

Natural variations in AAVHSC16 significantly reduce liver tropism and maintain broad distribution to periphery and CNS

Laura J. Smith,¹ Lindsay A. Schulman,¹ Samantha Smith,¹ Laura Van Lieshout,¹ Carmen M. Barnes,¹ Liana Behmoiras,¹ Meghan Scarpitti,¹ Monicah Kivaa,¹ Khanh L. Duong,¹ Ludo O. Benard,¹ Jeff L. Ellsworth,¹ Nancy Avila,¹ Deiby Faulkner,¹ April Hayes,¹ Jason Lotterhand,¹ Jose Israel Rivas,¹ Arnold V. Sengooba,¹ Alec Tzianabos,¹ Albert B. Seymour,¹ and Omar L. Francone¹

¹Homology Medicines, Inc., 1 Patriots Park, Bedford, MA 01730, USA

Adeno-associated viruses derived from human hematopoietic stem cells (AAVHSCs) are naturally occurring AAVs. Fifteen AAVHSCs have demonstrated broad biodistribution while displaying differences in transduction. We examine the structure-function relationships of these natural amino acid variations on cellular binding. We demonstrate that AAVHSC16 is the only AAVHSC that does not preferentially bind to terminal galactose. AAVHSC16 contains two unique amino acids, 501I and 706C, compared with other AAVHSCs. Through mutagenesis, we determined that residue 501 contributes to the lack of galactose binding. Structural analysis revealed that residue 501 is in proximity to the galactose binding pocket, hence confirming its functional role in galactose binding. Biodistribution analysis of AAVHSC16 indicated significantly less liver tropism in mice and non-human primates compared with other clade F members, likely associated with overall binding differences observed *in vitro*. AAVHSC16 maintained robust tropism to other key tissues in the peripheral and central nervous systems after intravenous injection, including to the brain, heart, and gastrocnemius. Importantly, AAVHSC16 did not induce elevated liver enzyme levels in non-human primates after intravenous injection at high doses. The unique glycan binding and tropism of AAVHSC16 makes this naturally occurring capsid an attractive candidate for therapies requiring less liver tropism while maintaining broad biodistribution.

INTRODUCTION

Adeno-associated viruses (AAVs) are small, non-enveloped viruses of the genus *Dependovirus* within the family *Parvoviridae*. They package a single-stranded DNA genome that is approximately 4.7 kb in length, consisting of *rep* and *cap* genes flanked by inverted terminal repeat regions (ITRs).^{1–4} AAVs are an ideal delivery system for gene therapy because they do not cause any disease, they cannot replicate without a helper virus, and the ITRs are the only regions of the viral genome necessary for packaging, thus allowing the *rep* and *cap* genes to be replaced by a therapeutic payload. This approach has led to US Food

and Drug Administration (FDA) approval of Luxturna (voretigene neparvovec-rzyl) and Zolgensma (onasemnogene abeparvovec-xioi), *in vivo* AAV gene therapies for Leber's congenital amaurosis and spinal muscular atrophy, respectively, as well as 145 AAV-based gene therapy trials being conducted.⁵

AAVs derived from human hematopoietic stem cells (AAVHSCs) are a group of 15 AAV capsids originally derived from human CD34⁺ cells of healthy hematopoietic stem cell (HSC) donors.⁶ The nucleotide sequences of AAVHSCs are most closely related to one another and to other members of clade F.^{6,7} AAVHSC capsids differ among one another by one to four capsid amino acids and they display differences in their broad biodistributions and transduction efficiencies.^{6,8–10} While some naturally occurring unique residues of the AAVHSCs have been linked to differences in tropism *in vivo*, the impact of the AAVHSC structure on transduction function, such as binding, trafficking, expression, and biodistribution, would aid in future capsid selection for therapeutic areas requiring key tissue tropism.^{11,12}

The first step in AAV infection occurs when an AAV particle binds to the cell-surface-exposed glycans prior to receptor recognition and cellular entry.¹³ Surface-membrane glycan receptors for AAVs include N-linked α 2,3-/ α 2,6-sialic acid for AAV1 and AAV6,¹⁴ heparan sulfate for AAV2,¹⁵ O-linked α 2,3-sialic acid for AAV4,¹⁶ and N-linked α 2,3-sialic acid for AAV5.¹⁷ The primary glycan receptor for AAV9, an additional clade F member, is β 1,4-galactose, which interacts with a defined pocket on the surface of the capsid.^{18–20} Glycan binding has been proposed to influence the tropism fate of an AAV, as either modulating glycan binding residues of AAV9 or altering the terminal glycans exposed on the surface of cells can affect binding and subsequent biodistribution.^{21–23}

Received 4 January 2022; accepted 27 June 2022;
<https://doi.org/10.1016/j.omtm.2022.06.013>.

Correspondence: Laura Jane Smith, PhD, 1 Patriots Park, Bedford, MA 01730, USA.

E-mail: LSmith@homologymedicines.com

Here, we sought to determine whether the naturally occurring variations in the capsid protein of the AAVHSCs can alter their glycan binding avidity *in vitro* and subsequent biodistribution *in vivo*. We report that AAVHSC16 was the only AAVHSC that did not preferentially bind surface-membrane-exposed galactose *in vitro*. Residue swapping and rational mutagenesis of AAVHSC16 were employed to understand the mechanism behind the abolished terminal galactose binding. These studies indicated residue 501, which is in close proximity to the galactose binding pocket, as the key residue leading to differences in AAVHSC16 binding *in vitro*. Biodistribution of AAVHSC16 was examined to understand how its binding contributed to *in vivo* tropism. Following a single intravenous (i.v.) injection in mice and non-human primates (NHPs), AAVHSC16 displayed reduced liver tropism compared with other clade F AAVs while maintaining strong tropism to other peripheral organs and central nervous system (CNS) tissues. Due to the lowered hepatic targeting, AAVHSC16 displayed no elevation of alanine aminotransferase (ALT) or aspartate aminotransferase (AST) in NHPs after i.v. injection. These data support the notion that single amino acid variations in the AAV capsid can contribute to dramatic differences in transduction and tissue tropism and that understanding the structure and function relationship of the naturally occurring AAVHSCs can facilitate their use in developing treatments for genetic disorders. In addition, the *in vivo* biodistribution and safety profile of AAVHSC16 has utility for diseases that require tissue transduction outside of the liver.

RESULTS

Variations in AAVHSCs lead to different glycan affinities and transduction efficiencies

AAVHSCs belong to clade F, a clade that utilizes terminally exposed galactose for cellular binding.^{7,18,20} To determine if the naturally occurring variations in the AAVHSCs affect their binding, we examined AAVHSC binding to CHO cell lines. Lec2 cells are the mutant form of their parent, Pro5, and bear a mutation that reduces translocation of CMP-sialic acid, leading to high levels of surface-exposed galactose on Lec2 cells (Figure 1A). We confirmed high and low levels of terminal galactose on Lec2 and Pro5 cells, respectively (Figures S1A and S1B).²⁴ Binding was performed on ice to prevent bound AAV internalization,²⁵ and cells were washed with ice-cold PBS to remove unbound particles prior to membrane-bound vector genome (vg) analysis. For eGFP expression, cells were resuspended in medium post washing, incubated overnight at 37°C, and assayed at 24 h by flow cytometry.

All AAVHSCs displayed improved binding and transduction of Lec2 cells with terminally exposed galactose over the parent cell line Pro5, apart from AAVHSC16 (Figures 1B and 1C). AAVHSC16 showed no difference in the number of vgs bound or eGFP expression for Lec2 and Pro5 cells, indicating that AAVHSC16 does not share the galactose-binding feature of clade F capsids. AAVHSC16 contains two unique residues compared with AAV9 and other AAVHSCs, 501I and 706C (Table 1). Both residues are surface exposed and therefore could affect glycan and/or receptor binding (Figure 2A).

While all AAVHSCs except for AAVHSC16 preferentially bound the Lec2 cells with the terminal galactose exposed, the fold change in binding was different between two groups of AAVHSCs. We found that the AAVHSCs containing the 505R residue (AAVHSC13, 15, 16, and 17) had a lower fold binding enhancement in Lec2 than the capsids with 505G (AAVHSC1, 3, 4, 6, 7, 8, and 9) (Table 1 and Figure 1D). These data indicate that galactose binding was stronger in capsids with a glycine at residue 505 and that arginine negatively affected galactose binding. Residue 505 is modeled within a loop on the surface of the capsid and part of a protrusion surrounding the galactose binding pocket (Figure 2A).⁶ The glycine-to-arginine variation increases the size and adds a positive charge at 505 and may affect galactose binding due to either steric hindrance or difference in charge (Figures 2C and 2D). AAVHSC4 displayed an increase in binding to galactose but had little eGFP expression in either cell line (Figures 1B and 1C). While the AAVHSC4 I119 variation in basic region 1 of VP1u could contribute to altered post-attachment kinetics, the unique residue 468 may also contribute to the low eGFP expression (Table 1).²⁶

Isoleucine at residue 501 on AAVHSC16 is responsible for altered terminal galactose binding

We set out to identify the contribution of each unique residue on AAVHSC16, 501I or 706C, to lowered galactose binding by mutagenizing residues 501 and 706 on AAVHSC15, an AAVHSC that also contains 505R (Tables 1 and 2) and the most studied capsid in the AAVHSC panel.^{8,27–29} As AAVHSC16 did not have robust binding to or transduction of Pro5 or Lec2 cells, we included additional mutant CHO cell lines to assess AAVHSC16 and mutagenized capsid binding to other terminal glycans (Figures 1A and 3).^{24,30,31} The lack of terminal galactose was confirmed in Lec1 and Lec8 cells (Figure S1B).

AAVHSC15 and AAVHSC16 did not display an increase in binding to Lec1 or Lec8 cells over Pro5 cells, indicating a lack of preferential binding to terminal N-acetylglucosamine (GlcNAc) or mannose (Figure 3A). The binding pattern of AAVHSC15 with the Y706C variation was similar to that of AAVHSC15, with increased binding only to galactose observed. AAVHSC15 with the F501I variation mimicked the binding pattern of AAVHSC16, with no preferential binding to any of the terminal glycans on the CHO cell lines. These data indicate that the 501I residue on AAVHSC16 is the main contributor to the abolished galactose binding and support previously published work on key residues of clade F AAVs that influence galactose binding.^{19,22,23,32} Although residues 501, 505, and 706 are located on the surface of the capsid, only 501 and 505 are in close proximity to key galactose-binding residues of AAV9 (Figure 2A).^{19,22} While the patterns between AAVHSC15 and AAVHSC15 Y706C were similar, AAVHSC15 Y706C had significantly increased galactose binding and expression in Lec2 cells compared with wild type (WT) AAVHSC15, indicating that a variation not in proximity to the galactose binding pocket altered binding and transduction of cells with high levels of surface-exposed galactose (Figures 3A, 3B, 2K, and 2L).

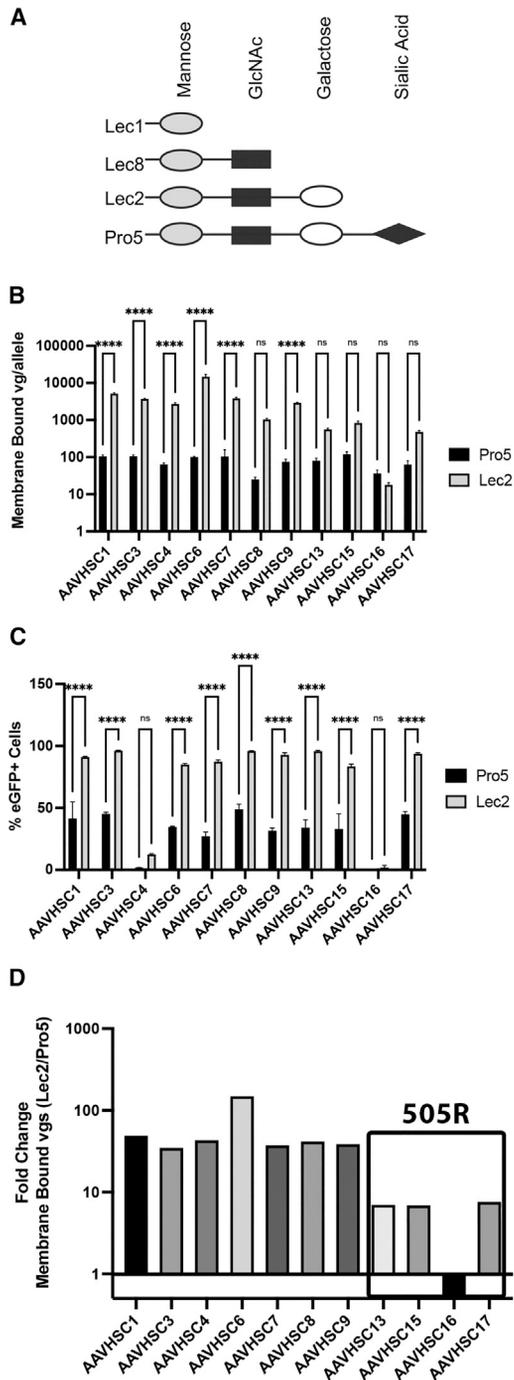


Figure 1. All AAVHSCs with the exception of AAVHSC16 have increased binding and transduction in Lec2 cells with terminally exposed galactose

Pro5 and Lec2 cells were transduced at an MOI of 2×10^4 on ice for 1 h. Surface-bound vector genomes (vgs) were determined after three washes with cold PBS. eGFP expression was determined by flow cytometry 24 h post washing after incubation at 37°C. (A) Diagram of cell-surface glycosidic chains associated with CHO cell mutants. (B) Vector genomes bound to surface of Pro5 and Lec2 cells at time of transduction. (C) eGFP expression at 24 h. (D) Fold change in cell-surface binding between Lec2 and Pro5 cells. Statistical analysis was performed using a two-way

Table 1. Amino acid variations in AAVHSCs investigated with respect to AAV9

AAVHSC serotype	Capsid region		
	VP1	VP2	VP3
AAVHSC1	A2T	–	–
AAVHSC3	–	G161D	–
AAVHSC4	F119L	–	P468S
AAVHSC6	–	–	Q590R
AAVHSC7	A68V	–	–
AAVHSC8	–	Q151R	–
AAVHSC9	–	–	C206G
AAVHSC13	–	–	G505R
AAVHSC15	–	–	G505R, T346A
AAVHSC16	–	–	F501I, G505R, Y706C
AAVHSC17	–	–	G505R

AAVHSC15 F501I expressed in all CHO cell lines better than AAVHSC16 (Figure 3B), indicating that remaining variations between these capsids, residues 346 and 706 (Table 1), contribute to additional functional differences. These variations led to detectable eGFP expression (1 to >10%) with AAVHSC15 and 501 and 706 variants in cell lines with low levels of bound vgs. The highly efficient eGFP expression of AAVHSC15 and 501 and 706 variants may be due to the T346A residue, which is not found on AAVHSC16. Residue 346 is located internally on the β strand of the capsid and may play a role in intramolecular interactions, altering expression.⁶ 346T is highly conserved among serotypes, including AAV2, AAV3b, AAV5, AAV6, AAV8, and AAV9.³³

To further confirm the contribution of 501I on AAVHSC16 to the lack of galactose binding, we generated and tested gain-of-function mutants AAVHSC16 I501F and AAVHSC16 C706Y for their binding and transduction efficiency on CHO cell lines (Figures 4A and 4B). In addition to the AAVHSC16 gain-of-function mutants, we investigated the influence on glycan binding of residue 501I on a non-505R AAVHSC (AAVHSC7; Tables 1 and 2; Figure 2E) as well as the effect of the unique AAVHSC15 residue 346A on transgene expression of AAVHSC16. AAVHSC16 Y706C did not display increased binding to surface-exposed galactose but did result in significantly increased expression in Lec2s, indicating a post-attachment role of the Y706C variation of AAVHSC16. While the AAVHSC16 I501F displayed partially restored binding to terminal galactose and expression in Lec2 cells, the number of bound vgs and eGFP expression were still significantly lower than those of WT AAVHSC15. This difference in binding and eGFP expression between WT AAVHSC15 and AAVHSC16 I501F once again highlights

analysis of variance (ANOVA). Statistical significance: ****p < 0.0001; ns, no statistical significance. All transductions were performed with biological triplicates and all molecular analysis was performed with technical triplicates. Error bars denote standard deviation.

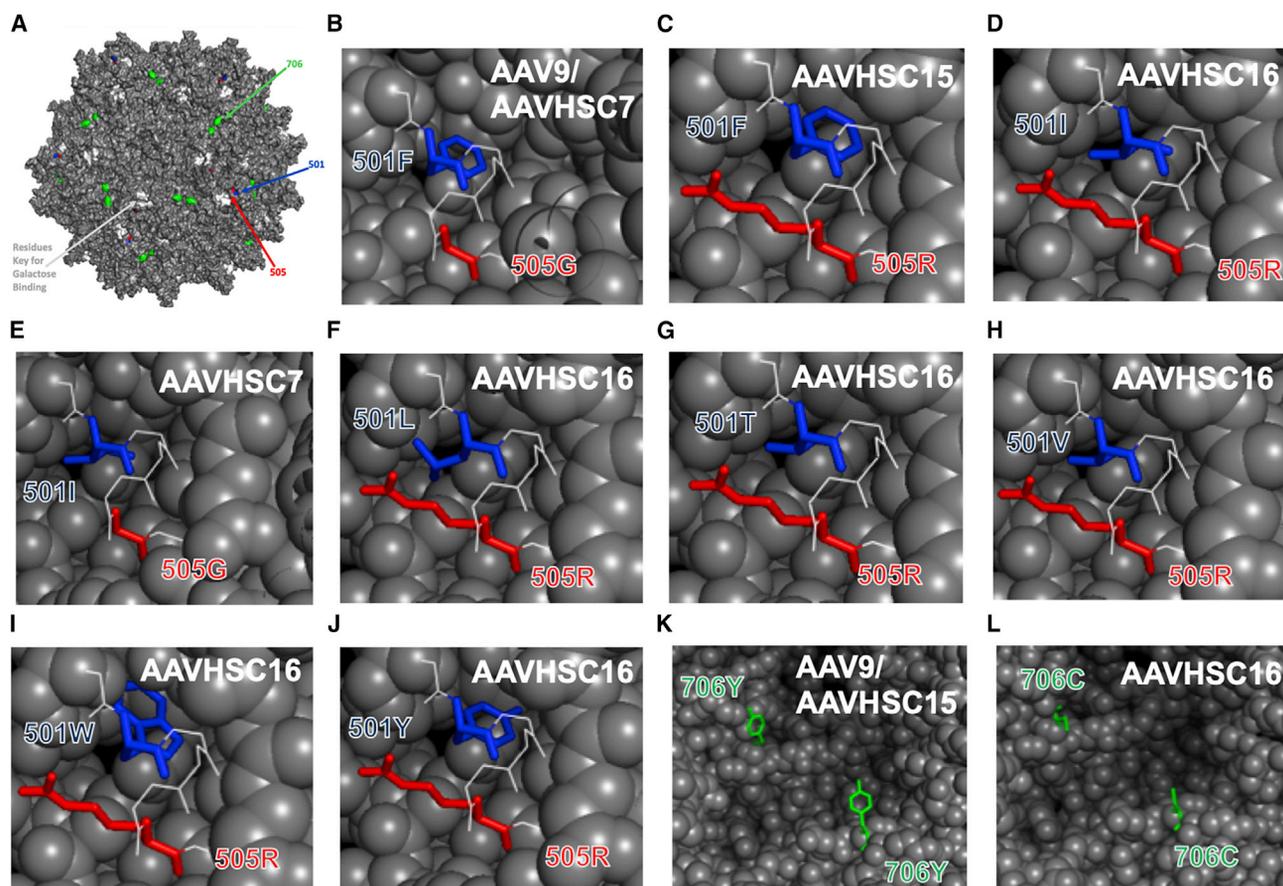


Figure 2. Structural modeling of 501, 505, and 706 residues

Structural predictions of key residue variations on AAV9 were created using PyMOL. (A) Exterior capsid surface of AAV9 with 501 (blue), 505 (red), and 706 (green) residues highlighted. In addition, for reference, seven residues previously identified as key residues for galactose binding (271, 272, 446, 470, 472, 473, and 503) are highlighted in white. For (B to J), the galactose binding pocket was zoomed in on using PyMOL. Residues 500 to 507 are highlighted in the white chain in (B to J). (B) AAV9/AAVHSC7 with predicted 501F (blue) and 505G (red) residues. (C) AAVHSC15, with predicted 501F (blue) and 505R (red) residues. (D) AAVHSC16 with predicted 501I (blue) and 505R (red). (E) AAVHSC7 F501I with predicted 501I (blue) and 505G (red) (F) AAVHSC16 I501L with predicted 501L (blue) and 505R (red). (G) AAVHSC16 I501T with predicted 501T (blue) and 505R (red). (H) AAVHSC16 I501V with predicted 501V (blue) and 505R (red). (I) AAVHSC16 I501W with predicted 501W (blue) and 505R (red). (J) AAVHSC16 I501Y with predicted 501Y (blue) and 505R (red). (K) AAV9/AAVHSC15 with predicted 706Y (green). (L) AAVHSC16 with predicted 706C (green).

the contribution of the unique residues 346 and 706 to functional differences between the capsids. The contribution of 346A, which is located internally on the capsid, to AAVHSC15 expression was highlighted in the AAVHSC16 T346A variant, where AAVHSC16 T346A showed higher eGFP expression over WT AAVHSC16 with no significant increase in membrane-bound vgs. Interestingly, AAVHSC7 F501I retained the ability to bind terminal galactose and transduce Lec2s (Figures 1B, 4A, and 4B). As AAVHSC7 has a glycine at residue 505 as opposed to a 505 arginine on AAVHSC15 and AAVHSC16 (Table 1), these data indicate that the obliteration of galactose binding due to 501I on AAVHSC16 is dependent on 505R.

The amino acid at residue 501 on AAVHSC16 contributes to overall binding function

To test the specific contribution of isoleucine at residue 501 to galactose binding, we mutagenized 501I on AAVHSC16 to leucine, threo-

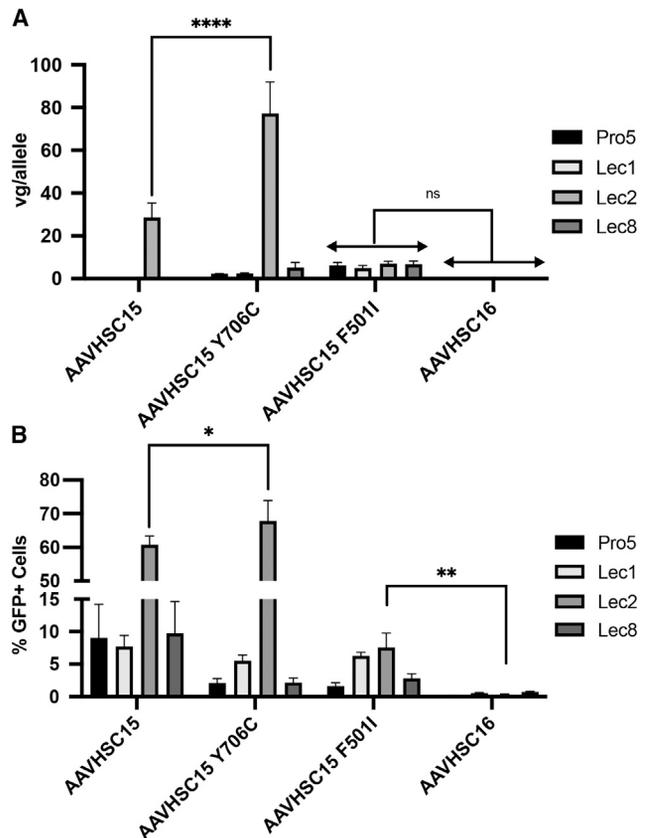
nine, valine, tryptophan, and two different tyrosine codons (Table 2). Leucine and valine are also hydrophobic amino acids, with leucine similar in size to isoleucine, but both larger than valine due to an extra carboxyl group. Leucine is as hydrophobic as phenylalanine, and both are more hydrophobic than valine and isoleucine. Threonine and valine have a similar branch structure with the exception of a hydroxyl side chain, but threonine and isoleucine are the only amino acids with two stereogenic centers.³⁴ Tryptophan and tyrosine both contain an aromatic ring structure in their side chain, making them most similar in structure to phenylalanine, with tryptophan being a binuclear ring structure and the largest of all amino acids. Modeling of the mutagenized 501 residues on AAVHSC16 revealed that the aromatic ring side chains of tryptophan and tyrosine were predicted to occupy similar space in the pocket compared with phenylalanine (Figures 2B and 2D compared with Figures 2F–2J). Binding analysis of the 501 mutagenized AAVHSC16 capsids revealed that leucine,

Table 2. Amino acid mutagenesis of AAVHSC capsid proteins

AAVHSC serotype	Amino acid variation	Reason for change
AAVHSC7	F501I	to investigate 501I on a capsid without 505R or 706C
AAVHSC15	F501I	to investigate 501I on a 505R capsid without 706C
AAVHSC15	Y706C	to investigate 706C on a 505R capsid without 501I
AAVHSC16	I501F	gain of function for 501 residue
AAVHSC16	C706Y	gain of function for 706 residue
AAVHSC16	I501L	to investigate different residues at position 501
AAVHSC16	I501T	to investigate different residues at position 501
AAVHSC16	I501V	to investigate different residues at position 501
AAVHSC16	I501W	to investigate different residues at position 501
AAVHSC16	I501YA (codon TAT)	to investigate different residues at position 501
AAVHSC16	I501YB (codon TAC)	to investigate different residues at position 501
AAVHSC16	T346A	to investigate the contribution of 346A to expression

threonine, and valine variants have preferential binding to the terminal GlcNAc of Lec8 cells, indicating a shift in glycan binding with leucine, threonine, and valine at 501 (Figure 5A). This increase in binding to the Lec8 cells also led to increased eGFP expression for the leucine, threonine, and valine AAVHSC16 mutants. Interestingly, data previously published showed that wheat germ agglutinin (WGA), which binds to GlcNAc, slightly reduced other clade F AAV binding to Pro5 cells, indicating that leucine, threonine, and valine at the 501 position may enhance clade F GlcNAc binding.¹⁸ The introduction of an aromatic ring at the 501 position with tryptophan or tyrosine allowed the AAVHSC16 I501W and I501Y variants to bind all terminal glycans along with restoration of preferential binding to terminal galactose as observed with other AAVHSCs containing a phenylalanine at 501.

eGFP expression was elevated in all CHO cell lines with 501-mutagenized AAVHSC16s compared with AAVHSC16, correlating with the improvement in overall binding (Figure 5B). Although tryptophan and tyrosine were able to bind to all the terminal glycans of all the CHO cell lines at high efficiencies, only the Lec2 cells with terminally exposed galactose displayed high eGFP expression at levels close to those of AAVHSC15, which also contains an aromatic ring at 501 with phenylalanine. No significant differences in binding or eGFP expression were observed between AAVHSC16 YA and YB, indicating that the codon usage and tRNA availability did not alter this specific function. Differences observed in binding versus expression are aligned with published data highlighting 501 on AAV9 as a key

**Figure 3. F501I amino acid variation abolishes terminal galactose binding on AAVHSC15**

Pro5, Lec1, Lec2, and Lec8 cells were transduced at an MOI of 5×10^3 on ice for 1 h. The cells were washed three times post transduction with cold PBS to determine surface-bound vgs. eGFP expression was determined at 24 h post incubation at 37°C post washing. (A) Cellular-surface-bound vgs. (B) eGFP expression at 24 h. Statistical analysis was performed using a two-way ANOVA. Statistical significance: * $p < 0.05$, ** $p < 0.01$, and **** $p < 0.0001$; ns, no statistical significance. All transductions were performed with biological triplicates and all molecular analysis was performed with technical triplicates. Error bars denote standard deviation.

residue for galactose binding and post-attachment processes up to expression.³² These data also imply that the specific amino acid and side chain at 501 influences the overall function of the virus.

AAVHSC16 has significantly reduced hepatic transduction in *in vivo* and *in vitro* models

Adachi et al. observed that modifications to residue 501 of AAV9 influenced not only glycan binding but also the *in vivo* tropism of that capsid.³² To test if differences in galactose binding observed *in vitro* affect tropism *in vivo*, we chose to compare AAVHSC15 and AAVHSC16 with two serotypes with weak and strong galactose binding, AAV2 and AAV9, respectively.^{18,20} Systemic biodistribution was assessed in albino C57 mice administered 1×10^{13} or 1×10^{14} vg/kg ssCBA.Luc vector by i.v. injection. At 6 weeks post dosing, total body bioluminescence was determined through

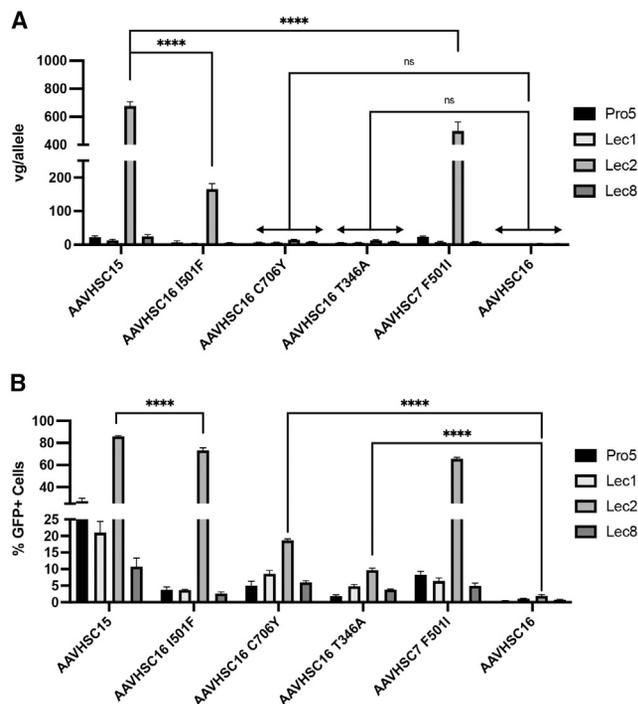


Figure 4. All naturally occurring unique residues on AAVHSC15 and AAVHSC16 contribute to the binding and expression profile in CHO

Pro5, Lec1, Lec2, and Lec8 cells were transduced at an MOI of 5×10^3 on ice for 1 h. The cells were washed three times post transduction with cold PBS to determine surface-bound vgs. eGFP expression was determined at 24 h post incubation at 37°C post washing. (A) Cellular-surface-bound vgs. (B) eGFP expression at 24 h. Statistical analysis was performed using a two-way ANOVA. Statistical significance: **** $p < 0.0001$; ns, no statistical significance. All transductions were performed with biological triplicates and all molecular analysis was performed with technical triplicates. Error bars denote standard deviation.

imaging (Figures 6A and S2A). Total body bioluminescence revealed that AAV2, which does not bind galactose, had the lowest level of systemic expression at all doses tested (Figure 6A). In contrast, AAVHSC16 and AAV9 had comparable levels of total body expression, despite their differences in galactose binding. This maintenance of total body expression for AAVHSC16 suggests that galactose binding is not sufficient to predict the overall systemic tropism *in vivo*. AAVHSC15, which binds strongly to terminal galactose, had the highest level of total body bioluminescence for both doses. To further characterize the tissue-specific tropism of each capsid, key therapeutically relevant organs were assessed for bioluminescence post harvest, as well as cellular vgs (Figures 6B–6E and S2B). The liver was the only tissue with significant differences in bioluminescence observed between the capsids tested. AAVHSC16 led to significantly lower levels of transgene expression in the liver compared with both AAV9 and AAVHSC15, with hepatic bioluminescence levels being insignificant compared with AAV2 at all doses tested. This decreased liver tropism of AAVHSC16 observed *ex vivo* was in contrast to the whole body luminescence images, which displayed higher levels of bioluminescence located in the abdomen (Figure S2A). It is possible

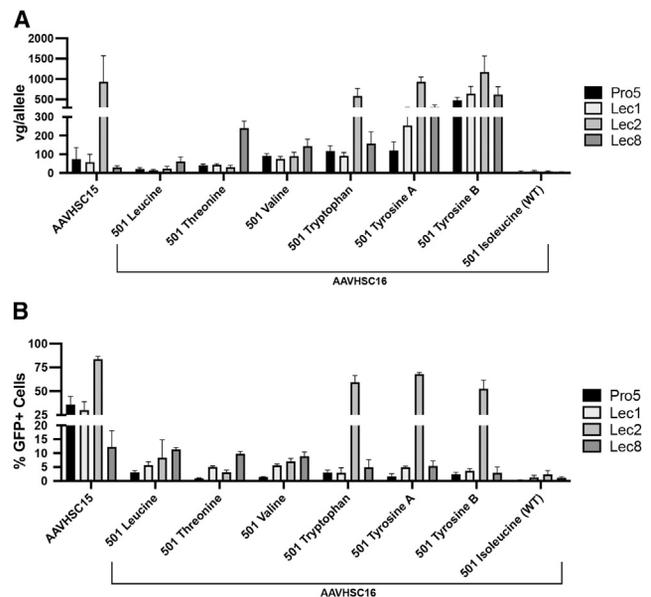


Figure 5. Amino acid at residue 501 contributes to overall binding of AAVHSC16

AAVHSC16 residue 501 was mutagenized to leucine, threonine, valine, tryptophan, and tyrosine. Pro5, Lec1, Lec2, and Lec8 cells were transduced at an MOI of 2×10^4 on ice for 1 h. The cells were washed three times with cold PBS to determine surface-bound vgs. eGFP expression was determined at 24 h post incubation at 37°C post washing. (A) Cellular-surface-bound vgs. (B) eGFP expression at 24 h. All transductions were performed with biological triplicates and all molecular analysis was performed with technical triplicates. Error bars denote standard deviation.

that other tissues or fluids in the abdomen are contributing to this localized higher transgene expression for AAVHSC16. By examining the ratio of vgs among the different organs assayed, it became evident that AAVHSC16 also had a different biodistribution pattern, with a lower ratio of vgs in the liver compared with the other clade F AAVs tested (Figures 6D and 6E). The ratio of vgs for AAV2 also revealed that it was less hepatic driven than galactose-binding AAV9 and AAVHSC15 with the lowest number of hepatic vgs. While AAVHSC16 had higher hepatic vgs than AAV2, approximately 6-fold higher for both doses, AAVHSC16 had 4- and 16-fold lower liver vgs compared with AAV9 and AAVHSC15, respectively. In contrast to the significantly lower liver tropism, AAVHSC16 maintained high levels of transduction in the CNS and periphery, with equivalent vgs compared with AAV9 and AAVHSC15 in the brain, heart, kidney, and quadriceps.

To further characterize tissue-specific expression of AAVHSC16 compared with AAV2, AAV9, and AAVHSC15, we injected CD1 mice *i.v.* with 7.5×10^{12} vg/kg of scCBA.eGFP vector with eGFP expression and determined vgs at 4 weeks post dose (Figures 7A–7C). A different strain and dose were chosen to determine if the lower hepatic transduction was influenced by either variable. Consistent with our prior study, liver vgs were lower with AAVHSC16 compared with AAV9, AAVHSC15, and, unexpectedly, AAV2. Hepatic eGFP

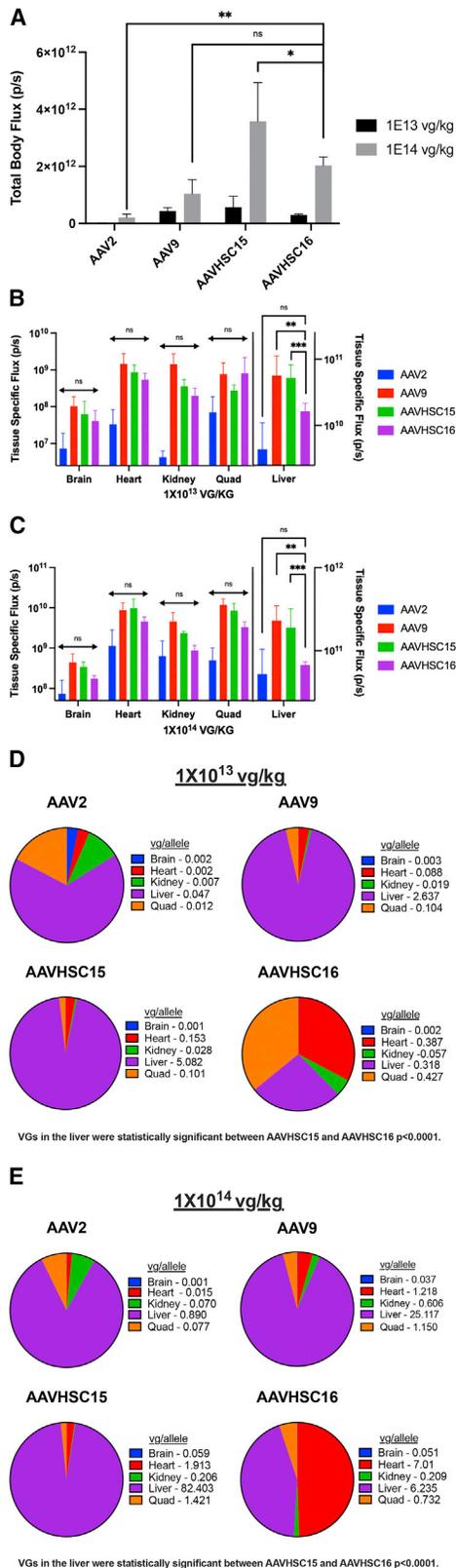


Figure 6. AAVHSC16 has widespread total body expression while maintaining reduced tropism to the liver

AAV2, AAV9, AAVHSC15, and AAVHSC16 ssCBA.Luc vectors were injected i.v. into albino C57 mice at either a 1×10^{13} or a 1×10^{14} vg/kg dose. Whole-body bioluminescence, tissue-specific luminescence, and tissue-specific vgs were determined 6 weeks post injection. (A) Whole-body bioluminescence imaging was performed to determine total body transgene expression. (B and C) Tissue-specific bioluminescence was assessed at time of harvest to determine tissue-specific transgene expression. (D and E) Specified tissue vgs were determined by digital droplet PCR (ddPCR). Statistical analysis was determined by two-way ANOVA. Statistical significance: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$; ns, no statistical significance; $n = 3-5$ animals per group. Error bars denote standard deviation.

expression among the different capsids correlated with hepatic vgs (Figures 7A–7C). AAVHSC16 vgs were between 2- and 3-fold lower in the brain, spinal cord, and gastrocnemius compared with AAVHSC15, yet were comparable to or slightly higher than those of AAV9 in these tissues. AAV9 displayed the most robust eGFP expression in the peripheral tissues (heart, liver, gastrocnemius) but had drastically lower eGFP expression in the CNS tissue (brain and spinal cord) compared with AAVHSC15 and AAVHSC16. Although AAVHSC16 had higher vgs in the heart compared with AAVHSC15 and AAV9, expression appeared lower with AAVHSC16. Differences in sampling of the heart for vg and expression analysis or capsid-specific expression mechanisms of action may explain this discrepancy. The vg and eGFP expression analysis of AAVHSC16 demonstrates the maintenance of a broad biodistribution of this AAVHSC with lowered liver tropism.

To determine if the unique tropism of AAVHSC16 is species specific, we investigated the biodistribution in NHPs. Biodistribution of AAVHSC15 and AAVHSC16 was examined in two cynomolgus macaques (*Macaca fascicularis*) per capsid at 2 weeks post i.v. delivery of ssCBA.eGFP vector at a 7×10^{13} vg/kg dose. eGFP expression was substantially greater in the liver of NHPs with AAVHSC15 compared with those that received AAVHSC16 (Figure 8A). NHPs administered AAVHSC16 had approximately 3 log fewer hepatic vgs than NHPs administered AAVHSC15 (Figure 8B). Despite variability between individual NHPs, AAVHSC16 maintained high and equivalent levels of transduction (vgs) and expression in the heart, gastrocnemius, and brain compared with AAVHSC15, indicating that high levels of non-hepatic and broad systemic tropism *in vivo* were not species constrained. The reduced hepatic tropism of AAVHSC16 also did not lead to increased levels of liver transaminases (Figures 8C–8E). NHPs administered AAVHSC16 ssCBA.eGFP at 7×10^{13} vg/kg or a vector expressing human phenylalanine hydroxylase driven by a liver-specific promoter (ssDNG.PAH) at 1×10^{14} vg/kg did not display elevated ALT or AST levels at any time point post dose compared with baseline levels or vehicle-treated control animals. As shown in Figure 8F, comparing AAVHSC16 liver transduction and ALT/AST levels with AAV9 and other AAVHSC serotypes further suggests that the reduced risk of hepatotoxicity of AAVHSC16 is associated with lower hepatic tropism.

In an effort to translate the lowered liver tropism of AAVHSC16 observed in the mice and NHPs to human cells, we transduced

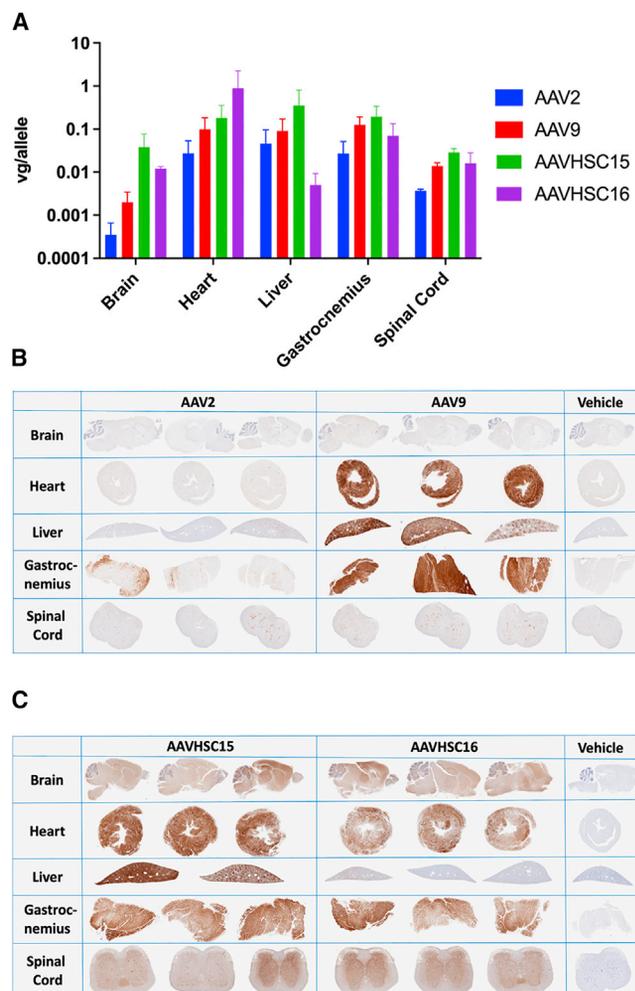


Figure 7. AAVHSC16 transduction and expression are comparable to those of AAVHSC15 in peripheral and CNS tissues while maintaining reduced tropism to the liver

AAV2, AAV9, AAVHSC15, and AAVHSC16 scCBA.eGFP vectors were injected i.v. into CD1 mice at a 7.5×10^{12} vg/kg dose. Tissue-specific vgs and eGFP expression were determined 4 weeks post injection. (A) Vector genomes in specified tissues. (B) Tissue-specific eGFP expression assessed by anti-eGFP immunohistochemistry for AAV2 and AAV9. (C) Tissue-specific eGFP expression assessed by anti-eGFP immunohistochemistry for AAVHSC15 and AAVHSC16. Liver from one animal dosed with AAVHSC15 was not included due to issues with anti-eGFP immunohistochemistry. Significant differences in transduction by vgs were observed only in the heart at this dose; statistical analysis was performed by two-way ANOVA; $n = 3$ animals per group. Error bars denote standard deviation.

primary hepatocytes with several WT AAVHSC and AAVHSC variant scCBA.eGFP vectors at three MOIs (Figure 9). At 4 days post transduction, AAVHSC15, AAV9, AAVHSC7, and AAVHSC7 F501I led to greater eGFP expression than AAVHSC16 and all AAVHSC16 variants at all MOIs tested. As observed *in vitro* with the CHO cells, the internally located 346A residue change alone on AAVHSC16 increased expression in primary human hepatocytes. The AAVHSC15 variants containing either AAVHSC16 501I or

706C revealed that the lower hepatocyte expression from AAVHSC16 is primarily due to the 706C variation and not the galactose-altering 501I variation. The AAVHSC16 gain-of-function mutants also revealed that the variation at residue 706 was the main contributor to lowered hepatocyte expression, as AAVHSC16 C706Y had more expression restored than AAVHSC16 I501F. While all 501 changes (L, T, V, W, and Y) partially restored expression of AAVHSC16 in the hepatocytes, the largest and most intrusive 501 change with tryptophan restored the least amount of hepatic eGFP expression. The lack of eGFP expression observed with AAVHSC16 in primary human hepatocytes along with the AAVHSC16 501 mutants reinforced that the significantly reduced hepatic tropism can be translated among species and that all unique amino acid variations contribute to the unique binding and biodistribution of AAVHSC16.

DISCUSSION

AAV gene therapy is an evolving field that has the potential to provide a cure for a range of genetic disorders. AAVHSCs are desirable therapeutic vectors because the panel of 15 distinct AAVs are naturally derived, are amenable to both gene transfer and gene editing modalities, can cross the blood-brain barrier (BBB), and have low pre-existing neutralizing antibody prevalence in humans.^{6,8,10,27} Here, we sought to better understand the structure-function relationship of the naturally occurring AAVHSC variations for cell-surface attachment and *in vivo* tropism.

We show that 10 AAVHSCs tested use galactose as their primary glycan receptor. AAVHSC16, on the other hand, does not display an increase in binding to cells with high levels of surface-bound terminal galactose *in vitro*. The unique AAVHSC16 residues were examined through mutagenesis to determine which naturally occurring residue was contributing to the lack of galactose binding. We have demonstrated, in concordance with literature from other groups, that a single amino acid difference can affect galactose binding, particularly residues in close proximity to the defined galactose binding pocket.^{19,21,32} Mutagenesis experiments revealed that residue 501I on AAVHSC16 is responsible for the obliteration of terminal galactose binding, but only in combination with the 505R neighboring residue, as AAVHSC7 F501I retained binding to terminal galactose on Lec2 cells. This contribution of residue 501 on clade F AAVs to terminal glycan binding is in line with Adachi et al., as their AAV9 F501A mutant did not bind terminal galactose.³² The AAVHSC16 residue 501 is an isoleucine, while other clade F AAVs and AAVHSCs have phenylalanine at residue 501. Interestingly, a phenylalanine at the 501 position is highly conserved among many other naturally occurring AAV serotypes (AAV1, AAV2, AAV3, AAV6, AAV7, AAV8, and AAV9).³² Our AAVHSC16 501 tryptophan and tyrosine data support the need for an aromatic ringed structure at the 501 position on capsids with a 505R residue for galactose binding and are supported by the fact that interactions between aromatic residues are widespread for carbohydrate recognition.³⁵ Results from mutagenesis studies also highlight the importance and difference between naturally occurring variations versus selective mutagenesis of capsids for structure and function studies.

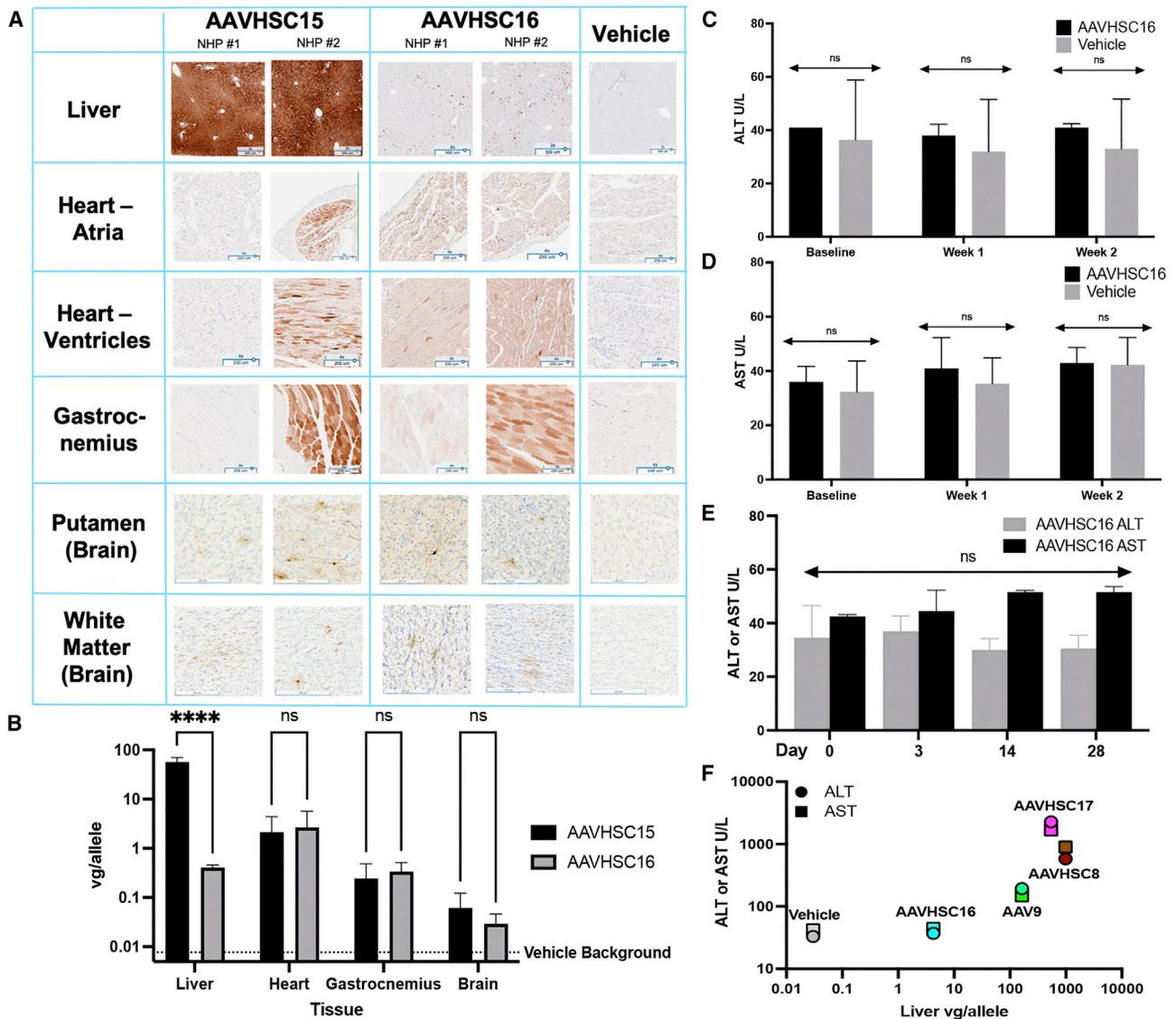


Figure 8. AAVHSC16 has reduced tropism in the liver in NHPs *in vivo* and does not lead to increased ALT and AST levels after high-dose i.v. injection

(A) Liver, heart, gastrocnemius, and brain biodistribution of AAVHSC15 and AAVHSC16 was assessed in cynomolgus macaques (*M. fascicularis*) following systemic delivery at a dose of 7×10^{13} vg/kg of scCBA.eGFP vector at 2 weeks post dose using anti-eGFP immunohistochemistry. Liver images are shown at a $2\times$ magnification with scale bars representing $500 \mu\text{m}$. All remaining peripheral tissue images are shown at a $4\times$ magnification with scale bars representing $250 \mu\text{m}$. Brain images have scale bars representing 200 to $250 \mu\text{m}$. (B) Vector genome measurement from specified tissues. (C) ALT levels in NHPs dosed at 7×10^{13} vg/kg with AAVHSC16 scCBA.eGFP vector or vehicle were assessed prior to dosing followed by 1 and 2 weeks post dosing. (D) AST levels in NHPs dosed at 7×10^{13} vg/kg with AAVHSC16 scCBA.eGFP vector or vehicle were assessed prior to dosing followed by 1 and 2 weeks post dosing. (E) ALT and AST levels in NHPs dosed at 1×10^{14} vg/kg with AAVHSC16 ssDNG.PAH vector were assessed prior to dosing followed by 3, 14, and 28 days post dosing. (F) Average ALT and AST levels versus average liver vgs in NHPs dosed at 1×10^{14} vg/kg with AAV9, AAVHSC8, AAVHSC16, and AAVHSC17 ssDNG.PAH vector were assessed at 3 days post dosing. $n = 2$ animals per group. Statistical analysis was determined by two-way ANOVA. Statistical significance: **** $p < 0.0001$ ns, no statistical significance; $n = 2$ animals per group. Error bars denote standard deviation.

The research presented here supports the idea that differences in cellular binding affect tropism *in vitro* and *in vivo*, consistent with previously published data.^{21–23,32} AAVHSC16 does not bind galactose *in vitro* and displays lower transduction of the liver in both mice and NHPs *in vivo* and human hepatocytes *in vitro* compared with other clade F AAVs tested. A similar relationship between lack of galactose

binding and lower hepatic transduction was also observed in other publications with a W503R AAV9 variant.^{22,23} The original proposed mechanism for the relationship between galactose binding and hepatic tropism was attributed to changes in distribution half-life or clearance, as increasing glycan affinity reduced transduction of non-hepatic tissues such as skeletal and heart muscle.²² Contradicting that proposed

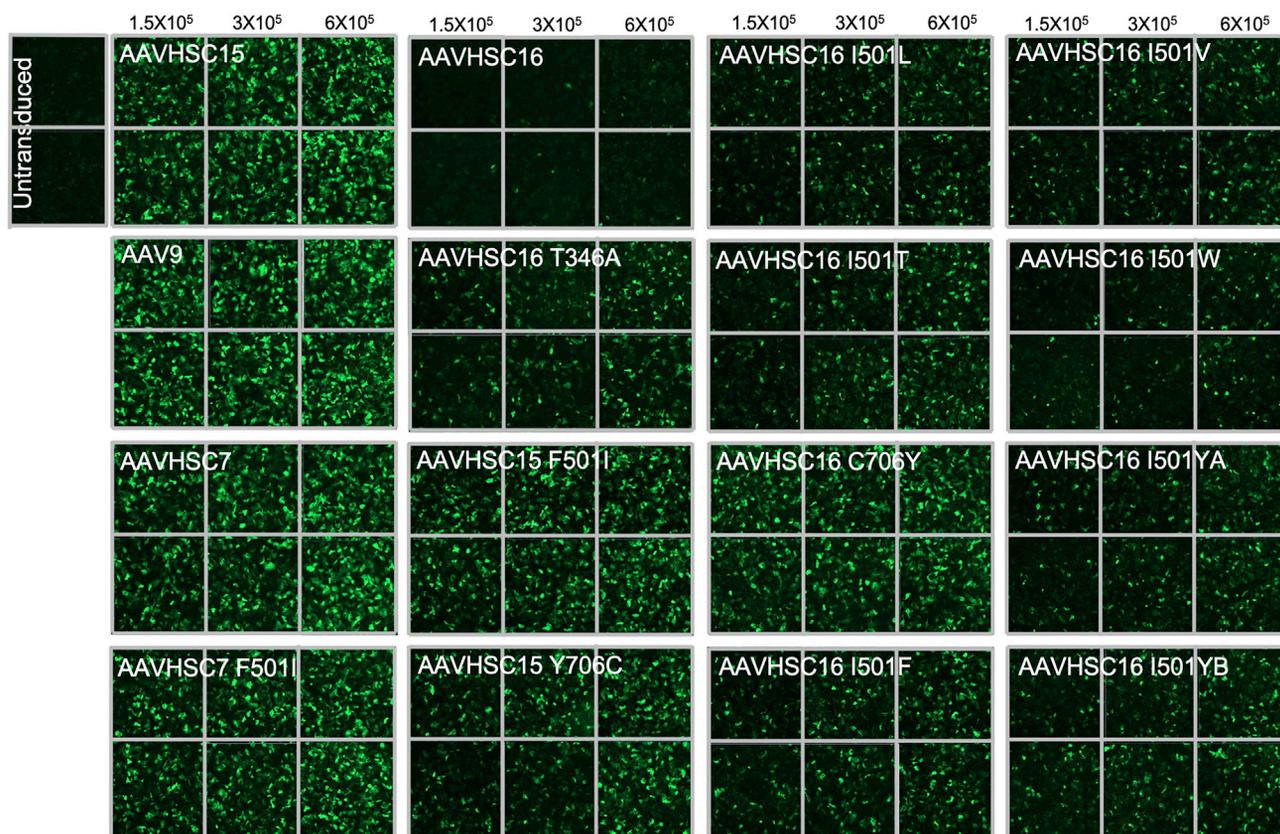


Figure 9. AAVHSC16 has reduced transduction efficiency in human hepatocytes *in vitro* due to the combination of naturally occurring amino acid variants
 Frozen primary human hepatocytes were plated at 2×10^5 cells/well in collagen-coated 48-well plates. Cells were transduced using WT AAV or AAVHSCs and variant AAVHSC scCBA.eGFP vectors at MOIs of 1.5×10^5 , 3×10^5 , and 6×10^5 . Fluorescence microscopy was used to detect eGFP expression at day 4 post transduction. For each MOI, transductions were performed with biological duplicates.

mechanism, further studies introducing key galactose binding residues of AAV9 onto hepatotropic or non-hepatotropic variants did not alter their *in vivo* biodistribution, indicating that galactose binding was not the driving factor to tropism.²¹ This is not surprising, as the prevalence of glycans with terminal galactose would be low *in vivo* due to natural sialylation.³⁶ Indeed, our mutagenesis further revealed that residue 706C of AAVHSC16 appeared to be the main influencer of tropism to human hepatocytes *in vitro*, with a modest contribution from the 50II residue. These data also align with previously published literature that demonstrated that changing the tyrosine to alanine at residue 706 of AAV9 affected liver tropism but not galactose binding, most likely due to alteration of post-attachment processes.³² To further support these structure and function findings that liver tropism for clade F viruses is multifaceted, Li et al. found that transferring AAV9 amino acids 705 to 735 to AAV2 did not transfer AAV9's high liver tropism to AAV2.³⁷ This indicates that variations in clade F AAV capsids can affect both glycan binding and tropism, but these two characteristics are not necessarily linked and are driven by multiple regions of the capsid.

While AAVHSC16 had lower transduction in the liver compared with other clade F members, AAV9 and AAVHSC15, it maintained a high

level of transduction of other key therapeutic tissues, including in the periphery and CNS. Previous publications have shown that engraftment of key galactose binding residues from AAV9 onto AAV2 did not increase brain-specific transduction, indicating that altering galactose binding does not affect BBB crossing or CNS transduction.²¹ Indeed, mutations of two other key residues for galactose binding, H527Y and R533S on AAV9, reduced peripheral tissue transduction but maintained CNS transduction.³⁸ As we have not yet established which glycan or receptor AAVHSC16 is preferentially binding to in these studies, the maintenance of high tropism to non-hepatic tissues, including the CNS, indicates AAVHSC16 is binding another common and highly expressed glycan(s). These data displaying maintenance of high-level tropism to non-hepatic peripheral and CNS tissue with lower liver transduction elevate AAVHSC16 as an attractive capsid to use for gene therapy or gene editing applications that can address diseases that require little to no hepatic contributions. In addition, the significant reduction in liver transduction and lack of induction of liver transaminases with AAVHSC16 in NHPs could increase the safety profile of the capsid, as elevated levels of transduction in the liver have been linked with liver toxicity in non-clinical studies.³⁹

Table 3. Titers of AAVHSC and AAVHSC variants

AAVHSC serotype	Titer (vg/mL)
AAVHSC7	1.25×10^{13}
AAVHSC7 F501I	4.21×10^{13}
AAVHSC15	3.64×10^{13a}
AAVHSC15 F501I	4.56×10^{13}
AAVHSC15 Y706C	1.75×10^{13}
AAVHSC16	3.33×10^{13}
AAVHSC16 I501F	2.83×10^{13}
AAVHSC16 Y706C	4.56×10^{13}
AAVHSC16 T346A	3.17×10^{13}
AAVHSC16 I501L	2.00×10^{13}
AAVHSC16 I501T	1.30×10^{13}
AAVHSC16 I501V	1.20×10^{13}
AAVHSC16 I501W	1.32×10^{13}
AAVHSC16 I501Ya	9.16×10^{12}
AAVHSC16 I501Ya	1.44×10^{13}

^aAverage of three separate productions.

These findings highlight the importance of evolutionarily directed natural variations in human-derived AAVHSCs and the mechanistic advantages that they naturally carry. These structure-function relationships can increase translational success by guiding capsid selection of AAVHSCs to best address tissue specificity and the safety requirements for developing an AAV therapy. Further studies will gain additional understanding of how the natural variations in the AAVHSC capsids influence their broad distribution and differences in transduction as well as additional characteristics, such as the immune profile across multiple species *in vivo*. These studies highlight how variations in a natural capsid isolated from humans, AAVHSC16, contribute to its desirable biodistribution and safety profile for the development of a safe and effective AAV-mediated gene therapy.

MATERIALS AND METHODS

Culturing of cell lines

CHO glycan mutants, Pro5, Lec2, Lec1, and Lec8 cells, were cultured in minimum essential medium α (MEM- α) with nucleosides (Thermo Fisher Scientific, Waltham, MA) and supplemented with 10% fetal bovine serum (FBS; Thermo Fisher Scientific, Waltham, MA) and 1% penicillin-streptomycin (Thermo Fisher Scientific, Waltham, MA).

rAAV production, purification, and titration

Recombinant AAV and AAVHSC vectors used in *in vitro* and *in vivo* studies were manufactured by Homology Medicines. Briefly, triple transfection of HEK293 cells was performed with a plasmid containing the payload of interest containing AAV2 ITRs, a second plasmid containing the capsid sequence and AAV2 Rep gene, and a third

plasmid containing adenovirus helper genes. At 72 h post transfection, cells were separated from supernatant by centrifugation and lysed for 1 h in a buffer containing Tris-HCl, sodium chloride (NaCl), Triton X-100, magnesium chloride (MgCl₂), and benzonase. Cell debris was clarified by centrifugation and purified using an AAV9 affinity resin. Vectors were enriched for full capsids by cesium ultracentrifugation and buffer exchanged and formulated in Dulbecco's phosphate-buffered saline (DPBS)-based buffer. Vectors were titrated by qPCR using primers targeting the SV40 region. All vectors were analyzed for VP1, 2 and 3 ratios by silver and Coomassie blue-stained sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), endotoxin levels (<10 EU/mL), and capsid content by enzyme-linked immunoassay (ELISA). All AAVHSC and altered AAVHSC capsids were able to yield acceptable titers when produced using the same process (Table 3).

Mutagenesis of AAVHSCs

Mutagenic oligonucleotides were synthesized to induce single-nucleotide changes (Integrated DNA Technologies, Coralville, IA). Mutations were introduced onto the WT AAVHSC7, AAVHSC15, and AAVHSC16 capsids using the QuickChange Lightning site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) according to the manufacturer's specifications. The mutagenized plasmid was then transformed into DH5 α Turbo Competent *Escherichia coli* (New England Biolabs, Ipswich, MA) and colonies were screened by DNA sequencing (Eton Biosciences, Charlestown, MA).

CHO cell surface binding assay

Cells were counted and resuspended in growth medium at a density of $1-2 \times 10^6$ cells/mL, plated in pre-chilled 96-well tissue culture plates, and then incubated on ice for 15 min prior to transduction. Cells were then transduced on ice by adding the scCBA.eGFP vectors directly to the cell culture medium at the indicated MOI. One hour following transduction, the cells were washed three times in 4°C PBS. After the final wash, half of the cells were resuspended in 100 μ L PBS for subsequent DNA extraction using the in-plate cell lysis and DNA extraction method. The remaining half of the cells were resuspended in medium and cultured at 37°C for 24 h.

Transduction assay in primary human hepatocytes

Frozen primary human hepatocytes were plated at 2×10^5 cells/well in a collagen-coated 48-well plate. At 6 h following plating, the culture medium was replaced with fresh medium and cells were transduced with scCBA.eGFP vectors at increasing MOIs of 1.5×10^5 , 3×10^5 , and 6×10^5 . After 48 h post transduction, the culture medium was replaced. Fluorescence microscopy was used to detect eGFP expression at day 4 post transduction.

Flow cytometric analysis

In vitro eGFP expression was measured at 24 h post-AAV scCBA.eGFP transduction by flow cytometric analysis on either the SONY SH800 (Sony Biotechnology) or the CytoFLEX S (Beckman Coulter, Brea, CA). Briefly, cells were washed once with PBS and

adherent CHO cells were dissociated using TrypLE reagent (Thermo Fisher Scientific, Waltham, MA). Cells were then stained to determine viability using the LIVE/DEAD fixable violet dead cell stain (Thermo Fisher Scientific, Waltham, MA) for 10 min at room temperature in PBS. Cells were washed and then resuspended in fluorescence-activated cell sorting (FACS) buffer (2% FBS in PBS) or cell staining buffer (BioLegend, San Diego, CA). A minimum of 1×10^4 cells were collected for each sample. The total number of cells and level of terminally exposed cell-membrane galactose were determined using lectin peanut agglutinin (PNA) (no. L32460; Thermo Fisher Scientific, Waltham, MA) from *Arachis hypogaea* (peanut), conjugated with Alexa Fluor 647. Cells were stained using a 1:1,000 dilution, incubated for 1 h, and then washed twice with PBS prior to flow cytometry analysis.

DNA isolation from cells

Direct quantification of DNA from lysed cells was performed through an in-plate lysis protocol. Briefly, to each 100- μ L cell suspension sample, 85 μ L of extraction solution (0.25% deoxycholate, 0.45% Tween/HEPES solution [25% Tween 20, 0.5 M HEPES], 1 mM Tris-HCl, 1 mM EDTA, 0.1% SDS, 0.3 mg/mL proteinase K) (Thermo Fisher Scientific, Waltham, MA) was added. The samples were then mixed by pipetting and sealed with an adhesive plate film. The plate was then transferred to a thermocycler and incubated at 37°C for 1 h, 55°C for 2 h, and 95°C for 30 min (to inactivate the proteinase K). For subsequent qPCR analysis, 10 μ L of digestion per sample was diluted in 90 μ L of Ultrapure water (Thermo Fisher Scientific, Waltham, MA).

Quantitation of vector genomes in transduced cells

The ratio of AAV vgs per allele was determined by qPCR of DNA isolated from cells. Analysis was performed on the QuantStudio 6 real-time PCR System (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA). Vector genome detection was done using TaqMan Universal PCR Master Mix (Thermo Fisher Scientific, Waltham, MA) and the following primers and probes (Integrated DNA Technologies, Coralville, IA):

eGFP, forward, 5'-CTG CTG CCC GAC AAC CA-3'; reverse, 5'-GAC CAT GTG ATC GCG CTT CT-3'; probe, 5'-TAC CTG AGC ACC CAG TCC GCC CT-3'

CHO actin β , forward, 5'-CCA TGT ACG TAG CCA TTC AGG-3'; reverse, 5'-CAT GAG GGA GAG CGT AGC C-3'; probe, 5'-TGT CCC TGT ATG CCT CTG GTC GTA-3'

Animal procedures for mice

Housing and all experimental procedures were in accordance with the Institutional Animal Care and Use Committee (IACUC) at Homology Medicines (Bedford, MA). B6(Cg)-Tyr^{c-2J}/J (B6 albino) mice (obtained from The Jackson Laboratory, Bar Harbor, ME, USA) and CD1 mice (obtained from Charles River Laboratories, Wilmington, MA) were maintained in the animal facility at Homology Medicines under standard laboratory conditions (temperature 20°C–24°C, relative hu-

midity 50%–60%, 12 h light/12 h dark cycle). Mice were supplied with standard chow (PicoLab 5058, LabDiet, St. Louis, MO) and sterile water *ad libitum*. Male B6 albino mice at 7 weeks of age and male and female CD1 mice at 4 weeks of age were administered vectors or vehicle by i.v. injection using a 1-mL syringe with 28G needle. Imaging of dosed B6 albino mice was performed at 6 weeks post injection. Xenolight D-luciferin-K+ salt bioluminescent substrate (no. 122799; Perkin Elmer, Waltham, MA) was administered via intraperitoneal (i.p.) injection (15 mg/mL in DPBS –Ca²⁺, –Mg²⁺, 150 mg/kg of body weight). *Ex vivo* bioluminescence imaging of tissues was performed immediately *ex vivo* after carbon dioxide (CO₂) euthanasia of mice injected i.p. with the luciferin substrate. Luciferase transgene expression was imaged using the IVIS Lumina LT series III pre-clinical *in vivo* imaging system (PerkinElmer, Waltham, MA) and whole body and tissue-specific luminescence was quantitated using Living Image (PerkinElmer, Waltham, MA) software. Tissues from dosed CD1 mice were analyzed by eGFP immunohistochemistry (Premier Laboratory, Longmont, CO).

Animal procedures for NHPs

Housing and all procedures were in accordance with the IACUC at the Mannheimer Foundation (Homestead, FL). All breeding, housing, and procedures were performed on cynomolgus macaques (*M. fascicularis*) at the Mannheimer Foundation under protocol 2016-05 as previously described.⁸ The Mannheimer Foundation IACUC specifically approved these studies. Sera from all animals were pre-screened for anti-AAVHSC neutralizing antibodies using the Huh-7 cell-based assay (Horae Gene Therapy Center, University of Massachusetts, Worcester, MA) over a range of serum dilutions from 1/10 to 1/1,250. Only antibody-negative animals were used in this work. Veterinary staff at the Mannheimer Foundation anesthetized each subject (young males, 4–5 months of age) with an intramuscular (i.m.) injection of ketamine (10 mg/kg). Each animal was fitted with a saphenous vein catheter for i.v. injection of AAVHSCs. Blood samples were collected from the femoral vein for assessment of clinical chemistry, complete blood counts (CBCs), and neutralizing antibodies at baseline. AAVHSCs were infused at 7–8 mL/kg through a saphenous vein catheter over 1.0 min using all-plastic syringes. Each animal was allowed to recover in a warmed incubator following injection and, when fully awake, was returned to the cage with its mother. At 2 weeks post dosing, all subjects were anesthetized with ketamine, heparinized (5,000 units, given 5–10 min prior to sacrifice), and euthanized with an overdose of i.v. sodium pentobarbital (80–100 mg/kg). ALT and AST levels in the serum of dosed NHPs were determined at VRL Laboratories (San Antonio, TX) using a clinical CBC and chemistry panel.

DNA isolation from murine tissues

DNA was extracted from tissues by the Maxwell RSC (Promega, Madison, WI) using the Maxwell RSC DNA FFPE kit (Promega, Madison, WI) and by the QIAamp Fast DNA tissue kit (Qiagen, Hilden, Germany) both as described by the manufacturer with the following modifications: approximately 50 mg of tissue was homogenized in a Pre-cellys soft tissue homogenizing CK14 2-mL tube with the

Precellys Evolution homogenizer (Bertin Instruments, Montigny-le-Bretonneux, France) for 30 s at 5,000 rpm prior to column extraction. DNA was eluted in 50 μ L ultrapure water.

DNA isolation from fixed NHP tissue

Approximately 30- μ g portions of fixed NHP tissue were isolated using a biopsy punch. Samples were homogenized in a Precellys soft tissue homogenizing CK14 2-mL tube with the Precellys Evolution homogenizer (Bertin Instruments, Montigny-le-Bretonneux, France) for 30 s at 5,000 rpm. Homogenized tissue was then digested for 24 h using proteinase K (Thermo Fisher Scientific, Waltham, MA), and DNA was isolated using the RecoverAll total nucleic acid isolation kit (Thermo Fisher Scientific, Waltham, MA). DNA was eluted in 50 μ L ultrapure water (New England Biolabs, Ipswich, MA).

Quantitation of vector genome biodistribution in animal model tissues

Vector genome copies per allele were determined by digital droplet PCR (ddPCR) using ddPCR Supermix for probes (no dUTP) (Bio-Rad Laboratories, Hercules, CA). Detection of targets in 10 ng of sample input was accomplished using the following primers and probes (Integrated DNA Technologies, Coralville, IA):

eGFP #1, forward, 5'-CTG CTG CCC GAC AAC CA-3'; reverse, 5'-GAC CAT GTG ATC GCG CTT CT-3'; probe, 5'-TAC CTG AGC ACC CAG TCC GCC CT-3'

eGFP #2, forward, 5'-GAA CCG CAT CGA GCT GAA-3'; reverse, 5'-TGC TTG TCG GCC ATG ATA TAG-3'; probe, 5'-ATC GAC TTC AAG GAG GAC GGC AAC-3'

Luciferase, forward, 5'-CTC AAA GTA TTC AGC ATA GGT GAT GTC-3'; reverse, 5'-AAC TGC ACA AGG CCA TGA AGA-3'; probe, 5'-TTG CCT TCA CTG ATG CTC ACA TTG AGG T-3'

Mouse ApoB, forward, 5'-CGT GGG CTC CAG CAT TCT A-3'; reverse, 5'-TCA CCA GTC ATT TCT GCC TTT G-3'; probe, 5'-CCT TGA GCA GTG CCC GAC CAT TG-3'

NHP ApoB, forward, 5'-TGA AGG TGG AGG ACA TTC CTC TA-3'; reverse, 5'-CTG GAA TTG CGA TTT CTG GTA A-3'; probe 5'-CGA GAA TCA CCC TGC CAG ACT TCC AT-3'

Post-droplet generation, 96-well plates were sealed with an adhesive plate film and transferred to a thermocycler for cycling and then to the QX200 droplet reader for analysis. Data were analyzed using QuantaSoft Analysis Pro software (Bio-Rad Laboratories, Hercules, CA).

Histology and microscopy

Paraffin-embedded tissues were sectioned at 4–5 μ m and were processed for eGFP immunohistochemistry (IHC) at Premier Laboratory (Boulder, CO) as previously described.⁸ eGFP-positive cells were stained by IHC with a rabbit anti-eGFP polyclonal antibody (no.

ab290; Abcam, Cambridge, UK). Endogenous peroxidase was inhibited by incubation in 3% H₂O₂. Samples were pre-treated with proteinase K, followed by a serum-free protein block (epitope retrieval). Primary antibody incubation was for 30 min at room temperature, followed by an EnVision+ rabbit HRP detection system (Agilent, Santa Clara, CA). eGFP was visualized with an EnVision+ rabbit HRP kit (Agilent, Santa Clara, CA) using diaminobenzidine to stain for eGFP and hematoxylin as counterstain. Slides were viewed with a Nikon Eclipse 80i microscope with an attached Nikon DXM 1200C digital camera (Nikon Instruments, Melville, NY). IHC images were acquired using Aperio-ImageScope (Leica Biosystems, Wetzlar, Germany) and Huron Viewer (Huron Digital Pathology, Ontario, CA) software.

Statistical analysis

All statistical analysis was performed using GraphPad Prism software using either a one-way or a two-way ANOVA.

Data availability

The authors confirm that the data supporting the findings of these studies are available within the article and its [supplemental information](#). Raw data were generated at Homology Medicines, and data supporting the findings of these studies are available from the corresponding author (L.J.S.) upon request.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtm.2022.06.013>.

ACKNOWLEDGMENTS

This work was funded by Homology Medicines, Inc. All authors were employed at Homology Medicines during their contributions to the manuscript. We would like to thank Dr. MiJeong Kim and Daisy Huynh for their contribution of liver enzyme levels for the ssDNG.PAH NHP experiment. We would like to thank Jason Lindemeyer for the generation of high-resolution images.

AUTHOR CONTRIBUTIONS

Conceptualization, L.J.S., S.S., C.M.B., L.O.B., J.L.E., O.F., and A.B.S.; methodology, L.J.S., L.A.S., S.S., C.M.B., L.B., M.S., M.K., K.D., L.O.B., and J.L.E.; validation, L.A.S.; formal analysis, L.J.S., L.A.S., S.S., L.B., M.S., M.K., K.D., L.O.B., and J.L.E.; investigation, L.A.S., L.V.L., L.B., M.S., K.D., D.F., A.H., J.L., J.I.R., A.V.S., and A.T.; writing – original draft, L.J.S. and S.S.; writing – review and editing, L.J.S.; visualization, L.J.S. and S.S.; supervision, L.J.S., S.S., C.M.B., L.O.B., J.L.E., N.A., O.F., and A.B.S.

DECLARATION OF INTERESTS

All authors have received salary with employee stock options and/or restricted stock units options during their employment at Homology Medicines, Inc. L.J.S. is an inventor on patents associated with AAVHSCs and their uses.

REFERENCES

- Hoggan, M.D., Blacklow, N.R., and Rowe, W.P. (1966). Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. *Proc. Natl. Acad. Sci. USA* 55, 1467–1474. <https://doi.org/10.1073/pnas.55.6.1467>.
- Rutledge, E.A., Halbert, C.L., and Russell, D.W. (1998 Jan). Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2. *J. Virol.* 72, 309–319. <https://doi.org/10.1128/JVI.72.1.309-319.1998>.
- Atchison, R.W., Casto, B.C., and Hammon, W.M. (1965). Adenovirus-associated defective virus particles. *Science* 149, 754–756. <https://doi.org/10.1126/science.149.3685.754>.
- Wu, Z., Asokan, A., and Samulski, R.J. (2006). Adeno-associated virus serotypes: vector toolkit for human gene therapy. *Mol. Ther.* 14, 316–327. <https://doi.org/10.1016/j.ymthe.2006.05.009>.
- Wang, D., Tai, P.W.L., and Gao, G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.* 18, 358–378. <https://doi.org/10.1038/s41573-019-0012-9>.
- Smith, L.J., Ul-Hasan, T., Carvaines, S.K., Van Vliet, K., Yang, E., Wong, K.K., Jr., Agbandje-McKenna, M., and Chatterjee, S. (2014). Gene transfer properties and structural modeling of human stem cell-derived AAV. *Mol. Ther.* 22, 1625–1634. <https://doi.org/10.1038/mt.2014.107>. Epub 2014 Jun 13.
- Gao, G., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X., and Wilson, J.M. (2004). Clades of Adeno-associated viruses are widely disseminated in human tissues. *J. Virol.* 78, 6381–6388. <https://doi.org/10.1128/JVI.78.12.6381-6388.2004>.
- Ellsworth, J.L., Gingras, J., Smith, L.J., Rubin, H., Seabrook, T.A., Patel, K., Zapata, N., Olivieri, K., O'Callaghan, M., Chlipala, E., et al. (2019). Clade F AAVHSCs cross the blood brain barrier and transduce the central nervous system in addition to peripheral tissues following intravenous administration in nonhuman primates. *PLoS One* 14, e0225582. <https://doi.org/10.1371/journal.pone.0225582>.
- Smith, L., Kauss, A., Wong, K.K., and Chatterjee, S. (2010). Enhanced hepatic transgene expression following intravenous delivery of novel stem cell-derived recombinant AAV vectors. *Mol. Ther. AAV Vectors: Vector Traffic.* 18, S1–S2.
- Smith, L.J., Wright, J., Clark, G., Ul-Hasan, T., Jin, X., Fong, A., Chandra, M., St Martin, T., Rubin, H., Knowlton, D., et al. (2018). Stem cell-derived clade F AAVs mediate high-efficiency homologous recombination-based genome editing. *Proc. Natl. Acad. Sci. USA* 115, E7379–E7388. Erratum in: *Proc Natl Acad Sci USA*. 2019 Jan 2;116:337. <https://doi.org/10.1073/pnas.1802343115>.
- Smith, L., Van Vliet, K., Miller, E., Kauss, A., Ul-Hasan, T., Wong, K.K., Agbandje-McKenna, M., and Chatterjee, S. (2011). Functional mapping of tissue tropism of naturally occurring adeno-associated virus isolates from human hematopoietic stem cells. *Mol. Ther.* 19, S127–S128.
- Smith, L., Ul-Hasan, T., Johnson, B., Bhargava, R., Chandra, M., Wong, K.K., and Chatterjee, S. (2014). Receptor binding and post-entry processing of stem cell-derived AAV isolates in CD34+ cells. *Mol. Ther. AAV Vector Biol.* 22, S27–S28.
- Huang, L.Y., Halder, S., and Agbandje-McKenna, M. (2014). Parvovirus glycan interactions. *Curr. Opin. Virol.* 7, 108–118. <https://doi.org/10.1016/j.coviro.2014.05.007>.
- Wu, Z., Miller, E., Agbandje-McKenna, M., and Samulski, R.J. (2006). Alpha2,3 and alpha2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *J. Virol.* 80, 9093–9103. <https://doi.org/10.1128/JVI.00895-06>.
- Summerford, C., and Samulski, R.J. (1998). Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J. Virol.* 72, 1438–1445. <https://doi.org/10.1128/JVI.72.2.1438-1445.1998>.
- Kaludov, N., Brown, K.E., Walters, R.W., Zabner, J., and Chiorini, J.A. (2001). Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J. Virol.* 75, 6884–6893. <https://doi.org/10.1128/JVI.75.15.6884-6893.2001>.
- Walters, R.W., Yi, S.M., Keshavjee, S., Brown, K.E., Welsh, M.J., Chiorini, J.A., and Zabner, J. (2001). Binding of adeno-associated virus type 5 to 2,3-linked sialic acid is required for gene transfer. *J. Biol. Chem.* 276, 20610–20616. Epub 2001 Mar 21. PMID: 11262413. <https://doi.org/10.1074/jbc.M101559200>.
- Bell, C.L., Vandenberghe, L.H., Bell, P., Limberis, M.P., Gao, G.P., Van Vliet, K., Agbandje-McKenna, M., and Wilson, J.M. (2011). The AAV9 receptor and its modification to improve in vivo lung gene transfer in mice. *J. Clin. Invest.* 121, 2427–2435. <https://doi.org/10.1172/JCI57367>. Epub 2011 May 16.
- Bell, C.L., Gurda, B.L., Van Vliet, K., Agbandje-McKenna, M., and Wilson, J.M. (2012). Identification of the galactose binding domain of the adeno-associated virus serotype 9 capsid. *J. Virol.* 86, 7326–7333. <https://doi.org/10.1128/JVI.00448-12>.
- Shen, S., Bryant, K.D., Brown, S.M., Randell, S.H., and Asokan, A. (2011). Terminal N-linked galactose is the primary receptor for adeno-associated virus 9. *J. Biol. Chem.* 286, 13532–13540. <https://doi.org/10.1074/jbc.M110.210922>.
- Shen, S., Horowitz, E.D., Troupes, A.N., Brown, S.M., Pulicherla, N., Samulski, R.J., Agbandje-McKenna, M., and Asokan, A. (2013). Engraftment of a galactose receptor footprint onto adeno-associated viral capsids improves transduction efficiency. *J. Biol. Chem.* 288, 28814–28823. <https://doi.org/10.1074/jbc.M113.482380>.
- Shen, S., Bryant, K.D., Sun, J., Brown, S.M., Troupes, A., Pulicherla, N., and Asokan, A. (2012). Glycan binding avidity determines the systemic fate of adeno-associated virus type 9. *J. Virol.* 86, 10408–10417. <https://doi.org/10.1128/JVI.01155-12>.
- Pulicherla, N., Shen, S., Yadav, S., Debbink, K., Govindasamy, L., Agbandje-McKenna, M., and Asokan, A. (2011). Engineering liver-detargeted AAV9 vectors for cardiac and musculoskeletal gene transfer. *Mol. Ther.* 19, 1070–1078. <https://doi.org/10.1038/mt.2011.22>.
- Deutscher, S.L., Nuwayhid, N., Stanley, P., Briles, E.I., and Hirschberg, C.B. (1984). Translocation across Golgi vesicle membranes: a CHO glycosylation mutant deficient in CMP-sialic acid transport. *Cell* 39, 295–299. [https://doi.org/10.1016/0092-8674\(84\)90007-2](https://doi.org/10.1016/0092-8674(84)90007-2).
- Bartlett, J.S., Wilcher, R., and Samulski, R.J. (2000). Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J. Virol.* 74, 2777–2785. <https://doi.org/10.1128/jvi.74.6.2777-2785.2000>.
- Popa-Wagner, R., Porwal, M., Kann, M., Reuss, M., Weimer, M., Florin, L., and Kleinschmidt, J.A. (2012). Impact of VP1-specific protein sequence motifs on adeno-associated virus type 2 intracellular trafficking and nuclear entry. *J. Virol.* 86, 9163–9174. <https://doi.org/10.1128/JVI.00282-12>.
- Ellsworth, J.L., O'Callaghan, M., Rubin, H., and Seymour, A. (2018). Low seroprevalence of neutralizing antibodies targeting two clade F AAV in humans. *Hum. Gene Ther. Clin. Dev.* 29, 60–67. <https://doi.org/10.1089/humc.2017.239>. PMID: .
- Chen, H.M., Resendes, R., Ghodssi, A., Sookiasian, D., Tian, M., Dollive, S., Adamson-Small, L., Avila, N., Tazearslan, C., Thompson, J.F., et al. (2020). Molecular characterization of precise in vivo targeted gene integration in human cells using AAVHSC15. *PLoS One* 15, e0233373. <https://doi.org/10.1371/journal.pone.0233373>.
- Ahmed, S.S., Rubin, H., Wang, M., Faulkner, D., Sengooba, A., Dollive, S.N., Avila, N., Ellsworth, J.L., Lamppu, D., Lobikin, M., et al. (2020). Sustained correction of a murine model of phenylketonuria following a single intravenous administration of AAVHSC15-PAH. *Mol. Ther. Methods Clin. Dev.* 17, 568–580. <https://doi.org/10.1016/j.omtm.2020.03.009>.
- Deutscher, S.L., and Hirschberg, C.B. (1986). Mechanism of galactosylation in the Golgi apparatus. A Chinese hamster ovary cell mutant deficient in translocation of UDP-galactose across Golgi vesicle membranes. *J. Biol. Chem.* 261, 96–100.
- Stanley, P., and Chaney, W. (1985). Control of carbohydrate processing: the lec1A CHO mutation results in partial loss of N-acetylglucosaminyltransferase I activity. *Mol. Cell Biol.* 5, 1204–1211. <https://doi.org/10.1128/mcb.5.6.1204-1211.1985>.
- Adachi, K., Enoki, T., Kawano, Y., Veraz, M., and Nakai, H. (2014). Drawing a high-resolution functional map of adeno-associated virus capsid by massively parallel sequencing. *Nat. Commun.* 5, 3075. <https://doi.org/10.1038/ncomms4075>.
- Meyer, N.L., Hu, G., Davulcu, O., Xie, Q., Noble, A.J., Yoshioka, C., Gingerich, D.S., Trzynka, A., David, L., Stagg, S.M., and Chapman, M.S. (2019). Structure of the gene therapy vector, adeno-associated virus with its cell receptor. *AAVR. Elife.* 8, e44707. <https://doi.org/10.7554/eLife.44707>.
- IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN) (1984). Nomenclature and symbolism for amino acids and peptides. Recommendations 1983. *Eur. J. Biochem.* 138, 9–37. Erratum in: *Eur J Biochem.* 1993 Apr 1;213:2. <https://doi.org/10.1111/j.1432-1033.1984.tb07877.x>.

35. Spiwok, V. (2017). CH/ π interactions in carbohydrate recognition. *Molecules* 22, 1038. <https://doi.org/10.3390/molecules22071038>.
36. Conte, F., van Buuringen, N., Voermans, N.C., and Lefeber, D.J. (2021). Galactose in human metabolism, glycosylation and congenital metabolic diseases: time for a closer look. *Biochim. Biophys. Acta Gen. Subj.* 1865, 129898. <https://doi.org/10.1016/j.bbagen.2021.129898>.
37. Li, W., Asokan, A., Wu, Z., Van Dyke, T., DiPrimio, N., Johnson, J.S., Govindaswamy, L., Agbandje-McKenna, M., Leichtle, S., Redmond, D.E., Jr., et al. (2008). Engineering and selection of shuffled AAV genomes: a new strategy for producing targeted biological nanoparticles. *Mol. Ther.* 16, 1252–1260. <https://doi.org/10.1038/mt.2008.100>.
38. Wang, D., Li, S., Gessler, D.J., Xie, J., Zhong, L., Li, J., Tran, K., Van Vliet, K., Ren, L., Su, Q., et al. (2018). A rationally engineered capsid variant of AAV9 for systemic CNS-directed and peripheral tissue-detargeted gene delivery in Neonates. *Mol. Ther. Methods Clin. Dev.* 9, 234–246. <https://doi.org/10.1016/j.omtm.2018.03.004>.
39. Hinderer, C., Katz, N., Buza, E.L., Dyer, C., Goode, T., Bell, P., Richman, L.K., and Wilson, J.M. (2018). 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. <https://doi.org/10.1089/hum.2018.015>.