. An AAV2-derived variant, named AAVv66, exhibits enhanced production yields, virion
56 stability, and central nervous system (CNS) transduction[15].

57 Two liver-targeted gene therapies delivered by AAV vectors have been approved for treating
58 hemophilia A and hemophilia B, respectively. In both cases, therapeutic levels of the blood clotting
59 factor VIII[16] or factor IX[17] secreted from the liver into the bloodstream could be attained. However,
60 transducing the liver to function as a bio-factory for producing and secreting alpha-1 antitrypsin (A1AT),
61 the second most abundant secreted serum protein[18], proved challenging in reaching the therapeutic
62 threshold in the bloodstream for treating A1AT deficiency[19], suggesting that more efficient liver targeting
63 may be required. Among the AAV capsid variants that we previously identified from human clinical
64 samples[15], a large portion are AAV8-derived (i.e., exhibiting the highest degree of sequence homology
65 to AAV8)[20]. Given AAV8's strong liver tropism, in this study, we tapped into this diverse repertoire of
66 AAV8 variants to characterize their liver tropism in mice using a barcoded library screen approach,
67 aiming to identify more potent liver-tropic capsid variants as potential vectors to deliver A1AT deficiency
68 gene therapy. In addition, ferret models of A1AT deficiency were recently generated and characterized
69 to be a platform for preclinical testing of therapeutics including gene therapy[21]. Therefore, we also
70 screened the same capsid library in ferrets to investigate cross-species tropism of emerging variants.

71 Although we found that no variants surpassed AAV8 for liver gene transfer, we identified one variant
72 with a liver detargeting phenotype mediated by a single N271D residue change in VR-1 (AAV8 VP1
73 numbering). This study adds to the recent literature showing the important role of VR-1 in liver gene
74 transfer[22-24]. Furthermore, several studies on systemic AAV vector administration have demonstrated
75 liver toxicity[25-28], suggesting the potential safety benefit with a liver-detargeting capsid to deliver gene
76 therapy to other organs, such as the CNS and muscle. The key VR-1 residues that influence liver gene
77 transfer identified by us and others may provide a rational avenue to develop AAV vectors with reduced
78 liver toxicity.

79

80 **RESULTS**

81 **Generation and characterization of an AAV8 vector library**

82 We first constructed a total of 159 packaging plasmids, each expressing a unique AAV8 capsid
83 variant identified from human tissues **(Supplementary Table 1)** along with the AAV2 Rep. These
84 plasmids were used individually to package an *EGFP* transgene cassette in a small-scale AAV vector
85 production assay to determine their vector production yield and to benchmark against the parental
86 AAV8 capsid. Overall, the AAV8 capsid variants were less efficient than AAV8, exhibiting 63% or lower
87 vector production yield **(Figure 1a)**. We arbitrarily selected the top 37 variants for the subsequent
88 library screen, as low vector production yield will pose a translational hurdle for a gene therapy delivery
89 vehicle.

90 Next, we constructed a series of gene-of-interest (GOI) plasmids, each expressing a barcoded,
91 non-coding Tough Decoy (TuD) RNA under the control of the U6 promoter as previously reported[29]
92 **(Supplementary Figure 1)**. We chose not to use a protein-coding reporter transgene, such as *EGFP,*
93 to avoid potential immunogenicity when conducting the screen in large animals such as ferrets. A pair of
94 unique packaging and GOI plasmids were used in large-scale AAV vector production, so that each

95 barcoded *TuD* transgene represented a unique capsid. These vectors were individually purified,
96 followed by vector genome (vg) titer determination. According to their vg titers, these vectors were then
97 pooled at various volumes with the goal of equal representation for each capsid vector in the library. We
98 extracted the pooled library vector DNA, amplified the barcode-containing region by PCR, and
99 performed high-throughput nanopore sequencing to quantify the relative abundance of each barcode.
100 This abundance served as a proxy for the relative representation of each capsid vector in the library,
101 with the potential caveat that different barcode sequences may introduce PCR bias. We found that the
102 abundance of the barcodes representing 39 capsid vectors (i.e., AAV8, AAV9, and the 37 AAV8
103 variants) deviated from the average by less than 2.5-fold **(Figure 1b)**, comparable to other barcoded
104 libraries generated in the same manner (approximately 7-fold)[30]. In contrast, purifying all capsid vectors
105 in bulk would likely result in skewed representation in favor of "good producers" (e.g., more than 200-
106 fold deviation from the average[30]), which may cause bias in subsequent functional screens.
107 Alternatively, individual capsid titers could be adjusted at the crude lysate stage prior to pooling for bulk
108 purification. However, we reasoned that purifying and storing individual vectors would provide greater
109 flexibility for downstream applications, such as enabling the use of single vectors or selected subsets
110 for specific tests.

111

112 **Vector library screen in mice and ferrets**

113 The pooled vector library was delivered to three adult wild-type (WT) mice at 2×10^{13} vg/kg via
114 tail vein. Four weeks post-treatment, the mice were euthanized for tissue collection. The total DNA
115 extracted from livers or the vector library was subjected to PCR to amplify the barcode-containing
116 region in vector DNA. The barcodes present in the amplicons were quantified by Illumina sequencing
117 and normalized to their relative abundance in the vector library. Importantly, the relative abundance of
118 each barcoded amplicon in the library was consistent with the results obtained by nanopore sequencing
119 **(Figure 1b, Supplementary Figure 2)**. This analysis revealed that no candidate variants could surpass

120 AAV8 for liver gene delivery **(Figure 2a)**. In parallel, the same vector library was screened in three WT
121 ferrets sero-negative for AAV8 and AAV9 **(Supplementary Figure 3)** in the same fashion, which led to
122 the similar finding that AAV8 outperformed all candidate variants **(Figure 2b)**. Notably, the candidate
123 variants generally exhibited a similar trend of liver gene delivery efficiency in mice and ferrets **(Figure**
124 **2c)**, suggesting the robustness of the screening pipeline. As several studies have shown that low AAV
125 vector DNA abundance in tissues does not necessarily result in low functional transduction (i.e.,
126 transgene expression level)[31-33], we also quantified the barcoded *TuD* RNA levels in mouse livers.
127 However, consistent with the vector DNA analysis, AAV8 outperformed all variants in terms of
128 transgene expression **(Figure 3a**, blue bars**)**.

129

130 **Liver detargeting by several capsid variants**

131 In both mice and ferrets, we consistently observed that several capsid variants, including v2, v5,
132 v12, v13, v15, v23, v25, and v36, showed very low or barely detectable levels of liver gene delivery
133 **(Figure 2a, b)**. As expected, these candidate variants also led to low or barely detectable *TuD* RNA
134 levels in the mouse livers **(Figure 3a)**. To investigate whether these capsid variants lost liver-specific or
135 pan-tissue tropism, we quantified the *TuD* RNA levels in the heart and tibialis anterior (TA) muscle from
136 the treated mice. Taken together, the RNA analysis revealed that several variants, including v2, v5, v23,
137 v25, and v36, showed a strong liver detargeting phenotype, albeit with moderate reductions in *TuD*
138 expression in the heart and TA muscle. In contrast, transduction by v12, v13, and v15 failed in all three
139 tissue types **(Figure 3a)**. We were not able to detect *TuD* transcripts in ferrets tissues, likely due to the
140 overall low gene delivery efficiency in ferrets as compared to that in mice **(Supplementary Figure 4)**.

141 The N271D (AAV8 VP1 numbering) residue change in v5 is located in VR-1 **(Supplementary**
142 **Table 2)**, which had been implicated in modulating liver tropism in recent publications[22-24] **(Figure 3b)**

143 (also see Discussion). Therefore, we focused on this variant for further validation by single vector
144 treatment **(Supplementary Figure 5)**.

145

146 **Validation of liver detargeting in mice by single vector administration**

147 To further characterize the impact of N271D on tissue tropism in mice, we generated a pair of
148 AAV8 and AAV8.N271D vectors packaging the same *EGFP* transgene cassette, and treated two groups
149 of mice with the two vectors, respectively. Four weeks post-treatment, the mice were euthanized to
150 compare vector DNA abundance, *EGFP* mRNA levels, and EGFP protein levels in the liver, heart, and
151 TA muscle. Consistent with the library screen results, the AAV8.N271D.EGFP vector showed >100-fold
152 reductions in vector DNA abundance and transgene expression in the liver as compared to the
153 AAV8.EGFP vector, while heart and TA muscle targeting was moderately impacted **(Figure 4a,**
154 **Supplementary Figure 6a)**.

155 As N271 (AAV8 VP1 numbering) is highly conserved among multiple serotypes **(Figure 3b)**, we
156 tested whether grafting the N271D residue change to other AAV capsids could confer liver detargeting.
157 To this end, we introduced the homologous N270D (AAV9 VP1 numbering) mutation into the AAV9
158 capsid and generated the AAV9.N270D.EGFP vector. Following administration to WT mice, we
159 observed a dramatic liver detargeting phenotype as compared to the parental AAV9.EGFP vector
160 **(Figure 4b, Supplementary Figure 6b)**. Regarding transducing the heart and TA muscle, N270D in
161 AAV9 capsid had a larger negative impact than N271D in AAV8 (compare the middle panels in **Figures**
162 **4a and 4b**). MyoAAV is an engineered capsid with a 7-mer peptide insertion in AAV9 capsid proteins,
163 and exhibits enhanced muscle tropism as compared with AAV9 in mice and monkeys[34]. To test whether
164 the enhanced muscle tropism of MyoAAV can compensate for the reduced muscle targeting by
165 AAV9.N270D, we generated a pair of MyoAAV.EGFP and MyoAAV.N270D.EGFP vectors and
166 administered to mice. We consistently observed liver detargeting by MyoAAV.N270D in mice **(Figure**

167 **4c, Supplementary Figure 6c)**. Although MyoAAV.N270D.EGFP led to lower transduction levels in the
168 heart and TA muscle as compared to MyoAAV.EGFP, it still outperformed AAV9.EGFP and AAV8.EGFP
169 (compare the middles panels of **Figures 4a-c**). Taken together, the N271D (AAV8 VP1 numbering)
170 variant consistently conferred liver detargeting in the context of all three parental capsids tested
171 **(Supplementary Figure 7)**.

172

173 **DISCUSSION**

174 Residues in VR-1 have been shown to influence liver tropism of AAV vectors. Warischalk tested
175 several VR-1 amino acid deletion mutants across multiple serotype capsids, and found that some
176 detargeted from transducing mouse liver by orders of magnitude[24]. Using a domain swapping strategy,
177 Cabanes-Creus et al. found that a single threonine insertion downstream of residue 264 of AAV3B
178 (AAV3B-265*insT*, **Figure 3b**) greatly enhanced liver targeting in mouse[22]. In another study, Zinn et al.
179 identified residue 267 (AAV9 VP1 numbering), named “liver toggle”, as a key determinant for liver
180 targeting[23] **(Figure 3b)**. For example, a single G267A mutation in the AAV9 capsid (AAV9-GA) resulted
181 in significant liver detargeting with respect to both vector genomes (1,522-fold) and expression (363-
182 fold), with no significant difference in terms of gene transfer to the heart and quadriceps, and a
183 moderate reduction in expression in these tissues (1.5-fold and 3.1-fold reductions in quadriceps and
184 heart RNA, respectively)[23]. The N271D (AAV8 VP1 numbering) variant identified in this study is in VR-1
185 and in proximity with the previously described residues **(Figure 3b)**, and also shows a profound impact
186 on liver targeting. Together, these converging evidence points to the important role of VR-1 in
187 determining liver targeting. Although the detailed mechanism remains to be elucidated, it may involve
188 the proposed VR-1 hydrogen bond network[24], and/or interactions with AAVR[35], a cellular factor that
189 binds to most AAV capsids and facilitates their cell entry and intracellular trafficking, as structural
190 studies have shown that VR-1 is part of the AAVR-AAV binding footprint[36,37]. Interestingly, AAV capsid
191 deamidation where an amide group in the side chain of amino acid is converted to a carboxylic acid

192 group, such as N-to-D changes, has been shown to impact transduction efficiency[38]. Further saturating
193 mutagenesis studies will help determine whether the N271D variant exerts its effect through the loss of
194 asparagine (N) or the gain of aspartic acid (D).

195 Liver toxicity observed in some preclinical and clinical studies has highlighted the need for AAV
196 capsids with liver-detargeting properties, particularly for gene delivery to extrahepatic tissues via
197 systemic administration[39]. Several engineered capsids, such as AAV2i8[13], AAVMYO[30], MyoAAV[34],
198 PHPeB[40], and VCAP-102[41], exhibit both enhanced muscle or CNS targeting and modest liver-
199 detargeting (approximately 3- to 30-fold reduction in liver tropism relative to their parental capsids).
200 However, these liver-detargeting effects were typically incidental findings during biodistribution profiling,
201 rather than outcomes of intentional design. Moreover, enhanced targeting of extrahepatic tissues does
202 not necessarily lead to reduced liver-detargeting. For example, the engineered capsid BI-hTFR1
203 engages the human transferrin receptor to facilitate blood-brain barrier penetration, but retains liver
204 tropism comparable to its parental AAV9 capsid[42]. Therefore, a modular capsid modification that confers
205 liver-detargeting and can be applied across different capsid backbones represents a promising
206 complementary strategy in capsid engineering. The dramatic liver detargeting phenotype of the
207 AAV8.N271D, AAV9.N270D, and MyoAAV.N270D vectors is accompanied by moderate reductions in
208 targeting the heart and TA muscle **(Figure 4)**. A similar trend (i.e., profound liver detargeting with a
209 moderate impact on heart and quadriceps targeting) was also observed when reprogramming AAV9
210 and Anc80 with the liver toggle residue changes[23]. This creates a dilemma in developing liver-
211 detargeting vectors for gene therapy delivery to muscle tissues. However, the MyoAAV.N270D vector
212 partially retains the excellent muscle targeting property of the parental MyoAAV capsid, and exhibits
213 higher gene transfer and expression levels in the heart and TA muscle than the AAV8 and AAV9
214 vectors. Furthermore, it may be possible to engineer residues in VR-1, either alone or in combination
215 with changes in other capsid regions, to develop capsids with true liver-specific detargeting capabilities.

216 For example, incorporating the tyrosine mutant Y445F[43] into the T265del AAV6 capsid can further
217 enhance cardiac transduction while maintaining a liver-detargeted biodistribution profile[24].

218 In summary, this study identifies a naturally occurring AAV capsid variant with a liver detargeting
219 phenotype, reinforces the critical role of VR-1 in modulating liver tropism and transgene expression,
220 and opens new avenues for engineering AAV vectors that minimize liver targeting and associated
221 toxicity.

222

223 **MATERIALS AND METHODS**

224 **AAV constructs**

225 The reporter transgene construct used in the small-scale AAV vector production assay contains
226 an *EGFP* transgene under the control of the CMV enhancer and chicken beta-actin promoter. The
227 same construct was packaged in multiple capsids for single vector administration in mice. The
228 barcoded transgene construct used for generating AAV vector library contains a *TuD* transgene under
229 the control of the U6 promoter (also see **Supplementary Figure 1**). The detailed design of barcode has
230 been described previously[29]. Briefly, the barcode contains randomized 8 nucleotides that are embedded
231 in the microRNA binding site (MBS) between stem I and stem II of the *TuD* transgene DNA, and it is
232 also present in the *TuD* RNA following expression. All AAV constructs were designed to package single-
233 stranded vector genomes.

234 **Small-scale AAV vector production assay**

235 The details of the procedure have been described previously[44]. Briefly, a pair of packaging
236 plasmid and gene-of-interest (GOI) plasmid were co-transfected to HEK293 cells along with a helper
237 plasmid in a 12-well plate using the calcium phosphate method. 72 hours post-transfection, cells and
238 culture media were collected to generate crude lysates following three freeze-and-thaw cycles. The

239 crude lysates were clarified by centrifugation and sequentially treated with DNase I and proteinase K,
240 followed by vector genome determination using droplet digital PCR (ddPCR).

241 **Nanopore sequencing of amplicons**

242 AAV vector DNA was extracted with the QIAamp DNA Micro Kit (Qiagen, 56304), and subjected
243 to PCR amplification. The amplicons were purified with the DNA Clean & Concentrator-5 kit (Zymo,
244 D4013) and sequenced with the Oxford Nanopore platform (Plasmidsaurus). Reads were mapped to
245 the reference sequence, demultiplexed by barcodes, and counted using the Geneious Prime software.

246 **Large-scale AAV vector production**

247 Adherent HEK293 cells were cultured in roller bottles in Dulbecco's Modified Eagle Medium
248 (DMEM) with 10% fetal bovine serum. Helper plasmid, packaging plasmid, and GOI plasmid were co-
249 transfected using the calcium phosphate method. Next day, the cell culture medium was replaced with
250 fresh, serum-free DMEM. Three days post-transfection, cells and culture media were harvested and
251 subjected to purification using two rounds of cesium chloride gradient centrifugation. AAV vectors were
252 dialyzed against phosphate buffered saline and sterilized by passing through a 0.22 μm filter.

253 **Animal work**

254 Wildtype, male C57BL/6J mice were treated with AAV vectors via tail vein at six weeks of age,
255 and euthanized four weeks later. For library screen, the dose was 5×10^{11} vg/mouse (approximately $2 \times$
256 10^{13} vg/kg); for single vector administration, the dose was 3×10^{11} vg/mouse. Wildtype ferrets were
257 subjected to anti-AAV8 and anti-AAV9 neutralizing antibody (NAb) screen. Three ferrets (one female
258 and two males) were sero-negative (<1:5 NAb titers), and therefore were used in experiment. They
259 were treated at 10 weeks of age with the AAV vector library via intravenous injection at the dose of $2 \times$
260 10^{13} vg/kg, and euthanized five weeks later. All animal studies were reviewed and approved by the
261 Institutional Animal Care and Use Committee of the University of Massachusetts Chan Medical School.

262 **Neutralizing antibody assay**

263 The neutralizing antibody (NAb) assay was performed as previously described[45]. Briefly, serum
264 samples were heat-inactivated under 56°C for 35 minutes, and 1:5 diluted in Dulbecco's Modified
265 Eagle's Medium (DMEM) followed by 2-fold serial dilutions. Each diluted serum sample was incubated
266 with 1.2×10^9 genome copies (GC) of rAAV8.LacZ (for anti-AAV8 assay) or rAAV9.LacZ (for anti-AAV9
267 assay) under 37°C/5% CO_2 for 1 hour. Huh7 cells grown in a 96-well plate were first infected with
268 wildtype adenovirus serotype 5 (100 viral particles per cell). Three to four hours later, 100 μl of pre-
269 incubated rAAV/serum mixture containing 1×10^9 GC of rAAV.LacZ was added to each well of Huh7
270 cell culture and incubated under 37°C/5% CO_2 for 1 hour. Cells were then cultured with 5% fetal bovine
271 serum under 37°C/5% CO_2 overnight. β-galactosidase activity in cell lysate was measured using the
272 Galacto-Star™ One-Step β-galactosidase Reporter Gene Assay System (Thermo Fisher Scientific, Cat.
273 No. T1014). The transduction inhibition effect of a test sample at each dilution was calculated by
274 comparing with the negative control at the same dilution (i.e., maximal β-galactosidase activity without
275 inhibition). The NAb titer was defined as a range between the lowest dilution factor that yielded more
276 than 50% transduction inhibition and the next dilution factor that could not inhibit transduction by more
277 than 50%.

278 **Quantification of barcodes by Illumina sequencing**

279 DNA and RNA were isolated from liver samples using the AllPrep DNA/RNA kit (Qiagen, 80204).
280 For muscle tissues, DNA was extracted with the AllPrep DNA/RNA kit (Qiagen, 80204), while RNA was
281 extracted using Trizol. RNA was subjected to treatment with DNase (Qiagen 79254) followed by a
282 clean-up step using the RNA Clean & Concentrator-5 kit (Zymo, R1014). Purified RNA was reverse-
283 transcribed into cDNA using the SuperScript III First-Strand Synthesis System (Thermo Fisher
284 Scientific, 18080051). PCR was performed with barcoded primers and the KOD Hotstart Master Mix
285 (Millipore, 71842) in a total volume of 25 μL that contained 12.5 μL of KOD Hotstart Master Mix, 0.75 μL
286 of forward primer (10 μM), 0.75 μL of reverse primer (10 μM), 50 ng of DNA. Primer binding sites are

287 shown in **Supplementary Figure 1**. PCR cycling condition: 95°C for 2 min, followed by 25 cycles of
288 95°C/20s, 58°C/10s, and 70°C/10s. PCR amplicons were gel purified and concentrations were
289 determined using a Qubit 3 Fluorometer (Thermo Fisher Scientific). PCR amplicons were pooled at an
290 equal ratio and sequenced using MiSeq at UMass Chan Deep Sequencing Core. Sequencing data
291 were analyzed using the Galaxy platform[46]. The data were first demultiplexed by the barcodes in PCR
292 primers **(Supplementary Table 3)** using the "barcode splitter" module. This step resulted in datasets
293 representing individual tissue samples. Next, the resulting datasets were demultiplexed by the
294 barcodes in the *TuD* gene **(Supplementary Table 4)**, which led to sequencing reads representing their
295 corresponding capsid variants. Barcode and PCR designs are shown in **Supplementary Figure 1**.

296 **Quantification of vector DNA and RNA**

297 DNA and RNA were isolated from mouse tissues using the AllPrep DNA/RNA kit (Qiagen,
298 80204). RNA was reverse-transcribed into cDNA with the High-Capacity cDNA Reverse Transcription kit
299 (Thermo Fisher Scientific, 43-688-13). To quantify vector DNA in the tissues collected from the mice
300 and ferrets treated with the vector library, a duplexing Taqman droplet digital PCR (ddPCR) assay was
301 performed with one reagent targeting stuffer B in the GOI plasmid [forward primer:
302 AGCCCTAGGGATGAACCAGT, reverse primer: AACCCAGGAGTCATTGCATC, probe:
303 AATCTGAGCCACTGAGCCAT, all synthesized by Integrated DNA Technologies (IDT)] and a Taqman
304 reagent targeting mouse *Tfrc* (Thermo Fisher Scientific, 4458367) or ferret *Tfrc* (forward primer:
305 GCTTTAGACCCAGCAGAAGC, reverse primer: GTTCAGCTGCCCATTCTGAG, probe:
306 TCTCCAGCACTCCAGCTGGCA, all synthesized by IDT). To quantify *EGFP* vector DNA in mouse
307 tissues, a duplexing Taqman ddPCR assay was performed with one reagent targeting *EGFP* (Thermo
308 Fisher Scientific, Mr00660654_cn), and the other targeting mouse *Tfrc* (Thermo Fisher Scientific,
309 4458367). cDNA was quantified using a duplexing Taqman ddPCR assay with the same *EGFP*-
310 targeting Taqman reagent as mentioned above, and a Taqman reagent targeting mouse *Gapdh*
311 (Thermo Fisher Scientific, 4352339E). ddPCR was performed with a QX200 instrument (Bio-Rad)

312 under the cycling condition: 95°C for 10 min, 40 cycles of 94°C/30s and 61°C/60s, followed by
313 98°C/10min.

314 **Western blot**

315 Mouse tissues were homogenized using TissueLyser II (Qiagen) in ice-cold T-PER (Thermo
316 Fisher Scientific, 78510) with protease inhibitor (Roche, 4693159001). Total protein concentration in
317 tissue lysate was determined with Pierce BCA Protein Assay Kit (Pierce, 23225). Protein lysates
318 normalized for total protein amount were boiled with 4× Laemmli sample buffer (Bio-Rad, 1610747) at
319 99 °C for 5 min. Primary antibodies: mouse anti-EGFP Antibody (Abcam, ab184601, 1:5000), rabbit
320 anti-GAPDH (Abcam, ab9485, 1:10000). Secondary antibodies: LICOR IRDye 680RD goat anti-mouse
321 IgG (H + L) (LI-COR Biosciences, 926-68070, 1:7000), LICOR IRDye 800CW goat anti-rabbit IgG (H +
322 L) (LI-COR Biosciences, 926-32211, 1:7000). Blot membranes were imaged using a LI-COR scanner
323 (Odyssey) and quantified by Li-Cor software.

324 **Data availability**

325 Nanopore sequencing and Illumina sequencing data are deposited in NCBI Sequence Read
326 Archive (SRA) and available with accession ID PRJNA1268899.

327

328 **ACKNOWLEDGEMENTS**

329 This study was supported by a grant from the National Institutes of Health (NIH) (P01HL158506
330 to D.W.). The Gao Lab is supported by grants from the NIH (R01NS076991-01, P01AI100263-01,
331 P01HL131471-02, 35 R01AI121135, UG3HL147367-01, R01HL097088, and U19AI149646-01) and
332 Cystic Fibrosis Foundation.

333

334 **AUTHOR CONTRIBUTIONS**

R.X., D.W. designed the research. R.X., Mengyao X. conducted the majority of the experiments. D.R., A.D., A.M.G. performed injections in ferrets and other ferret work. R.X., M.C., P.W.L.T. contributed to Illumina sequencing data analysis. G.X., L.L., Meiyu X., F.Z., P.W.L.T., Y.W., G.G. provided the AAV8 variants. N.L. produced AAV vectors. J.L. performed mouse injections and other mouse work. T.R.F., G.G., D.W. supervised the study. R.X., Mengyao X., D.W. analyzed data and wrote the manuscript.

DECLARATION OF INTERESTS

G.X., L.L., P.W.L.T., Y.W., G.G. are inventors of a patent application regarding the AAV8 variants described in this study (US20230374545A1). G.G. is a scientific co-founder of Voyager Therapeutics, Adrenas Therapeutics and Aspa Therapeutics and holds equity in these companies.

KEYWORDS

AAV, liver detargeting, liver tropism, variable region

FIGURE LEGENDS

Figure 1. Generation and characterization of the AAV8 vector library. (a) Dot plot showing the production yield of each AAV8 capsid variant in a small-scale vector production assay packaging the same *EGFP* transgene cassette. Each dot represents a unique capsid variant. The production yield is normalized to the AAV8 vector level (defined as 100%) and ranked from the highest to the lowest. Data were based on one biological repeat. **(b)** Bar graph showing the count of Nanopore sequencing reads mapped to the unique vector transgene barcodes packaged in AAV8, AAV9, or AAV8 variants. The dashed line shows the mean value. Data were based on one biological repeat.

357

358 **Figure 2. Vector library screen in mice and ferrets. (a, b)** Scatter dot plots showing the relative
359 vector genome abundance of each barcoded capsid vector in mouse (a) or ferret (b) livers. Data are
360 normalized to the AAV8 vector level (defined as 1.0), and presented as mean ± SD. Each dot
361 represents an individual animal. v1 to v37 denote the identifiers of AAV8 variants. The variants in bold
362 indicate that they resulted in low or barely detectable levels of vector DNA. **(c)** Scatter plot showing the
363 relationship between the relative vector DNA abundance of each capsid variant in mouse liver (x-axis)
364 and ferret liver (y-axis). Data are presented as mean and standard deviation of three animals. The
365 linear regression statistics are shown.

366

367 **Figure 3. AAV8.v5 shows liver detargeting phenotype in mice. (a)** Bar graph showing the relative
368 vector RNA (cDNA) abundance of each barcoded capsid vector in mouse liver (blue), heart (orange),
369 and tibialis anterior (TA) muscle (green). Data are normalized to the AAV8 vector level in respective
370 tissues (defined as 1), and presented as mean ± SD. Each dot represents an individual animal. v1 to
371 v37 denote the identifiers of AAV8 variants. The variants in bold indicate that they resulted in low or
372 barely detectable levels of vector RNA expression in the liver, with the ones underlined indicating well-
373 detectable levels of vector RNA expression in the heart and/or TA muscle. **(b)** Alignment of the amino
374 acid sequences of multiple AAV capsids. Only variable region I (VR-1) and surrounding residues are
375 shown as single-letter abbreviations. Dashes indicate gaps. The 265*insT* (AAV3B VP1 numbering)
376 residue described in Cabanes-Creus M. et al. 2021[22], the liver toggle residue described in Zinn E. et al.
377 2022[23], and the N271D (AAV8 VP1 numbering) residue described in this study are highlighted with gray
378 background. The residues that differ from parental capsids described in these studies are highlighted in
379 bold. The Warischalk study[24] investigated a wide range of VR-1 mutants across multiple serotype
380 capsids.

Figure 4. Characterization of vector performance in mice following individual vector administration. (a-c) Box plots showing the vector DNA abundance (left), *EGFP* mRNA levels (middle), and EGFP protein levels (right) in the liver, heart, and tibialis anterior (TA) muscle collected from the mice treated with AAV8 or AAV8.N271D vectors (a), AAV9 or AAV9.N270D vectors (b), and MyoAAV or MyoAAV.N270D vectors (c). Each dot represents an individual mouse. The box extends from the first to the third quartiles with the line inside denoting median. The whiskers end at minimum and maximum values. The fold changes of medians and p values are labeled. Statistical analysis is performed using non-parametric Mann-Whitney test.

REFERNCES

1 Muzyczka, N. & Berns, K. in *Fields Virology* Vol. 2 (eds D Knipe *et al.*) 2327-2359 (Lippincott, Williams and Wilkins, 2001).

2 Wang, D., Tai, P. W. L. & Gao, G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nature reviews. Drug discovery* **18**, 358-378, doi:10.1038/s41573-019-0012-9 (2019).

3 Gao, G. *et al.* Adeno-associated viruses undergo substantial evolution in primates during natural infections. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 6081-6086, doi:10.1073/pnas.0937739100 (2003).

4 Gao, G. *et al.* Clades of Adeno-associated viruses are widely disseminated in human tissues. *Journal of virology* **78**, 6381-6388, doi:10.1128/JVI.78.12.6381-6388.2004 (2004).

5 ASGCT. *Gene, Cell, & RNA Therapy Landscape Report (Q4 2024 Quarterly Data Report)*, <https://www.asgct.org/global/documents/asgct-citeline-q4-2024-report.aspx> (2025).

6 Li, C. & Samulski, R. J. Engineering adeno-associated virus vectors for gene therapy. *Nature reviews. Genetics* **21**, 255-272, doi:10.1038/s41576-019-0205-4 (2020).

7 Agbandje-McKenna, M. & Kleinschmidt, J. AAV capsid structure and cell interactions. *Methods Mol Biol* **807**, 47-92, doi:10.1007/978-1-61779-370-7_3 (2011).

8 Bowles, D. E. *et al.* Phase 1 gene therapy for Duchenne muscular dystrophy using a translational optimized AAV vector. *Molecular therapy : the journal of the American Society of Gene Therapy* **20**, 443-455, doi:10.1038/mt.2011.237 (2012).

9 Li, C. *et al.* Single amino acid modification of adeno-associated virus capsid changes transduction and humoral immune profiles. *Journal of virology* **86**, 7752-7759, doi:10.1128/JVI.00675-12 (2012).

10 Grimm, D. *et al.* In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *Journal of virology* **82**, 5887-5911, doi:10.1128/JVI.00254-08 (2008).

11 Lerch, T. F. *et al.* Structure of AAV-DJ, a retargeted gene therapy vector: cryo-electron microscopy at 4.5 A resolution. *Structure* **20**, 1310-1320, doi:10.1016/j.str.2012.05.004 (2012).

12 Large, E. E., Silveria, M. A., Zane, G. M., Weerakoon, O. & Chapman, M. S. Adeno-Associated Virus (AAV) Gene Delivery: Dissecting Molecular Interactions upon Cell Entry. *Viruses* **13**, doi:10.3390/v13071336 (2021).

13 Asokan, A. *et al.* Reengineering a receptor footprint of adeno-associated virus enables selective and systemic gene transfer to muscle. *Nature biotechnology* **28**, 79-82, doi:10.1038/nbt.1599 (2010).

14 Tseng, Y. S. & Agbandje-McKenna, M. Mapping the AAV Capsid Host Antibody Response toward the Development of Second Generation Gene Delivery Vectors. *Frontiers in immunology* **5**, 9, doi:10.3389/fimmu.2014.00009 (2014).

15 Hsu, H. L. *et al.* Structural characterization of a novel human adeno-associated virus capsid with neurotropic properties. *Nat Commun* **11**, 3279, doi:10.1038/s41467-020-17047-1 (2020).

16 Mahlangu, J. *et al.* Two-Year Outcomes of Valoctocogene Roxaparvovec Therapy for Hemophilia A. *The New England journal of medicine* **388**, 694-705, doi:10.1056/NEJMoa2211075 (2023).

17 Pipe, S. W. *et al.* Gene Therapy with Etranacogene Dezaparvovec for Hemophilia B. *The New England journal of medicine* **388**, 706-718, doi:10.1056/NEJMoa2211644 (2023).

18 Strnad, P., McElvaney, N. G. & Lomas, D. A. Alpha(1)-Antitrypsin Deficiency. *The New England journal of medicine* **382**, 1443-1455, doi:10.1056/NEJMra1910234 (2020).

19 Blackwood, M. *et al.* Biodistribution and safety of a single rAAV3B-AAT vector for silencing and replacement of alpha-1 antitrypsin in Cynomolgus macaques. *Mol Ther Methods Clin Dev* **32**, 101200, doi:10.1016/j.omtm.2024.101200 (2024).

20 Gao, G., Xu, G., Tai, P. W., Wei, Y. & Luo, L. AAV capsid designs. US20230374545A1 (2023).

21 He, N. *et al.* Ferret models of alpha-1 antitrypsin deficiency develop lung and liver disease. *JCI Insight* **7**, doi:10.1172/jci.insight.143004 (2022).

22 Cabanes-Creus, M. *et al.* Single amino acid insertion allows functional transduction of murine hepatocytes with human liver tropic AAV capsids. *Mol Ther Methods Clin Dev* **21**, 607-620, doi:10.1016/j.omtm.2021.04.010 (2021).

23 Zinn, E. *et al.* Ancestral library identifies conserved reprogrammable liver motif on AAV capsid. *Cell Rep Med* **3**, 100803, doi:10.1016/j.xcrm.2022.100803 (2022).

24 Warischalk, J. K. *Uncovering structural components of the adeno-associated viral capsid that can be modified to improve clinical gene therapy outcomes* Doctor of Philosophy thesis, University of North Carolina at Chapel Hill, (2015).

25 Chand, D. *et al.* Hepatotoxicity following administration of onasemnogene abeparvovec (AVXS-101) for the treatment of spinal muscular atrophy. *J Hepatol* **74**, 560-566, doi:10.1016/j.jhep.2020.11.001 (2021).

26 Hudry, E. *et al.* Liver injury in cynomolgus monkeys following intravenous and intrathecal scAAV9 gene therapy delivery. *Molecular therapy : the journal of the American Society of Gene Therapy* **31**, 2999-3014, doi:10.1016/j.ymthe.2023.07.020 (2023).

27 Hordeaux, J. *et al.* High-dose systemic adeno-associated virus vector administration causes liver and sinusoidal endothelial cell injury. *Molecular therapy : the journal of the American Society of Gene Therapy* **32**, 952-968, doi:10.1016/j.ymthe.2024.02.002 (2024).

28 Whiteley, L. O. An Overview of Nonclinical and Clinical Liver Toxicity Associated With AAV Gene Therapy. *Toxicol Pathol* **51**, 400-404, doi:10.1177/01926233231201408 (2023).

29 Xu, M. *et al.* High-Throughput Quantification of In Vivo Adeno-Associated Virus Transduction with Barcoded Non-Coding RNAs. *Human gene therapy* **30**, 946-956, doi:10.1089/hum.2018.253 (2019).

30 Weinmann, J. *et al.* Identification of a myotropic AAV by massively parallel in vivo evaluation of barcoded capsid variants. *Nat Commun* **11**, 5432, doi:10.1038/s41467-020-19230-w (2020).

31 Salganik, M. *et al.* Adeno-associated virus capsid proteins may play a role in transcription and second-strand synthesis of recombinant genomes. *Journal of virology* **88**, 1071-1079, doi:10.1128/JVI.02093-13 (2014).

32 Gonzalez-Sandoval, A. *et al.* The AAV capsid can influence the epigenetic marking of rAAV delivered episomal genomes in a species dependent manner. *Nat Commun* **14**, 2448, doi:10.1038/s41467-023-38106-3 (2023).

33 Loeb, E. J. *et al.* Capsid-mediated control of adeno-associated viral transcription determines host range. *Cell Rep* **43**, 113902, doi:10.1016/j.celrep.2024.113902 (2024).

34 Tabebordbar, M. *et al.* Directed evolution of a family of AAV capsid variants enabling potent muscle-directed gene delivery across species. *Cell* **184**, 4919-4938 e4922, doi:10.1016/j.cell.2021.08.028 (2021).

35 Pillay, S. *et al.* An essential receptor for adeno-associated virus infection. *Nature* **530**, 108-112, doi:10.1038/nature16465 (2016).

36 Zhang, R. *et al.* Adeno-associated virus 2 bound to its cellular receptor AAVR. *Nat Microbiol* **4**, 675-682, doi:10.1038/s41564-018-0356-7 (2019).

37 Meyer, N. L. *et al.* Structure of the gene therapy vector, adeno-associated virus with its cell receptor, AAVR. *eLife* **8**, doi:10.7554/eLife.44707 (2019).

38 Giles, A. R. *et al.* Deamidation of Amino Acids on the Surface of Adeno-Associated Virus Capsids Leads to Charge Heterogeneity and Altered Vector Function. *Molecular therapy : the journal of the American Society of Gene Therapy* **26**, 2848-2862, doi:10.1016/j.ymthe.2018.09.013 (2018).

39 Asokan, A. & Shen, S. Redirecting AAV vectors to extrahepatic tissues. *Molecular therapy : the journal of the American Society of Gene Therapy* **31**, 3371-3375, doi:10.1016/j.ymthe.2023.10.005 (2023).

40 Deverman, B. E. *et al.* Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nature biotechnology* **34**, 204-209, doi:10.1038/nbt.3440 (2016).

41 Moyer, T. C. *et al.* Highly conserved brain vascular receptor ALPL mediates transport of engineered AAV vectors across the blood-brain barrier. *Molecular therapy : the journal of the American Society of Gene Therapy*, doi:10.1016/j.ymthe.2025.04.046 (2025).

42 Huang, Q. *et al.* An AAV capsid reprogrammed to bind human transferrin receptor mediates brain-wide gene delivery. *Science* **384**, 1220-1227, doi:10.1126/science.adm8386 (2024).

43 Qiao, C. *et al.* Adeno-associated virus serotype 6 capsid tyrosine-to-phenylalanine mutations improve gene transfer to skeletal muscle. *Human gene therapy* **21**, 1343-1348, doi:10.1089/hum.2010.003 (2010).

44 Ai, J., Ibraheim, R., Tai, P. W. L. & Gao, G. A Scalable and Accurate Method for Quantifying Vector Genomes of Recombinant Adeno-Associated Viruses in Crude Lysate. *Hum Gene Ther Methods* **28**, 139-147, doi:10.1089/hgtb.2016.173 (2017).

45 Wang, D. *et al.* Adeno-Associated Virus Neutralizing Antibodies in Large Animals and Their Impact on Brain Intraparenchymal Gene Transfer. *Mol Ther Methods Clin Dev* **11**, 65-72, doi:10.1016/j.omtm.2018.09.003 (2018).

46 Galaxy, C. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic acids research* **52**, W83-W94, doi:10.1093/nar/gkae410 (2024).

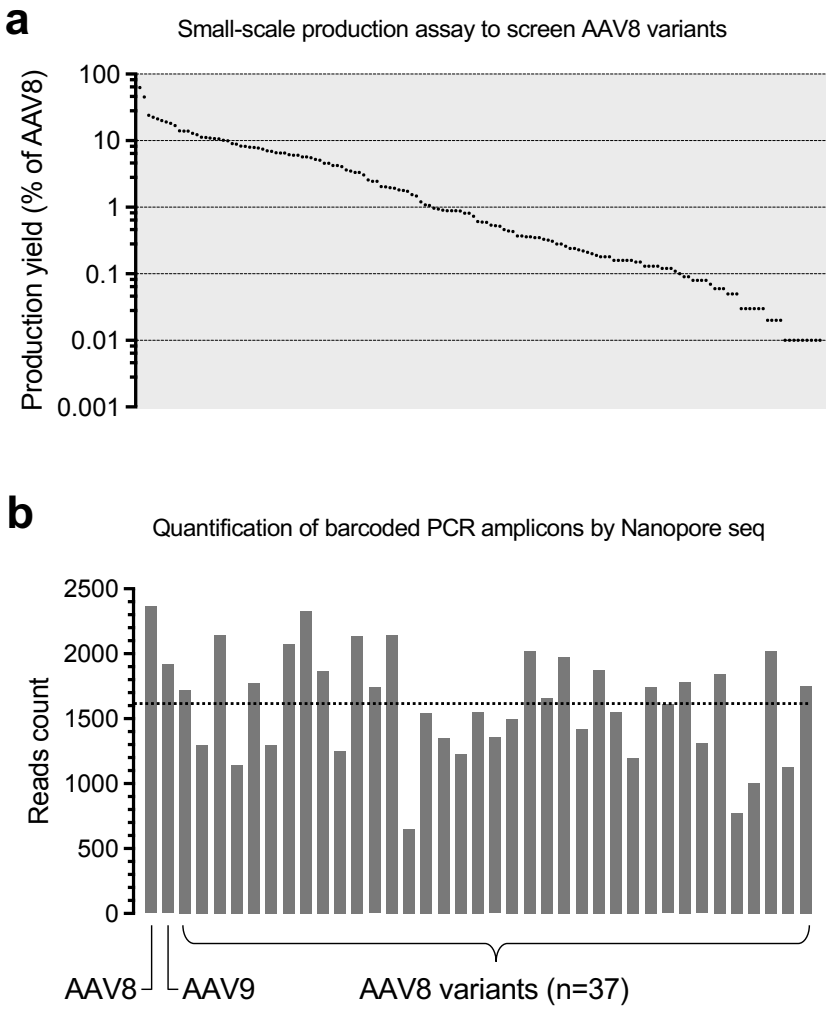

Figure 1. Generation and characterization of the AAV8 vector library.

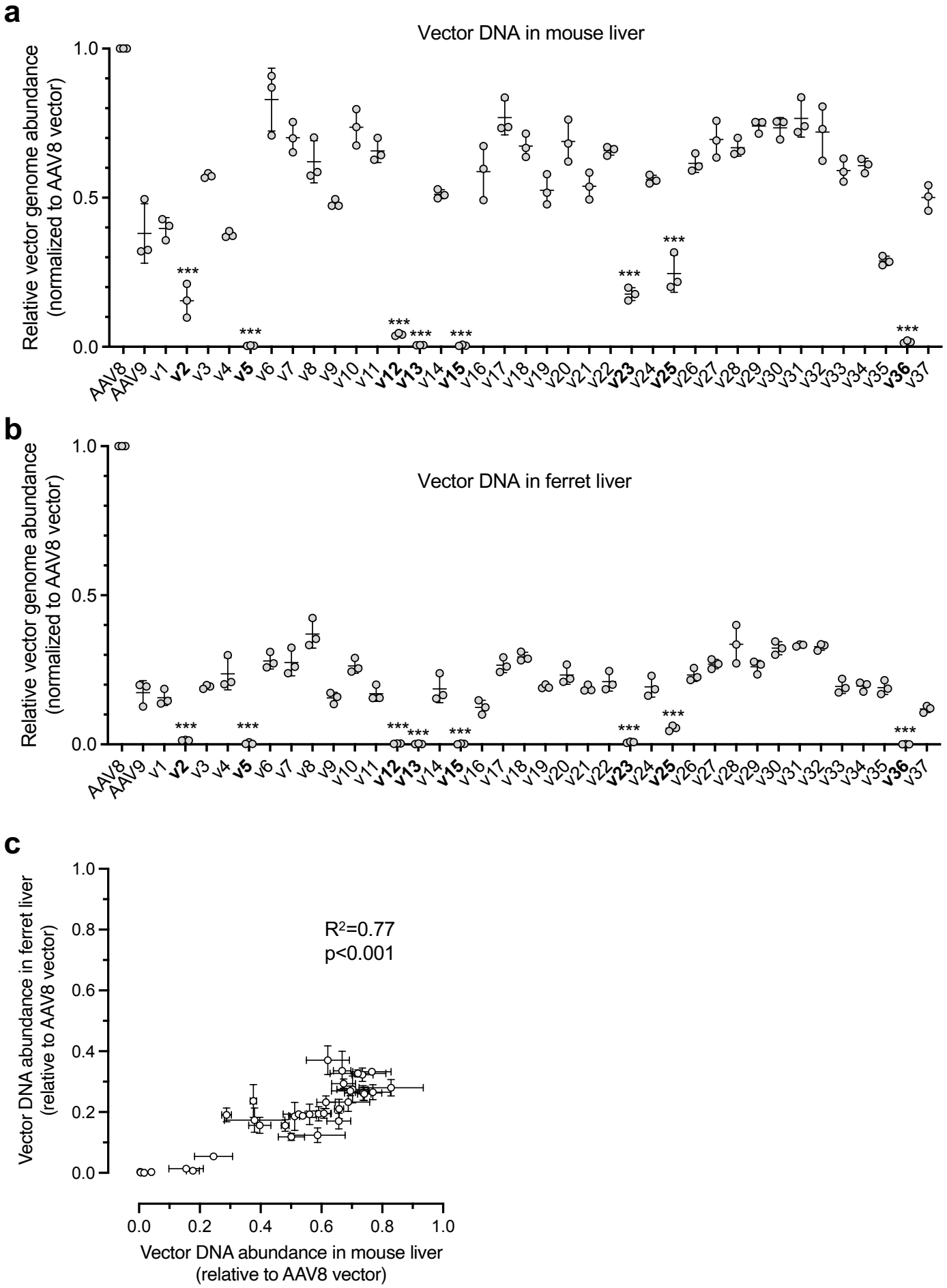

Figure 2. Vector library screen in mice and ferrets.

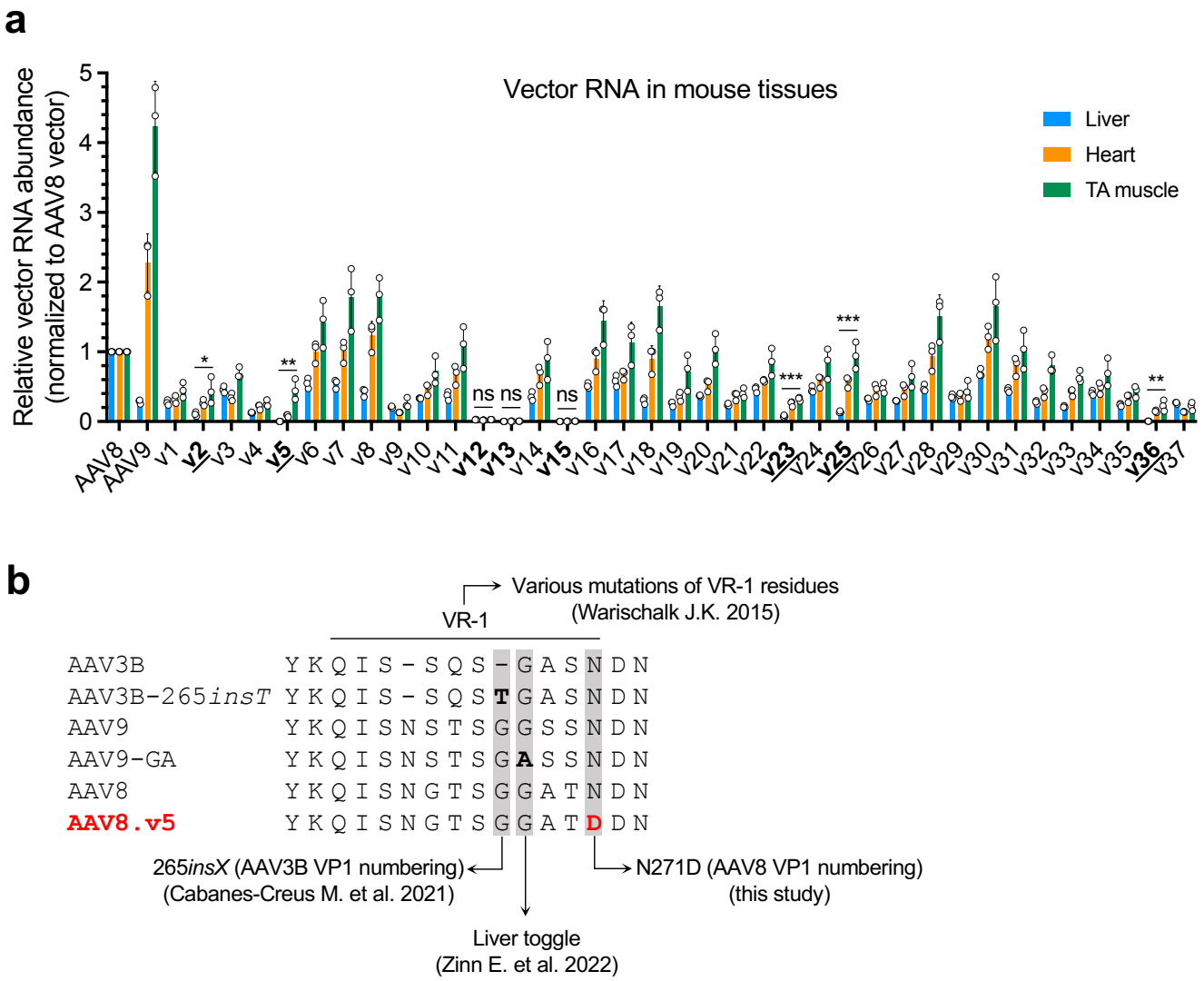

Figure 3. AAV8.v5 shows liver detargeting phenotype in mice.

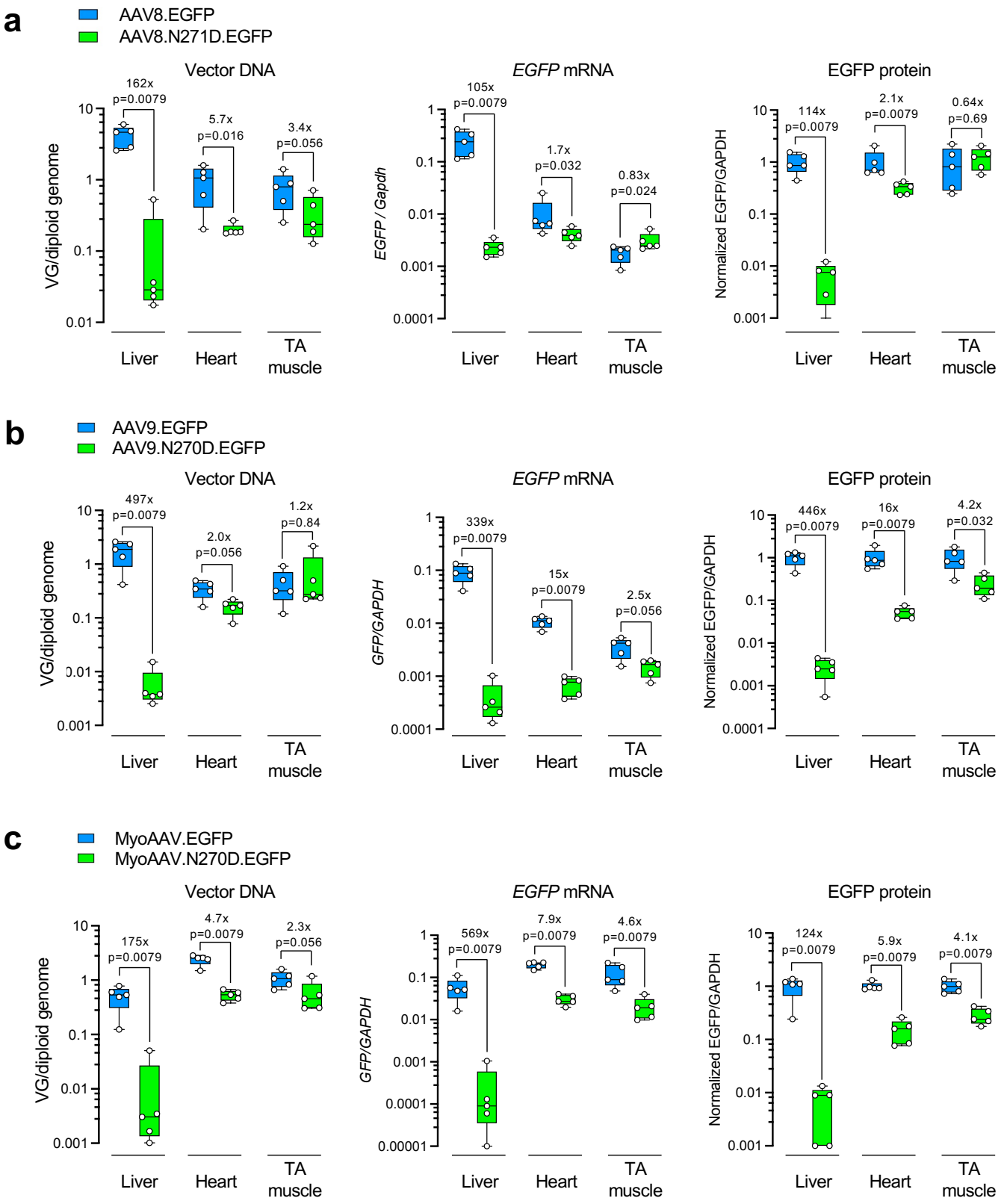

Figure 4. Characterization of vector performance in mice following individual vector administration.