An engineered AAV targeting integrin alpha V beta 6 presents improved myotropism across species

Ai Vu Hong, Laurence Suel, Eva Petat, Auriane Dubois, Pierre-Romain Le Brun, Nicolas Guerchet, Philippe Veron, Jérôme Poupiot, Isabelle Richard

nature communications

Article

https://doi.org/10.1038/s41467-024-52002-4

An engineered AAV targeting integrin alpha V beta 6 presents improved myotropism across species

Received: 25 October 2023

Accepted: 22 August 2024

Published online: 11 September 2024

Check for updates

Ai Vu Hong[1,2,4,5] ✉, Laurence Suel[1,2,4], Eva Petat[1,2], Auriane Dubois[1,2], Pierre-Romain Le Brun[1,2], Nicolas Guerchet[1,2], Philippe Veron[1,2], Jérôme Poupiot[1,2] & Isabelle Richard[1,2,3,5] ✉

Current adeno-associated virus (AAV) gene therapy using nature-derived AAVs is limited by non-optimal tissue targeting. In the treatment of muscular diseases (MD), high doses are often required but can lead to severe adverse effects. Here, we rationally design an AAV capsid that specifically targets skeletal muscle to lower treatment doses. We computationally integrate binding motifs of human integrin alphaV beta6, a skeletal muscle receptor, into a liver-detargeting capsid. Designed AAVs show higher productivity and superior muscle transduction compared to their parent. One variant, LICA1, demonstrates comparable muscle transduction to other myotropic AAVs with reduced liver targeting. LICA1's myotropic properties are observed across species, including non-human primate. Consequently, LICA1, but not AAV9, effectively delivers therapeutic transgenes and improved muscle functionality in two mouse MD models (male mice) at a low dose (5E12 vg/kg). These results underline the potential of our design method for AAV engineering and LICA1 variant for MD gene therapy.

Over 50 years since their discovery, adeno-associated viruses (AAVs) have shown great promise as an effective viral vector for gene delivery and gene therapy, leading to recent approval of therapeutic products[1,2]. Due to unmet medical needs and natural AAV tropism, many AAV-based gene therapies focus on treating muscle diseases (MD)[3]. Systemic treatment in such diseases aims to primarily target skeletal muscle, which accounts for > 40% of body mass, and therefore often requires very high doses (≥1E14 vg/kg) to achieve meaningful therapeutic efficacy[3–6]. In addition, most recombinant AAVs built on natural-occurring variants lack specificity and often accumulate in the liver, with the concomitant risk of hepatotoxicity[7]. Other key challenges of rAAV use persist, including manufacturing, immunological barriers, and associated toxicity[1,2,8,9].

AAV is a small non-pathogenic single-stranded DNA parvovirus. Multiple open reading frames (ORFs) were identified in its genome, including *Rep*, *Cap*, *AAP* and *MAAP*[1,10]. The single *Cap* ORF expresses three capsid proteins - virion protein 1 (VP1), VP2 and VP3, which assemble into an icosahedral 60-mer capsid. Structurally, the VP3 monomer core contains a highly conserved eight-stranded β-barrel motif[11]. Inserted between the β-strands, nine surface-exposed variable regions (VR1-9) result in local topological differences between serotypes and dictate virus-host interaction. Consequently, genetically modifying VRs can drastically change the AAV, transduction, antigenic profile, and fitness[10,12,13]. VR4 and VR8, that cluster together spatially, forming the most prominent protrusion at the 3-fold axis, have been widely subjected to modifications, notably by inserting short peptides into the loop apices[14]. This resulted in some highly efficient capsid variants for transducing a variety of cell types and tissues[1,12]. Among these, remarkably, AAVMYOs[15,16] and MYOAAVs[17] transduce skeletal muscles, deliver therapeutic transgenes efficiently, and were shown to

[1]Genethon, 1 bis rue de l'internationale, Evry, France. [2]INTEGRARE research unit UMR_S951 (INSERM, Université Paris-Saclay, Univ Evry), Evry, France. [3]Atamyo Therapeutics, 1 bis rue de l'internationale, Evry, France. [4]These authors contributed equally: Ai Vu Hong, Laurence Suel. [5]These authors jointly supervised this work: Ai Vu Hong, Isabelle Richard. ✉e-mail: avuhong@genethon.fr; richard@genethon.fr

correct dystrophic phenotypes in MD mouse models at relatively low doses (2E12–1E13 vg/kg).

Importantly, the myotropic AAVs[15–17] identified by muscle-directed high-throughput screening (HTS) were shown to share an Arg-Gly-Asp (RGD) motif, presumably targeting the integrin complex[17–20]. Integrins are a group of heterodimeric proteins composed of an α- and a β subunit that serve various cellular functions, including cell adhesion, cell migration, and cell signaling[21]. As adhesion molecules, integrins also mediate cell-pathogen interactions, and are therefore exploited by many viruses, including natural AAV, to infect cells[22–24]. Indeed, many of these viruses use an RGD motif on their viral envelope glycoproteins or capsids for cell attachment, endocytosis, entry, and endosomal escape[18,22,25]. The discovery that RGD-dependent integrin-targeting AAV variants can acquire myotropism therefore represents a potential candidate approach for a rational design to target skeletal muscle.

This study introduces a computational method for a rational AAV design targeting skeletal muscle, which resulted in a myotropic vector for MD gene therapy. First, the human skeletal muscle-enriched integrin complex alpha V beta 6 (αVβ6) was selected as the target receptor. Inspired by one-sided protein design[26,27], we computationally designed a previously developed liver-detargeting hybrid capsid between AAV9 and AAVrh74 (Cap9rh74) as an αVβ6 binder. The VR4 loop was completely modified, in which new sequences were iteratively selected to simultaneously optimize for free energy, while hosting αVβ6-binding RGDLXXL/I motifs. All designed AAVs were well-produced, at higher titers than their parent. The designed AAVs were confirmed to require the binding to αVβ6 from multiple species for cellular transduction. The most promising variant, renamed LICA1, was selected for further analysis and showed superior transduction in human differentiated myotubes and strong myotropism in several mouse models and non-human primate (NHP). We evaluated this variant by delivering therapeutic transgenes in two MD mouse models at a very low dose of 5E12 vg/kg, in comparison to AAV9. In both cases, LICA1 presents higher efficacy than AAV9 in correcting dystrophic phenotypes, global transcriptomic changes and restoring muscle function, thanks to improved transduction and transgene expression in skeletal muscles. Collectively, the study provides a proof-of-concept for a rational AAV design pipeline leveraging protein design tools, which resulted in a myotropic AAV with high potential for gene therapy for muscle diseases.

Results

Selection of the cellular receptor for rational design

Several myotrophic AAVs have recently been developed, notably, the insertion into the AAV9 VR-VIII loop of P1 peptide (RGDLGLS)[15,16], and a series of RGD-containing sequences identified by directed evolution[17]. Importantly, these modified capsids shared a common RGD motif, which suggested their affinity to integrin (ITG), cell-surface heterocomplexes that interact with the extracellular matrix[28]. Using publicly available datasets, we aimed to select relevant integrin subunits for a subsequent rational AAV design targeting skeletal muscle.

Chemello and colleagues previously performed single-nucleus RNA sequencing, comparing gene expression of all cell types in the skeletal muscle of wild-type (WT) and Duchenne muscular dystrophy mouse models (D51)[29]. We extracted RNA levels of all integrin alpha and beta genes from these data (Fig. S1A). Among all subunits, only the α- subunits *Itgav*, *Itga7* and the β-subunits *Itgb6*, *Itgb1*, and *Itgb5* show relatively high expression in the myogenic nuclei. Of interest is the fact that the expression level of *Itgb6* is highly enriched in myonuclei, and significantly upregulated in the dystrophic condition, whereas *Itgb1* and *Itgb5* expression are ubiquitous in all cell types, and significantly lower than the *Itgb6* level in all myonuclei. Among the two expressed α-subunits, only *Itgav* was known to associate with *Itgb6* to form αvβ6 heterocomplexes—a member of the RGD-binding integrin family[30]. Furthermore, the cell type-enriched transcriptome profiles from bulk RNA sequencing data of multiple human tissues[31] confirmed high expression and enrichment of both *Itgav* and *Itgb6* in skeletal and cardiac muscles, and low expression of *Itgb6* in the liver and spleen, two preferred targets of natural AAV (Fig. S1B, C, GTEx V8, dbGaP Accession phs000424.v8.p2). We therefore hypothesize that AAV transduction in skeletal muscle can be improved by rationally designing an AAV capsid that specifically binds to αvβ6.

Rational design of a hybrid capsid, Cap9rh74, with a high affinity to the αVβ6 complex

As we aim to specifically target the skeletal muscle, we selected a hybrid capsid that we previously developed and that has a liver-detargeting property as the parental capsid in our design (Patent Number: EP18305399.0). This hybrid capsid of AAV9 and AAV.rh74 (AAV9rh74) was constructed by replacing the AAV9 sequence of VR4 to VR8 with that of AAV-rh74. The hybrid capsid showed similar infectivity in skeletal and cardiac muscles but was strongly de-targeted from the liver. The latter property is of particular interest in skeletal muscle gene transfer since the majority of administrated viral vector will not accumulate in the liver, as is the case for natural AAVs[32,33].

After selection of the cellular receptor of interest and capsid backbone, AAV capsids were computationally engineered (Fig. 1A). First, the 3D structure of the parental capsid, of with structure was unknown, was modeled using AlphaFold2[34,35]. The structural prediction of the Cap9rh74 aa 219–737 monomer performed using AlphaFold2 was at a high level of confidence, with predicted local distance difference test (lDDT-Cα), a per-residue measure of local confidence, of 97.04 and low predicted aligned error (PEA) of 4.32 (Fig. S1D). This structure is thus suitable for the next steps in the design.

Second, we extracted the 3D structure or sequences of binding motifs of the human integrin complex from PDB. Importantly, αVβ6 was previously shown to bind with high affinity to the RGDLXXL/I motif found in the human TGF-β1 and TGF-β3 prodomains[36,37]. Binding peptides with eight amino acid residues, aa214–221 in TGF-β1 (PDB: 5ffo) and aa240–247 in TGF-β3 (PDB: 4um9), were isolated from the corresponding crystal structures before grafting into the Cap9rh74 VR4 loop. Both motifs bind to αVβ6 dimer at a very similar position (Fig. S1E).

Third, the defined binding motifs were then grafted into the VR4 loop (residues 453–459) of the capsid protein based on the RosettaRemodel protocol[38]. In the grafting-remodel process, many rounds of backbone optimization and sequence design iteratively search for low-energy sequence–structure pairs (Fig. 1B). The lowest-energy designs in grafting experiments of each TGF-β motif showed convergence in both structure and sequence (Fig. 1C, D and Fig. S1F, G). The new VR4 loops include the binding peptide and two flanking 2-amino acid linkers and retain the LXXL/I motif as an α-helix, which is important to bind in the β6 subunit's pocket[37].

Retrospective docking simulations of the two AAV_ITGs with the best scores, namely Cap9rh74_5ffo and Cap9rh74_4um9, on the αVβ6 complex showed highly similar binding positions of the new VR4 loop to its corresponding inserted motifs (Fig. 1E, F). This suggests that the new capsids can bind to αVβ6 thanks to VR4-included RGDLXXL/I motif. Sequences with the best scores, which reflect the thermodynamic stability of one static protein conformation[39], were subjected to experimental validation.

All designed AAV_ITGs showed higher productivity and enhanced cellular transduction via αVβ6 binding

The two AAVs with the best design were then tested for productivity and the effectiveness of using αVβ6 as a cellular receptor. They were produced by tri-transfection with pITR-CMV-GFP-Luciferase as the expression cassette. Thanks to energy optimization, all the designed AAV-ITG variants significantly increase their titers compared to their

https://doi.org/10.1038/s41467-024-52002-4

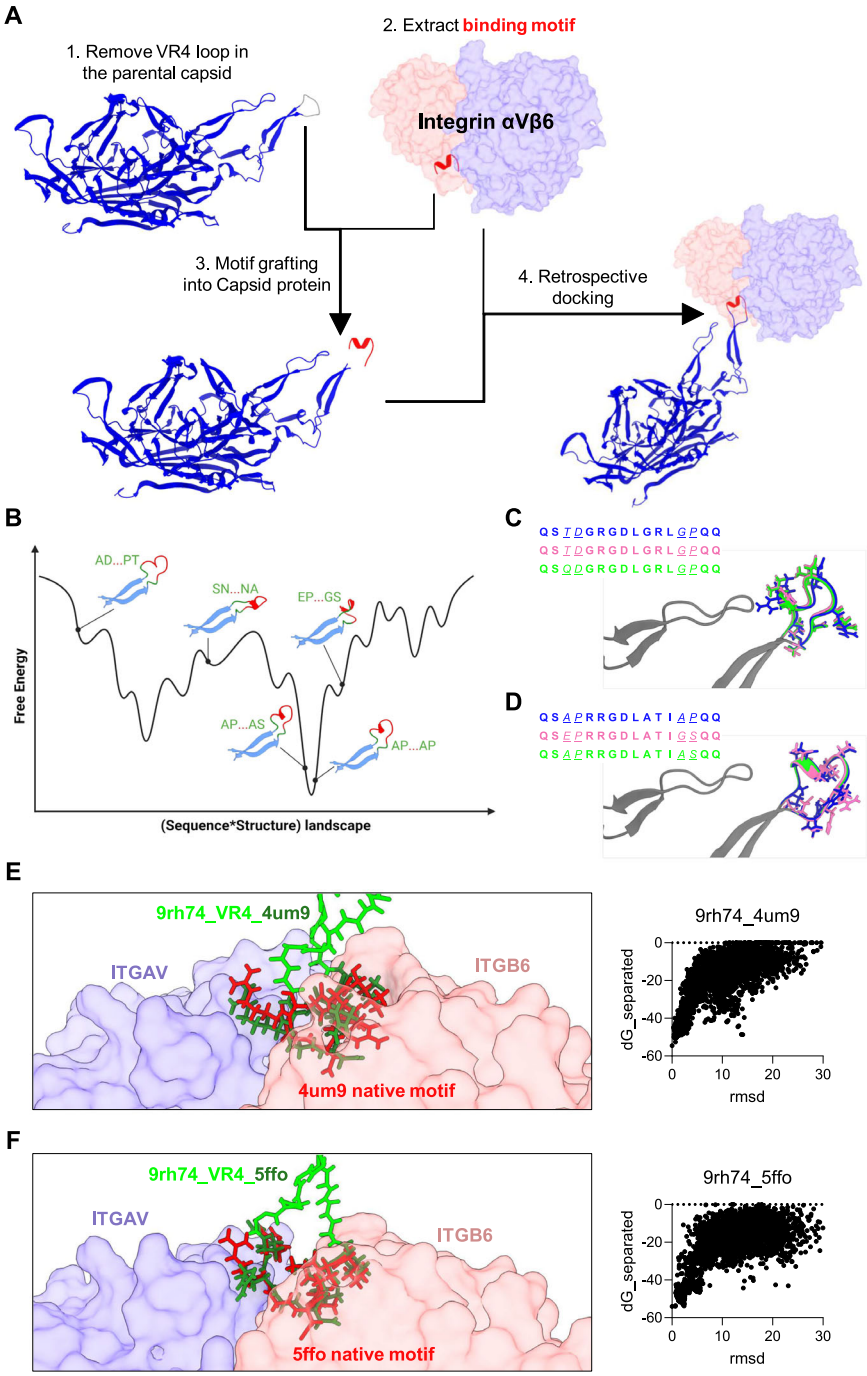


Fig. 1 | Computational rational AAV capsid design to bind to αVβ6 integrin. **A** Overview of the design pipeline, including three steps: 1. Capsid 3D structures were obtained either from the PDB database or predicted by AlphaFold2. 2. The capsid VR4 loop was completely replaced by integrating the binding motif, which was extracted from receptor's natural binder, using RosettaRemodel protocol. 3. Top scored designs from the previous grafting step were docked onto the intended receptor in silico to verify the binding potential of the designed capsid. **B** An illustration of the sampling for low-energy sequence-structure pairs during motif-grafting process. Capsid VR4 after removing the loop was colored in blue, extracted binding motif was colored in red. The sampled linkers and sequences (Fig. S1F) were labeled in green. Created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en). **C**, **D** The three lowest energy designs after grafting TGFβ3 (**C**) and TGFβ1 (**D**) into the capsid VR4. All top designs showed convergence in structures and sequences, suggesting sampling approached the global optimum. **E**, **F** Retrospective docking of motif-grafted capsids (**E** Cap9rh74_4um9) and (**F** Cap9rh74_5ffo) onto the αVβ6 structure. The left panels are illustrations of the structures with the lowest energy at the interface of capsid and integrin proteins (dG_separated: difference in free energy of two proteins). Both two newly designed VR4s (colored in green) were predicted to bind to the αVβ6 complex at very similar positions to natural binding motifs (colored in red). The right panels are scatter plots of dG_separated energy versus root-mean-square deviation (RMSD) from the lowest energy structure of all sampled docking positions.

parental hybrid capsid, to levels similar to those for AAV9 (Fig. 2A, Fig. S2A, and Supplementary Note). In addition, all modified AAV-ITG variants retain proportions of VP1, VP2, VP3 capsid proteins with a similar ratio of AAV9 (Fig. 2B). This suggests that the designed sequences result in more stable AAV capsid complexes thanks to their estimated low energy structure, and therefore better production efficacy.

Next, we examined whether these AAV-ITGs can effectively use αVβ6 as a cellular receptor upon transduction. First, a HEK293 cell line (293_αVβ6) constitutively overexpressing both integrin subunits, αV and β6, was created using the PiggyBac system (Fig. S2B, C). The designed AAVs were then tested for their transduction in this cell line. As expected, transduction of AAV_ITGs in 293_αVβ6 cells, as defined by vector copy numbers (VCN), was higher than for AAV9 and AAV9rh74 (Fig. 2C). Both AAV_ITGs dramatically improved the luciferase activity ($FC_{9rh74\_4um9/AAV9} = 60.50$, $FC_{9rh74\_5ffo/AAV9} = 25.99$, $FC_{9rh74\_4um9/9rh74} = 63.99$, and $FC_{9rh74\_4um9/9rh74} = 27.49$, Fig. 2D). To investigate how specific AAV_ITGs used αVβ6 as a cellular receptor, we tested their transduction under binding competition conditions. The number of AAV_ITG viral vectors entering the cells was significantly reduced when blocked by the recombinant protein αVβ6 before transduction, but no change occurred with AAV9 or AAV9rh74 (Fig. 2E). This result suggests that efficient transduction of AAV_ITGs requires specific binding to a αVβ6 complex.

During myogenesis, αVβ6 is only expressed in late differentiation, but not in the myoblast stage (Fig. S2D, E). We therefore hypothesized an enhanced transduction of AAV_ITGs in differentiated myotubes, but not myoblasts. We infected both human myoblasts and myotubes with AAV_ITGs. Low levels of luciferase activity were observed in all AAVs tested in human myoblasts (Fig. 2G, I). On the other hand, in human differentiated myotubes (hMT), VCN and luciferase activities in both AAV9rh74_4um9 and _5ff0 were significantly higher than for AAV9 or AAV9rh74 (Fig. 2F, H, J). In particular, variant AAV9rh74_4um9 showed a 16.56 ($p < 0.0001$) and 25.02-fold ($p < 0.0001$) improvement in luciferase activity compared to AAV9 and AAV9rh74, respectively, which is in agreement with its superior transduction efficiency and transgene expression seen in 293_αVβ6 cells.

We next examined the dependence of AAV_ITGs on the AAV's human cell entry receptor (AAVR)[40]. Structural and molecular studies have previously highlighted the importance of AAVR binding in cellular transduction of multiple natural[40,41] and engineered capsids[17,42] where VR4 plays a partial role in AAVR binding. We created a 293_AAVR-KO cell line (Fig. S2H) and test the viral transduction efficiency in comparison to 293_WT. Similar to other engineered capsids with peptide-inserted VR8[17,42], AAV9rh74_4um9 and _5ff0 reduced the AAV entry by 65.37% and 74.73%, reduced the transgene activity by 94.10% and 96.90% with the absence of AAVR (Fig. S2I, $n = 4–6$, $p < 0.0001$). It suggests that AAV_ITGs both require AAVR for effective transduction.

In summary, the two designed AAV_ITGs were both well-produced and function via αVβ6-specific binding and AAVR-dependent pathway, thus enhancing their transduction efficiency in 293_αVβ6 and human differentiated myotubes.

AAV_ITGs enhanced transduction in skeletal muscle following systematic administration

AAV_ITGs, together with AAV9 and AAV9rh74, were administrated systematically via intravenous injection (transgene: CMV_GFP-Luciferase, dose: 1E13 vg/kg, age at injection: 6 weeks, $n = 4$) in C57Bl6 mice to examine their biodistribution 3 weeks post-injection (Fig. 3A).

In agreement with a previous study, AAV9rh74 slightly reduces transduction in skeletal muscle compared to AAV9 but accumulates much less in the liver (Fig. 3B–D). Thanks to the liver-detargeting capsid and in accordance with the fact that αVβ6 is weakly expressed in the liver, we expected poor entry into the liver for designed AAV_ITGs. Indeed, AAV_ITGs is strongly detargeted from the liver, both at VCN and mRNA levels, even further than the parental capsid (Fig. 3C, D). In contrast, enhanced transduction was observed in all skeletal muscles that were tested, including the tibialis anterior (TA), quadriceps (Qua) and diaphragm (Dia) (Fig. 3B–D). The two AAV_ITGs both showed a substantial increase in VCN and luciferase activity compared to both AAV9 and AAV9rh74. Similar to the results obtained in in vitro models, AAV9rh74_4um9 is the best transducer among the two AAV_ITGs. Compared to AAV9, the variant 9rh74_4um9 significantly increased VCN 5.31/7.21/2.48-fold and increased luciferase activity 15.2/13.2/23.57-fold in Qua, TA, and Dia ($p < 0.05$), respectively. Compared to the original backbone AAV9rh74, this variant even magnified the difference by increasing VCN 5.53/2.85/7.69-fold and increasing luciferase activity 152.35/106.68/60.43-fold ($p < 0.05$). Furthermore, AAV9rh74_4um9, but not AAV9rh74_5ffo, significantly increased transduction in the heart ($FC_{VCN} = 4.15$, $FC_{LUC} = 15.43$, $p < 0.05$). All AAVs that were tested showed poor delivery and transgene expression in the lungs and kidneys. No alteration of TGFβ and integrin signaling was observed at 1 month post-injection in all AAVs being tested (Fig. S2F, G). Overall, these data indicate that AAV_ITGs, especially the 9rh74_4um9 variant, are strongly liver-detargeted and exhibit enhanced tropism towards skeletal and cardiac muscles.

AAV9rh74_4um9 exhibits highest tropism towards the skeletal muscle among tested myotropic AAVs in mice

Several engineered myotropic AAVs (mAAVs), including AAVMYO[15], MYOAAV-1A and −2A[17], have demonstrated superior efficacy for in vivo delivery of muscle compared to natural AAVs. To evaluate the properties of these AAVs compared to ours, we performed in vitro and in vivo experiments (Fig. 4A, B). Viral preparations were produced using the same reporter transgene (CMV_GFP-Luc). All mAAVs were well-produced in 400 ml suspension, with higher titers than AAV9rh74. However, AAVMYO productivity was significantly lower than 9rh74_ITGs and MYOAAVs (Fig. S3A). Since all investigated mAAVs shared a common integrin-targeting RGD motif, these AAVs were then evaluated for their transduction via integrin complexes in myotubes and in cell lines where integrin complexes were stably overexpressed by the PiggyBac system. In 293_αVβ6 cells as well as in hMT, where αVβ6 is highly expressed, AAV9rh74_4um9 showed the highest transduction among the tested myotropic AAVs, with the sole exception that luciferase activity of MYOAAV2A was higher in hMT (Fig. S3B, C). We also tested AAV transduction efficiency in two other cell lines, 293_WT, where αVβ6 expression is low, and 293_α7β1 that stably overexpresses a non-RGD-targeting α7β1 integrin. In both conditions, MYOAAV2A and AAV9rh74_4um9 showed the highest transduction (Fig. S3D, E). These results suggest that, as intended with the rational design, AAV9rh74_4um9 uses αVβ6 more preferentially for cellular transduction than others, yet it can also efficiently use other integrin(s) similar to MYOAAV2A.

Following in vivo injection in the same setting as described above (6-week-old WT mice, dose: 1E13 vg/kg, $n = 4$), the three mAAVs and 9rh74_4um9 all showed strong liver-detargeting, high enrichment in both skeletal and cardiac muscles, and negligible transduction levels in other organs that were tested (kidneys, lungs, and brain) (Fig. 4C, D). No significant difference was observed in either VCN or luciferase activity between all three mAAVs and 9rh74_4um9 in the skeletal muscles that were tested. In heart muscle, MYOAAV2A showed a significant increase in VCN compared to other myotropic vectors, but no difference in luciferase activity, in agreement with the original observation[17]. The most striking difference is the level of liver-detargeting between these vectors. The VCN for 9rh74_4um9 in liver is 3.34/22.05/13.85 times lower than for AAVMYO ($p = 0.0022$), MYOAAV-1A ($p = 0.0013$) and −2A ($p = 0.033$), respectively (Fig. 4C). Recently, AAVMYO2 and AAVMYO3 myotropic variants were identified to be more liver-detargeted while transducing skeletal and cardiac muscles less effectively than AAVMYO[16]. We then conducted a similar

https://doi.org/10.1038/s41467-024-52002-4

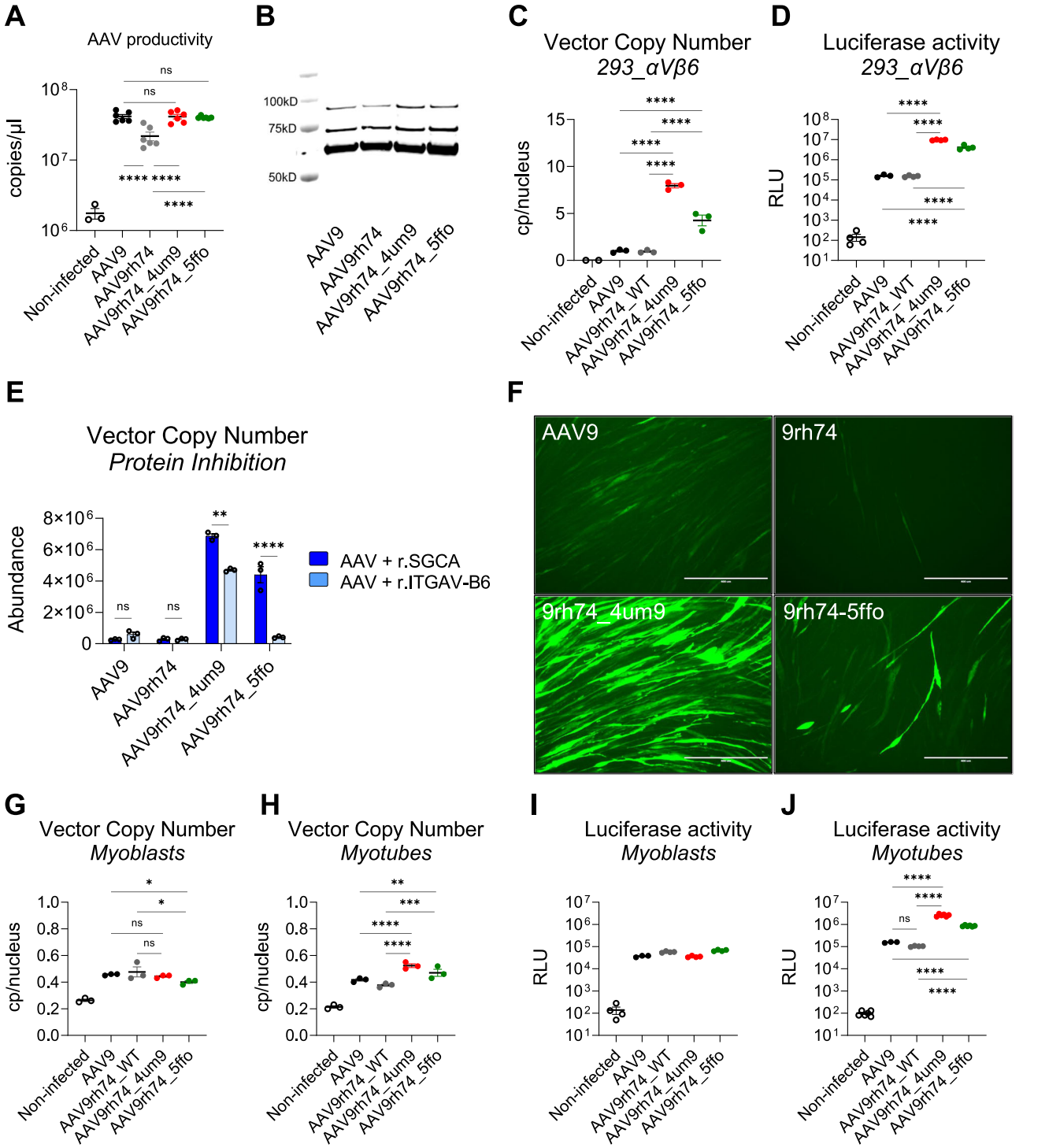


Fig. 2 | Designed AAV_ITGs were well-produced and improved transduction via αVβ6 binding. A AAV titers of different AAV variants in bulked small-scale production in suspension 3-day post-triple-transfection (2 ml production, 6 biological replicates, one-way ANOVA followed by FDR correction). **B** Western blot of VP proteins from purified AAVs showed similar VP ratios for designed AAV_ITGs capsids compared to AAV9 and AAV9rh74, suggesting successful capsid assembly. **C**, **D** VCN (**C**) and luciferase activity (**D**) of 293_αVβ6 after AAV infection (3 and 4 biological replicates, one-way ANOVA followed by FDR correction). Both the two designed AAV_ITGs showed enhanced VCN and luciferase activities compared to AAV9rh74 and AAV9. **E** Inhibition of cell entry of designed AAV_ITGs, but not for AAV9 or AAV9rh74, in 293_αVβ6 cells by αVβ6 recombinant protein. AAVs were preincubated with human αVβ6 recombinant protein (r.ITGAV-B6) for 30 min at 37 °C before infection (4 biological replicates, two-way ANOVA followed by FDR correction, 1 μg protein per 5E9vg AAV, 2E5 vg per cell). The same condition treated with recombinant SGCA protein (r.SGCA) was used as the control. **F**–**J** Enhanced transduction of AAV_ITGs in in vitro human differentiated myotubes, but not in myoblasts. **F** Representative images of the GFP signal of myotubes 48 h post-infection (scale bar: 400 μm). **G**–**J** VCN and luciferase activities of AAV_ITGs in comparison with AAV9 and AAV9rh74 in myoblasts (**G**, **I**, 3 and 3/4 biological replicates, respectively) and myotubes (**H**, **J**, 3 and 3/4/6 biological replicates, respectively) (one-way ANOVA followed by FDR correction). Data are presented as mean ± SEM. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$; ns not significant. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-024-52002-4

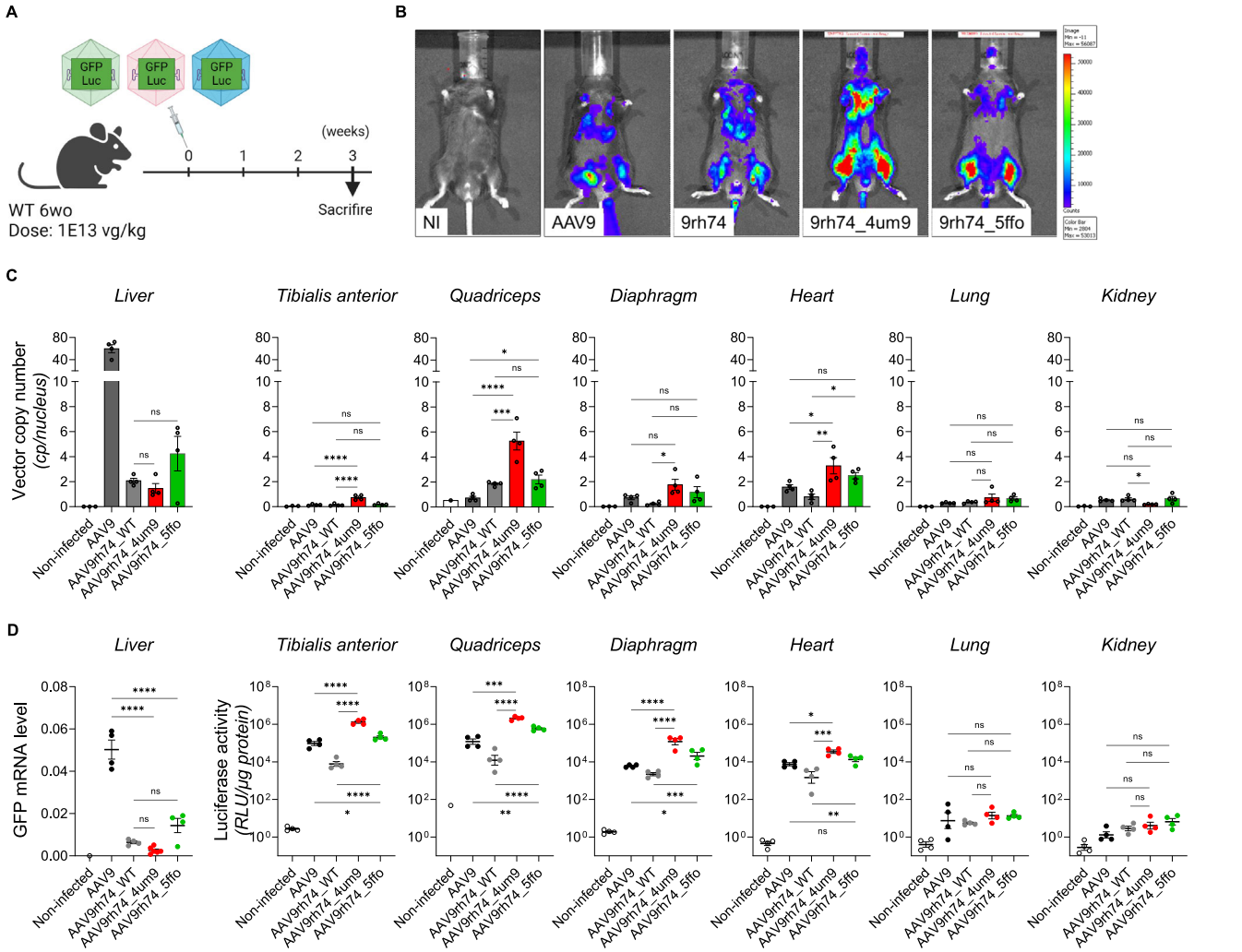


Fig. 3 | Designed AAV_ITGs showed enhanced transduction in skeletal and cardiac muscles while strongly liver-detargeted in vivo. A Scheme of in vivo experiment. AAVs (CMV_GFP-Luciferase) were injected intravenously into 6wo C57BL6 mice at the dose of 1E13 vg/kg. Created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en). **B** Representative images of the bioluminescence signal 20 days post-infection. **C**, **D** VCN (**C**) and gene expression (**D**) (GFP mRNA level in the liver and luciferase activity in other organs) for different AAVs in liver, skeletal muscles, heart, lung, and kidney (4 biological replicates, one-way ANOVA followed by FDR correction). Both designed AAV_ITGs strongly detargeted from the liver compared to AAV9, while they significantly improved VCN and luciferase activities over AAV9rh74 (and AAV9 with AAV9rh74_4um9 variant) in skeletal and cardiac muscles, and were detected and expressed at low levels in lung and kidney. Data in (**C**, **D**) are presented as mean ± SEM. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$; ns not significant. Source data are provided as a Source Data file.

head-to-head comparison of these two variants with 9rh74_4um9 (6-week-old WT mice, dose: 1E13 vg/kg, $n = 4$). All three capsids showed very low transduction in kidney, lung and brain (Fig. S3F). Similar to previous AAVMYO comparison[16], AAVMYO2 and AAVMYO3 both showed significantly lower VCN (3.24–7.44/2.61–6.16 times, respectively, $p < 0.001$) and transgene activity (12.41–18.42/11.1–15.91 times, respectively, $p < 0.05$) in all tested skeletal/cardiac muscles compared to LICA1 (Fig. S3F, G). On another hand, the liver transduction of AAVMYO2/3 is not significantly different from that of LICA1, yet the transgene expression is significantly lower (Fig. S3F,G). Altogether, LICA1 and AAVMYO2/3 showed highest ratio of muscle-to-liver transduction in all tested myotropic capsids (Fig. S3H, I), yet LICA1 is considerably better in transgene delivery and expression in skeletal muscles than AAVMYO2/3. These data indicate higher muscle specificity for the 9rh74_4um9 variant compared to other myotropic vectors that have been investigated to date.

In summary, the 9rh74_4um9 variant, hereafter referred to as LICA1 (linked-integrin-complex AAV), consistently showed enhanced transduction and strong liver-detargeting. Therefore, we then attempted to evaluate LICA1 as a delivery vector for muscular dystrophies, in comparison with AAV9. Two different setups will be investigated: the transfer of microdystrophin (µDys)—an incomplete transgene—in mdx, a mild mouse model of Duchenne muscular dystrophy (DMD) and of the full-length human α-sarcoglycan (SGCA) in a severe mouse model of limb-girdle muscular dystrophy R3 (LGMD-R3).

Low-dose LICA1-µDys gene transfer is effective in specifically overexpressing microdystrophin in dystrophic muscle but not sufficient to fully correct the underlying pathology

DMD is caused by mutations in the DMD gene, which encodes for dystrophin protein—a key player in the dystrophin-glycoprotein complex (DGC), which is critical for the structural stability of skeletal muscle fibers[43]. Lack of dystrophin can result in progressive loss of muscle function, respiratory defects, and cardiomyopathy. The most commonly used DMD animal model is the mdx mouse, with a lifespan reduced by 25%, milder clinical symptoms than those seen in human patients, with the exception of the diaphragm muscle[44]. Among many therapeutic strategies to restore dystrophin expression, high-dose

https://doi.org/10.1038/s41467-024-52002-4

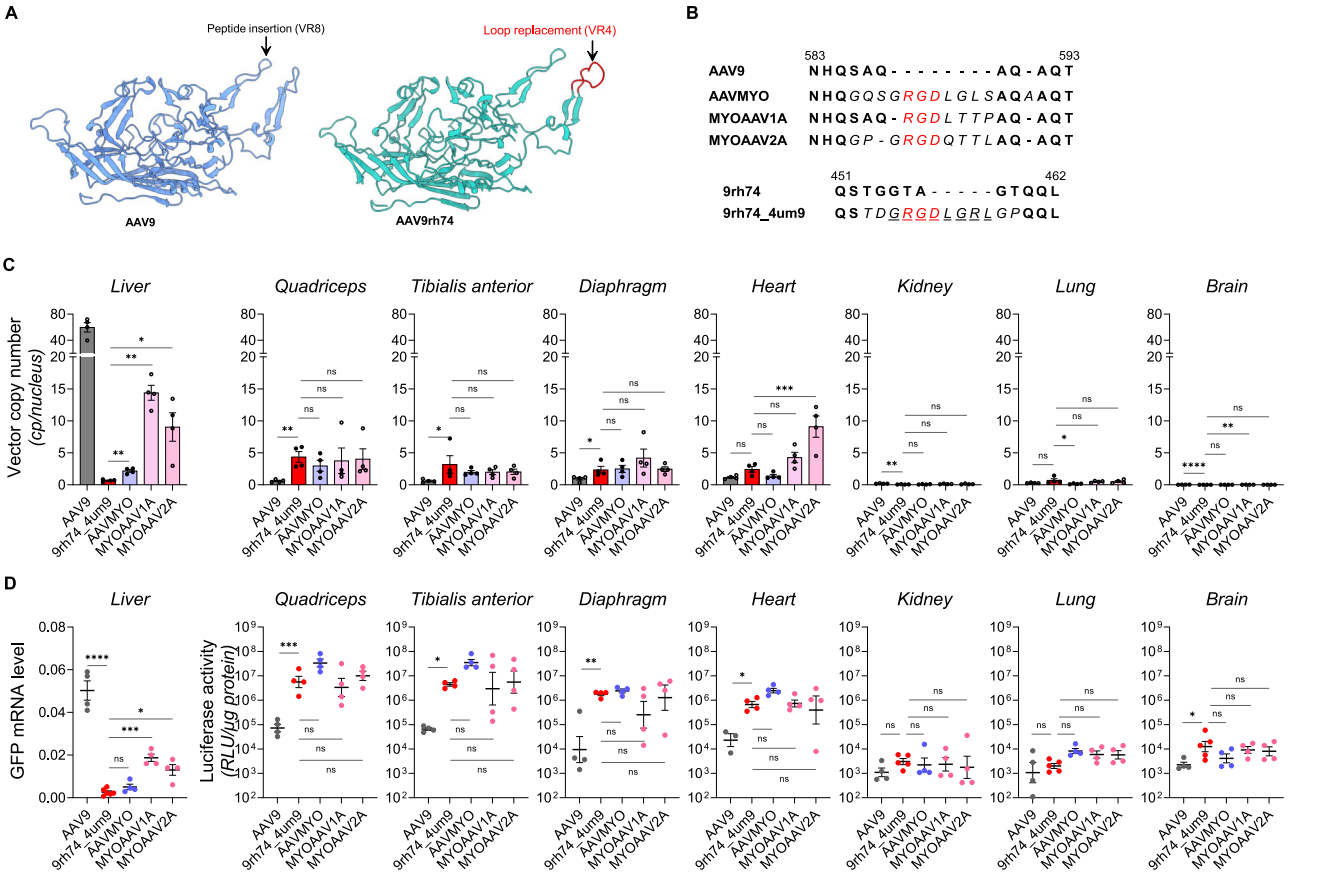


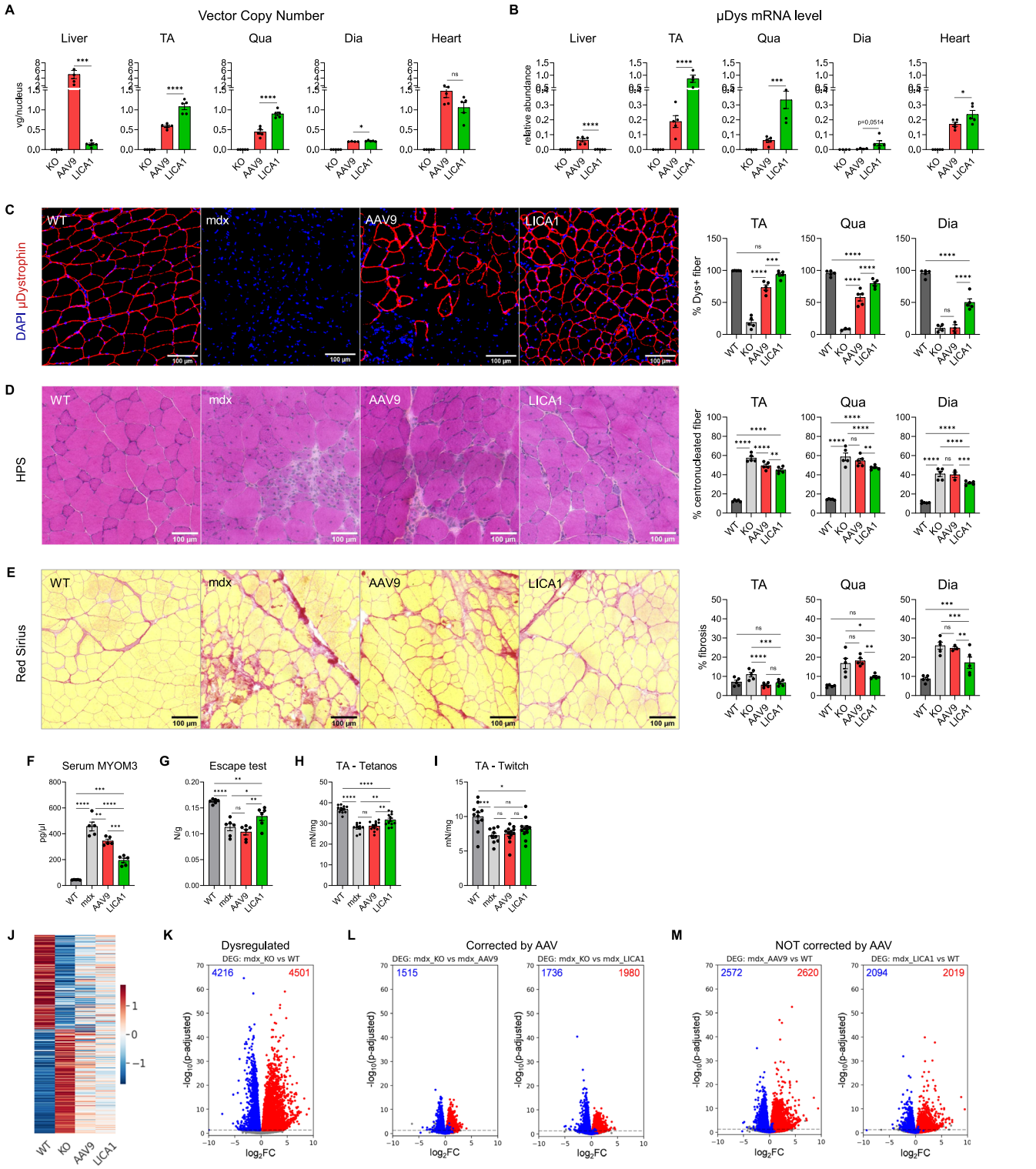
Fig. 4 | AAV9rh74_4um9 exhibits highest tropism towards the skeletal muscle among tested myotropic AAVs in mice. Comparison of the AAV9rh74_4um9 variant with other public myotropic AAVs (mAAVs)[15,17]. **A** Illustration of the differences between mAAVs and AAV9rh74_4um9 at modification sites in capsid protein and modification methods. **B** The VR8 loop sequences of mAAVs compared to VR8 of their backbone AAV9, and VR4 of AAV9rh74_4um9 compared to VR4 of AAV9rh74. **C**, **D** VCN (**C**) and gene expression (**D**) (GFP mRNA level in liver and luciferase activity in other organs) of different AAVs in liver, skeletal muscles, heart, lung, kidney, and brain (4 biological replicates, one-way ANOVA followed by FDR correction). AAV9rh74_4um9 showed similar VCN and gene expression in skeletal muscle to other mAAVs, while being significantly more strongly detargeted from the liver. Data in (**C**, **D**) are presented as mean ± SEM. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$; ns not significant. Source data are provided as a Source Data file.

AAV-based gene transfer of shortened functional forms of the dystrophin ORF provided excellent results in animal models, but unsatisfactory conflicting data in current clinical trials[6]. Severe toxicities, even patient death, have been reported from these trials (NCT03368742, NCT04281485), assumed to be related to the dose of ≥ 1E14 vg/kg. We therefore explored the possibility of low-dose µDys gene transfer[45] in mdx mice using LICA1 in comparison to AAV9 (Fig. S4A, age at injection: 4 weeks, dose: 5E12 vg/kg, treatment duration: 4 weeks, $n = 5$). Three muscles with increasing levels of severity—TA, Qua, and Dia—were used to study AAV transduction and treatment efficacy.

LICA1 showed better µDys gene transfer than AAV9 in this model. LICA1-treated mice exhibited a significantly higher VCN in all 3 muscles that were tested, 1.85/2.02/1.07 times higher in TA ($p < 0.0001$), Qua ($p < 0.0001$), and Dia ($p = 0.020$), respectively (Fig. 5A). RNA levels indicated even greater differences and were 4.56–7.57 times higher in the LICA1-treated group (Fig. 5B; TA: FC = 4.56, $p < 0.0001$; Qua: FC = 5.46, $p = 0.0001$; Dia: 7.57, $p = 0.05$). Consequently, LICA1 can transduce almost 100% in TA and Qua, and 49.98% in Dia, while substantially lower numbers were seen in AAV9-treated muscles, at 73.22% ($p = 0.0001$), 57.8% ($p < 0.0001$), 10.34% ($p < 0.0001$) in TA, Qua, Dia, respectively (Fig. 5C, Fig. S4B). Furthermore, while infection levels and expression of the transgene in liver were high for the AAV9 vector (despite the use of muscle-specific promoter), the VCN and mRNA levels in LICA1-treated liver were extremely low (Fig. 5A, B, $FC_{VCN:AAV9/LICA1} = 36.8$, $p = 0.0002$; $FC_{mRNA:AAV9/LICA1} = 64.7$, $p < 0.0001$). In the heart, there were no difference in VCN but slightly significantly higher in mRNA level in LICA1-treated group compared to AAV9 (Fig. 5A, B, $FC_{VCN} = 0.72$, $p = 0.073$; $FC_{mRNA} = 1.39$, $p = 0.021$). These data again confirmed the transduction efficiency and specificity towards skeletal muscle for the LICA1 vector, even with low-dose treatment.

The histological features and muscle functionality after AAV treatment were restored accordingly. The centronucleation index (percentage of centronucleated fibers)—an indicator of the regeneration/degeneration process—did not change with AAV9 (except in TA) but was significantly reduced upon LICA1 treatment (reduction of 21.68%, 19.05%, 22.88% in TA, Qua, Dia, respectively) (Fig. 5D and Fig. S4C). Similarly, the fibrosis level in two severely affected muscles, Qua and Dia, only exhibited a significant reduction with LICA1, but not AAV9 (Fig. 5E and Fig. S4D). The serum biomarker MYOM3 level, an indicator of muscle damage[46], showed a reduction for both AAV treatments, with a considerable further reduction seen in the LICA1-treated group (Fig. 5F, $FC_{AAV9/KO} = 0.75$, $FC_{LICA/KO} = 0.43$, $p_{AAV9\text{-}LICA1} < 0.0001$). More importantly, AAV9 treatment did not affect any muscle functionality being tested (Fig. 5G–I), while significant improvements with LICA1-µDys treatment were observed in escape test—a measure of global force (Fig. 5G, $FC_{LICA1/mdx} = 1.19$, $p_{LICA1/mdx} = 0.02$) and in situ TA mechanical force measurement (Fig. 5H, $FC_{LICA1/mdx} = 1.14$, $p_{LICA1/mdx} = 0.0006$). However, none of the treatment normalized to the WT functional levels. These data indicate that

https://doi.org/10.1038/s41467-024-52002-4



LICA1 is better than AAV9 at restoring dystrophic histological features and muscle functions.

We also investigated the molecular alteration in Qua upon AAV treatment using RNA-seq (Supplementary Data 1). On the two first principal components (PCs) of the PCA, a clear distinction between four transcriptome groups (WT, mdx, AAV9, LICA1) was observed, while LICA1-treated muscles were clustered closer to the WTs than others (Fig. S4E). To our surprise, despite excellent transgene expression by LICA1, global transcriptomic restoration was relatively modest (Fig. 5J). Nevertheless, a substantial improvement can still be seen for LICA1 compared to AAV9. Among 4216 down- and 4501 upregulated differentially expressed genes (DEGs) identified in mdx muscle, 1515 (35.9%) and 1728 (38.4%) were restored by AAV9, while LICA1 was able to correct 1736 (41.2%) and 1980 (44.0%), respectively (Fig. 5K, L). In addition, a greater number of genes were either not or insufficiently corrected by AAV9 than by LICA1 (Fig. 5M). A total of 2572 genes were downregulated (61.0%) and 2620 (58.2%) incompletely restored, while significantly lower numbers were seen for LICA, with 2094 (49.67%) down- and 2019 (44.86%) upregulated. Interestingly, some known dysregulated pathways, including α- and ϒ-interferon

Fig. 5 | Low-dose gene transfer by LICA1 was more effective and better at restoring dystrophic phenotypes than AAV9 in the DMD mouse model. **A**, **B** Comparison of transduction efficacy between AAV9 and LICA1 in three skeletal muscles and heart, in terms of VCN (**A**), and μDys RNA level (**B**) ($n = 4/5$). **C** Comparison of percentage of successfully transduced (dystrophin-positive) fibers in the skeletal muscles ($n = 3/5$). **D**, **E** Comparison of restoration levels in dystrophic histological features between AAV9 and LICA1 in terms of percentage of centro-nucleated fibers (**D**) and fibrosis level (**E**) ($n = 3/5$). Illustrated images in (**C**–**E**) are of quadriceps muscles (scale bar: 100 μm). **F** Serum MYOM3 level at 4 weeks post-injection ($n = 5$). **G**–**I** Comparison of functional restoration between AAV9 and LICA1 by Escape test—global force measurement (**G**, $n = 6$), tetanus force of TA muscle (**H**, $n = 10/12$), and twitch force of TA muscle (**I**, $n = 9/11/12$). **J**–**M** Comparison of restoration in global transcriptomic changes in quadriceps muscle between AAV9 and LICA1 ($n = 4$, adjusted p-values < 0.05). **J** The heatmap shows the log2 fold-change (log2FC) in comparison to WT muscle for all 8717 DEGs in mdx muscle, displayed as row Z-scores from blue (lowest) to red (highest). **K**–**M** Volcano plots of multiple comparisons illustrate transcriptomic changes before and after AAV treatment. As a reference, 4216 downregulated and 4501 upregulated DEGs found in mdx were colored blue and red, respectively, in all volcano plots. Among these DEGs, the number of genes found to be significantly different in each pair-wise comparison were labeled in the upper corners. **K** mdx versus WT. **L** mdx versus AAV treatments (significant DEGs are the genes correctly restored). **M** AAV treatments versus WT (significant DEGs are the genes that are not or incompletely restored). Data in (**A**–**I**) are presented as mean ± SEM. n represents the number of biological replicates. Statistics in (**A**, **B**) and (**F**–**I**) were performed using one-way ANOVA followed by FDR correction (**A**, **B**) or post-hoc test (**F**–**I**). Statistics in (**C**–**E**) were performed using two-way ANOVA (-serotypes * IHF slides) followed by FDR correction. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns not significant. Source data are provided as a Source Data file.

responses and oxidative phosphorylation, were significantly better normalized by LICA1 than by AAV9 (Fig. S4F).

In summary, at 5E12 vg/kg, LICA1-μDys, but not AAV9, was efficient in transducing close to 100% myofibers, except in the diaphragm. This effective improvement in transduction can significantly reduce some dystrophic features in all muscles that were tested, yet restoration in the global transcriptome remains modest. However, greater improvements in functional, histological, and transcriptomic restoration were achieved with LICA1 compared to AAV9.

Low-dose LICA1-SGCA treatment restored the muscle functionality, dystrophic phenotypes, and transcriptomic dysregulation in a severe SGCA mouse model

LGMDR3 is caused by mutations in the SGCA gene[47]—another component of the DGC complex. Defects in the SGCA protein therefore lead to muscle weakness and wasting. A LGMDR3 mouse model has been established, which closely represents patient's clinical phenotypes[48]. Similar to the setting in mdx mice, low-dose AAV treatment with 5E12 vg/kg was investigated in this mouse model. AAV9 or LICA1 encoding human SGCA (hSGCA) under control of a muscle-specific human Acta1 promoter were injected into 4-week-old SGCA-KO mice (Fig. 6A). Analysis was performed 4 weeks post-treatment.

In all three muscles that were tested, TA, Qua, Dia (in order of increasing severity), transduction in various measures, VCN, mRNA level, and percentage of SGCA+ myofibers, was significantly greater in the LICA1-treated group than for AAV9 (Fig. 6B–D, Fig. S5A). Of note is the fact that the differences in transduction efficacy (%SGCA + myofibers) between LICA1 and AAV9 are greater in more severely affected muscles (Fig. 6D). At such a low dose, AAV9 was able to transduce > 80% myofibers in TA while LICA1 can reach close to 100% ($p < 0.0001$). While LICA1 still transduced almost 100% of fibers in Qua (the muscle affected with intermediate severity), only 58.1% fibers were transduced by AAV9 on average ($p < 0.0001$). In the most severely affected muscle, Dia, both vectors displayed reduced efficiency; however, LICA1 continued to demonstrate much better transduction ($\mu_{AAV9} = 22.1\%$, $\mu_{LICA1} = 59.5\%$, $p < 0.0001$). Besides, slight significant increase was observed in both VCN and mRNA level in LICA1-treated group compared to AAV9 in the heart (Fig. 6B, C, $FC_{VCN} = 2.01$, $p = 0.0069$; $FC_{mRNA} = 5.11$, $p = 0.018$).

The differences in transgene delivery and expression positively correlated with levels of histological and functional restoration. Different dystrophic histological features, including percentage of centronucleated fibers (Fig. 6E, Fig. S5B), percentage of fibrosis area (Fig. 6F, Fig. S5C), and fiber size distribution (Fig. 6G), were all significantly better normalized by LICA1 than AAV9, especially in more severely affected muscles. Importantly, no significant improvement was observed in the AAV9-treated group in centronucleation index and fibrosis level in Dia, while LICA1 reduced these parameters by half (Fig. 6E, F). Fiber sizes were also restored to near-WT distribution by LICA1 in this muscle (Fig. 6G). No difference in body weight was seen between groups with or without AAV treatment (Fig. S5D). At the functional level, however, the escape test—a measure of global force—showed a significant increase in AAV9-treated mice ($FC = 1.42$, $p = 0.0072$) and was even higher in LICA1-treated group ($FC = 1.72$, $p < 0.0001$) (Fig. 6H). On the other hand, in situ TA mechanical forces were both improved in the two AAV groups at similar levels (Fig. 6I), possibly due to > 80% transduction rate by both vectors. Similar to the global force, the serum MYOM3 level was greatly reduced in the LICA1-treated group but not for AAV9, indicating less muscle damage (Fig. 6J). No difference was seen in the anti-capsid antibody between the two AAV treatments (Fig. S5E). These results indicate that better and significant functional and histological restoration in the LICA1-treated mice was achieved, even at low-dose treatment, thanks to superior transduction efficacy.

We further investigated the molecular alterations following AAV treatment by transcriptomic profiling of the quadriceps muscle (Supplementary Data 2). The first principal component (PCs) of the PCA was able to separate a group including WT and LICA1 with a group including SGCA-KO and AAV9, suggesting close proximity between elements within these 2 groups (Fig. S5F). A heatmap of all 8591 significant DEGs (4035 downregulated and 4556 upregulated) further highlighted the restorative effect of LICA1 on gene expression levels (Fig. 6K). LICA1-treated muscles, in particular, demonstrated a significant correction of 69.9% (2821/4035) and 66.5% (3028/4556) of down- and upregulated DEGs, respectively, compared to 12.4% (500/4035) and 9.21% (420/4556) corrected by AAV9 treatment (Fig. 6L, M). Conversely, not all DEGs were significantly restored or returned to WT levels. The number of such transcripts in AAV9-treated muscles was much higher than in the LICA1-treated group (Fig. 6N): 2541 (63.0%) downregulated DEGs and 3045 (66.8%) upregulated DEGs for AAV9, with only 483 (12.0%) downregulated DEGs and 1038 (22.8%) upregulated DEGs in the LICA1-treated group. These data illustrate that low-dose LICA1 treatment can effectively normalize the majority of the dysregulated transcriptome and is much more efficient in correcting gene expression dysregulation than AAV9 at the same dose.

In summary, low-dose (5E12 vg/kg) AAV gene transfer using LICA1 in the LGMDR3 mouse model is effective in restoring muscle function, dystrophic histology, and the dysregulated transcriptome. The efficacy was much greater than for AAV9 at the same dose due to enhanced transduction.

LICA1 showed enhanced muscle transduction across multiple species

While the initial design was targeting the human integrin αVβ6 complex, we further investigated whether AAV_ITG variants, including LICA1, exhibit myotropism in multiple species.

First, protein sequences of both integrin subunits at the binding interface with TGFβ1/3-derived binding motifs are shown to be

https://doi.org/10.1038/s41467-024-52002-4

identical between human, primate, pig, rat, and mouse (Fig. 7A, B). Structure prediction by AlphaFold2 showed highly similar structure of αVβ6 heterocomplex in different species compared to the human form (Fig. S6A, RMSD ≤ 0.431 Å, pLDDT>95). In addition, pre-incubation of AAV_ITGs with recombinant αVβ6 proteins from human, mouse, and rat all resulted in strong significant reduction of AAV transduction in dose-dependent manner in two different cell lines (293_WT and 293_ αVβ6) (Fig. 7C and Fig. S6B, $n = 3$, two-way ANOVA). It is noteworthy that the inhibition by recombinant αVβ6 in 9rh74_5ffo variant is significantly stronger than LICA1 in most conditions tested (Figs. 2E, 7C, and Fig. S6B). Together, it indicates that both AAV_ITGs can efficiently bind to αVβ6 receptors from multiple species and the AAV_ITGs/αVβ6 interaction is essential for cellular transduction.

We then conducted a comparative multiplexed study in in vitro human myotubes and in vivo in C57Bl6 mouse and NHP (Fig. 8A). An expression cassette flanked by ITR sequences consists of an ORF coding for a human inactive CAPN3 protein (C129S mutation, denoted as hC3i) under the control of skeletal muscle-specific tMCK

Fig. 6 | Low-dose gene transfer by LICA1 was better at restoring dystrophic phenotypes and functionality than AAV9 in the LGMDR3 mouse model. **A** Scheme of in vivo experiment. Created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en). **B**–**D** Comparison of transduction efficacy between AAV9 and LICA1 in terms of VCN (**B**, $n = 4/5$), hSGCA mRNA level (**C**, $n = 4/5$), and percentage of succesfully transduced (SGCA-positive) fibers (**D**, $n = 3/4/5$). **E**–**G** Comparison of restoration levels in dystrophic histological features between AAV9 and LICA1 in all three muscles that were tested in terms of percentage of centro-nucleated fibers (**E**, $n = 3/4/5$), fibrosis level (**F**, $n = 3/4/5$), and fiber size distribution (**G**, $n = 3/4/5$). Illustrated images in (**D**–**F**) are of quadriceps muscles (scale bar: 100 μm). **H**–**J** Comparison of functional restoration between AAV9 and LICA1 using the escape test—global force measurement (**H**, n = 3/5), tetanus force of TA muscle (**I**, $n = 3/4/5$), and serum MYOM3 level—indicator of muscle damage (**J**, $n = 3/4/5$). **K**–**N** Comparison of restoration in global transcriptomic changes in quadriceps muscle between AAV9 and LICA1 ($n = 4$, adjusted p-values < 0.05). **K**. The heatmap presents the log2FC in comparison to WT muscle for all 8591 DEGs found in KO muscle (compared to WT), displayed as row Z-scores from blue (lowest) to red (highest). **L**–**N** Volcano plots of multiple comparisons illustrate transcriptomic changes before and after AAV treatment. As a reference, 4035 downregulated and 4556 upregulated DEGs found in KO were colored blue and red, respectively, in all volcano plots. Among these DEGs, the number of genes found to be significantly different in each pair-wise comparison were labeled in the upper corners. **L** KO versus WT. **M** KO versus AAV treatments (significant DEGs are the genes correctly restored). **N**. AAV treatments versus WT (significant DEGs are the genes that are not or incompletely restored). Data in (**B**–**F**) and (**H**–**J**) are presented as mean ± SEM. n represents the number of biological replicates. Statistics in (**B**, **C**) and (**H**–**J**) were performed using one-way ANOVA followed by FDR correction. Statistics in (**D**–**F**) were performed using two-way ANOVA (-serotypes * IHF slides) followed by FDR correction. *$p < 0.05$; ** $p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns not significant. Source data are provided as a Source Data file.

promoter[49]. A unique DNA sequence (barcode/BC) was inserted after the stop-codon and before the SV40 polyA, which allows the quantification of the transgene expression mRNA level in a multiplex manner. Each capsid variant was produced separately with 2-3 unique barcodes, before pooling at the same viral genome number per capsid and then being co-infected in in vitro myotubes or co-injected intravenously in vivo.

First, a pool of AAV8, AAV9, AAV9rh74, LICA1, 9rh74_5ffo, AAV-MYO, and MYOAAV2A were coinjected in C57Bl6 mice ($n = 3$, dose: 5E12 vg/kg per capsid variant, duration: 21 days). Thank to the tMCK promoter, total hC3i mRNA level showed strong enrichment in skeletal muscle (Dia, Qua, TA), low expression in liver, heart, kidney, lung, dorsal root ganglion, and brain, as expected[49] (Fig. S6C). Similar to the individual comparison (Figs. 3, 4), LICA1 again showed strong liver-detargeting, greatly improved muscle transduction in all tested skeletal muscles than AAV8, AAV9, AAV9rh74 (Fig. 8B, $FC_{LICA1/AAV9} = 17.86/12.55/9.54$, $FC_{LICA1/AAV8} = 148.50/114.43/13.24$, $FC_{LICA1/AAV9rh74} = 53.37/34.00/30.74$, respectively in Qua, TA, and Dia). The transduction level of LICA1 in skeletal muscles showed no significant difference compared to two myotropic benchmarks, AAVMYO and MYOAAV2A, except slightly but significantly higher than AAVMYO in Quadriceps ($FC_{LICA1/AAVMYO} = 1.35$, padj = 0.0097).

In in vitro human myotubes, a pool of AAV8, AAV9, AAV9rh74, LICA1, 9rh74_5ffo, and MYOAAV2A were co-infected at 2 doses, 2×10^{10} and 2×10^{11} vg per 12-wp well (Fig. 8C, $n = 2$). At both concentrations, AAV_ITGs showed considerably enhanced transgene expression compared to natural capsids, AAV8 and AAV9, and their backbone, AAV9rh74. Unexpectedly, 9rh74_5ffo showed highest transgene expression level, 3.22–4.65-fold higher than LICA1 and 1.63–2.42-fold higher than MYOAAV2A. This observation contradicts with the individual comparison in various models. In addition, 9rh74_5ffo showed a more-sensitive inhibition by recombinant αVβ6 as described above and all three myotropic capsids utilizes the same receptor—αVβ6, suggesting a potential competition between different capsids in the multiplex setup, as previously described[50].

We then evaluated gene transfer efficacy of LICA1 in *Macaca fascicularis* ($n = 3$). Due to the potential risk of competition and the profiles of pre-existing neutralizing antibodies (Nabs) against different capsid variants (Table S3, Fig. S6D, no primate was Nab-negative for all capsid tested), minimal number of AAVs were injected intravenously in NHPs presenting the lowest Nab titers for LICA1. One NHP was injected with a pool of AAV9 + LICA1 (NHP1) and two others with a pool of AAV8 + LICA1 (NHP2/3) (Fig. 8A, 3.2E12 vg/kg per capsid variant, duration: 45 days). Among all tissues tested, transgene expression measured by hC3i mRNA level is restricted in the skeletal muscle, relatively low in the heart, lymph nodes, spleen, kidney, lung, and brain, as expected[49] (Fig. S6E). Heterogeneity of transgene expression among primates might be due to the low dose injected. Unexpectedly, a discrepancy of the expression in the liver was observed between NHP and mouse: hC3i liver expression is highest among tested tissues in NHP while almost suppressed in mice, presumably due to tMCK promoter. Despite, LICA1 showed a strong reduction of liver transgene expression compared to AAV-9/−8 in all three NHPs (Fig. 8D, $FC_{AAV9/LICA1} = 46.41$, $FC_{AAV8/LICA1} = 50.64$). Conversely, the mRNA level of LICA1 is significantly higher than that of AAV-9/−8 in all tested skeletal muscles (Fig. 8D, $FC_{LICA1/AAV9} = 2.94–60.75$, $FC_{LICA1/AAV8} = 4.47–21.85$), except in TA. Importantly, despite the slight presence of Nab against LICA1 in NHP2/3 (Table S3, Fig. S6D), LICA1 still express transgene efficiently and significantly better than AAV8. This data confirms that LICA1 is more effective in transducing skeletal muscle and more liver-detargeted than natural AAV8/9 in NHP.

Altogether, our experiments shows that AAV_ITGs, particularly LICA1, can target αVβ6 across multiple species and exhibit myotropism as well as strong liver-detargeting in both murine and NHP.

Discussion

Given the severe complications observed with very high dose AAV treatment, lowering the dose by increasing vector specificity via capsid modification is one way to address these issues. This study investigated the possibility of altering AAV tropism towards skeletal muscle by targeting integrin. We designed an AAV as a αVβ6 binder, which resulted in a myotropic AAV variant, namely LICA1. LICA1 showed greatly enhanced transduction in skeletal muscle across species, including WT and two MD mouse models as well as NHP. Consequently, by improving the delivery of therapeutic transgenes (hSGCA and μDys) in these MD mouse models, LICA1 was able to correct dystrophic phenotypes, global transcriptional dysregulations and significantly restore muscle function.

AAV capsid sequence design method that ensures high AAV production

AAV tropism is commonly altered by inserting a small peptide into the VR4 or VR8 loop without any sequence constraints. Since no consideration regarding AAV capsid stability is included in this method, the resulting AAV can suffer from instability, reduced productivity, and increased AAV genome fragmentation[17,51] (ASGCT 2023). In the current study, a physics-based protein sequence design method was used to graft the binding motifs from TGFβ−1 and −3 into the VR4 loop of the hybrid capsid AAV9rh74. The major differences to the classical peptide insertion method are that the entire VR4 loop was modified to include a new binding motif and the amino acids around this motif (linkers) were selected to minimize the potential energy. Low-energy sequences ensure the stability and intended folding of the designed proteins, presumably leading to improved stability of the AAV particle[39]. Six AAVs designed using this method were tested experimentally and all showed better productivity than their parent, Cap9rh74, and similar

https://doi.org/10.1038/s41467-024-52002-4

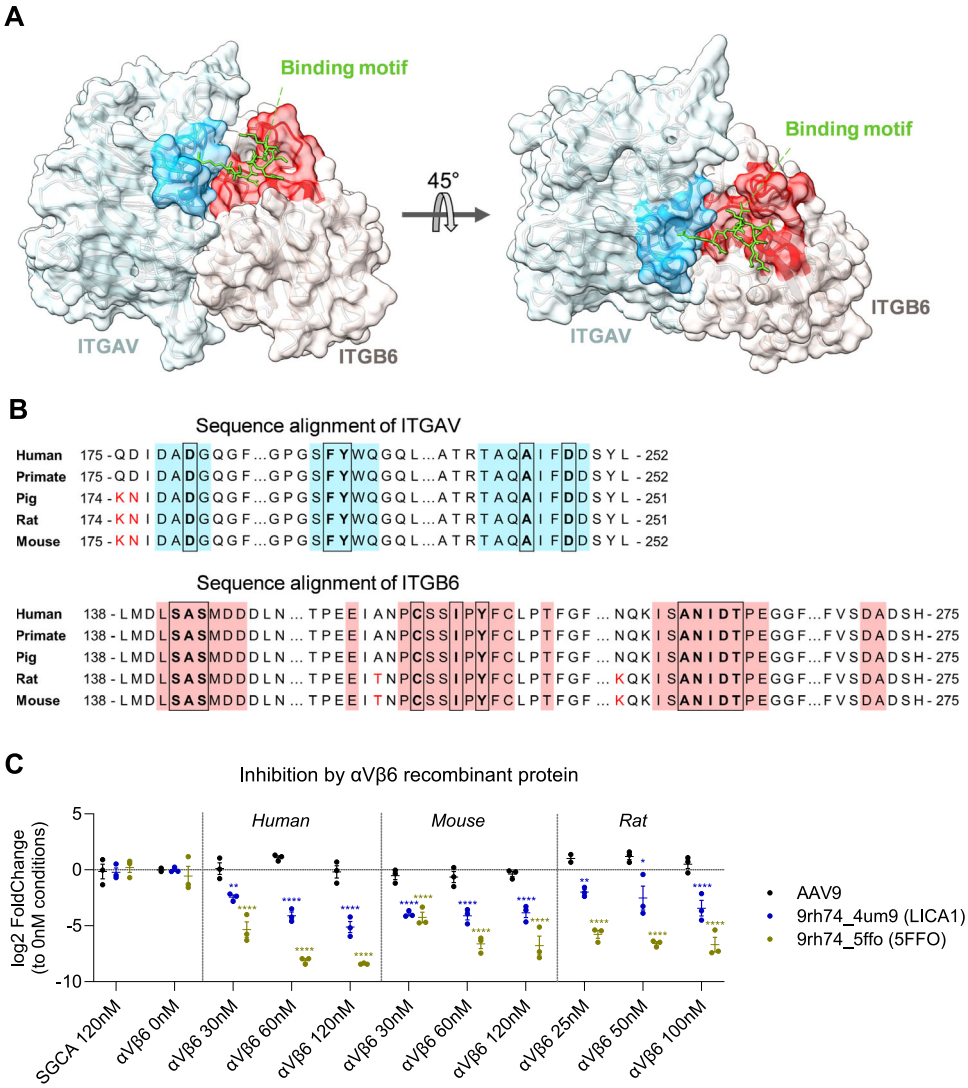


Fig. 7 | LICA1 showed conserved interaction with αVβ6 from multiple species. **A** Structure of human αVβ6 binding to human TGF-β3-derived motif (pdb code: 4um9). The binding interface, defined as all amino acid in αVβ6 with distance to the binding motif of < 8 Å, is highlighted. **B** Aligment of ITGAV (upper panel) and ITGB6 (lower panel) protein sequence around the binding interface from multiple species. The binding interface defined in **A** is highlighted, and is identical across species being examined. Amino acids with distance to the binding motif of < 6 Å, are in bold and boxed. The sequence mismatches, only found outside the binding interface, are colored in red. **C** Transduction efficiency, measured by luciferase activity, of AAV_ITGs but not AAV9 was inhibited by pre-incubating AAVs before infection with recombinant αVβ6 protein from both human, rat, and mouse. 2E8vg AAVs were incubated with different αVβ6 concentrations (0–120 nM) for 1 h at 37 °C before added directly into cell medium (dose: 1E4 vg/cell, 96-well plate, duration: 24 h, cell line: 293_WT, $n = 3$ biological replicates). Same incubation conditions using 120 nM of recombinant SGCA protein were used as the control, which showed no significant difference with 0 nM αVβ6 condition. Data are presented as mean ± SEM. The statistics were performed to compare with the condition of no αVβ6 protein during incubation (0 nM αVβ6) by using two-way ANOVA (-AAV serotypes * αVβ6 treatments) followed by FDR correction. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns not significant. Source data are provided as a Source Data file.

levels to well-produced AAV9. This suggests that low Rosetta energy correlates with high stability of capsid protein, and thereby high AAV production.

Integrin αVβ6 as a myotropic AAV receptor for skeletal muscle across species

Virus-host interaction is the foundation for improved viral vectors, yet skeletal muscle receptors that allow effective AAV transduction are poorly defined. However, top hits from two independent studies with different screening schemes identified myotropic AAVs with a common RGD motif[15,17,19],. In addition, it has previously been described that integrin functions as cellular receptor for natural AAV[23,24]. Coincident with our screening for possible integrin receptor, only αVβ6 is highly expressed and enriched in skeletal muscle (Fig. S1). By including αVβ6 binding motifs, AAV_ITGs efficiently utilized αVβ6 for cellular infection. It is worth-noting that the binding interface of αVβ6 with TGFβ-derived peptides is highly conserved across species (Fig. 7A, B). Enhanced transduction was observed in conditions with high (either ectopic or natural) αVβ6 expression and can be blocked by recombinant αVβ6 from multiple species, suggesting similar efficiency in targeting αVβ6 receptor in these species. Indeed, this improvement in skeletal muscle targeting was validated in various models from multiple species: human differentiated myotubes, murine skeletal muscles of WT and two other MD mouse models, and NHP.

Despite high enrichment in the human skeletal muscle, ITGB6—the subunit with more binding interaction with TGFβ-derived motifs, also shows relatively high enrichment in gastric mucous cells (stomach), urothelial cells (prostate), ductal cells (pancreas), and alveolar

https://doi.org/10.1038/s41467-024-52002-4

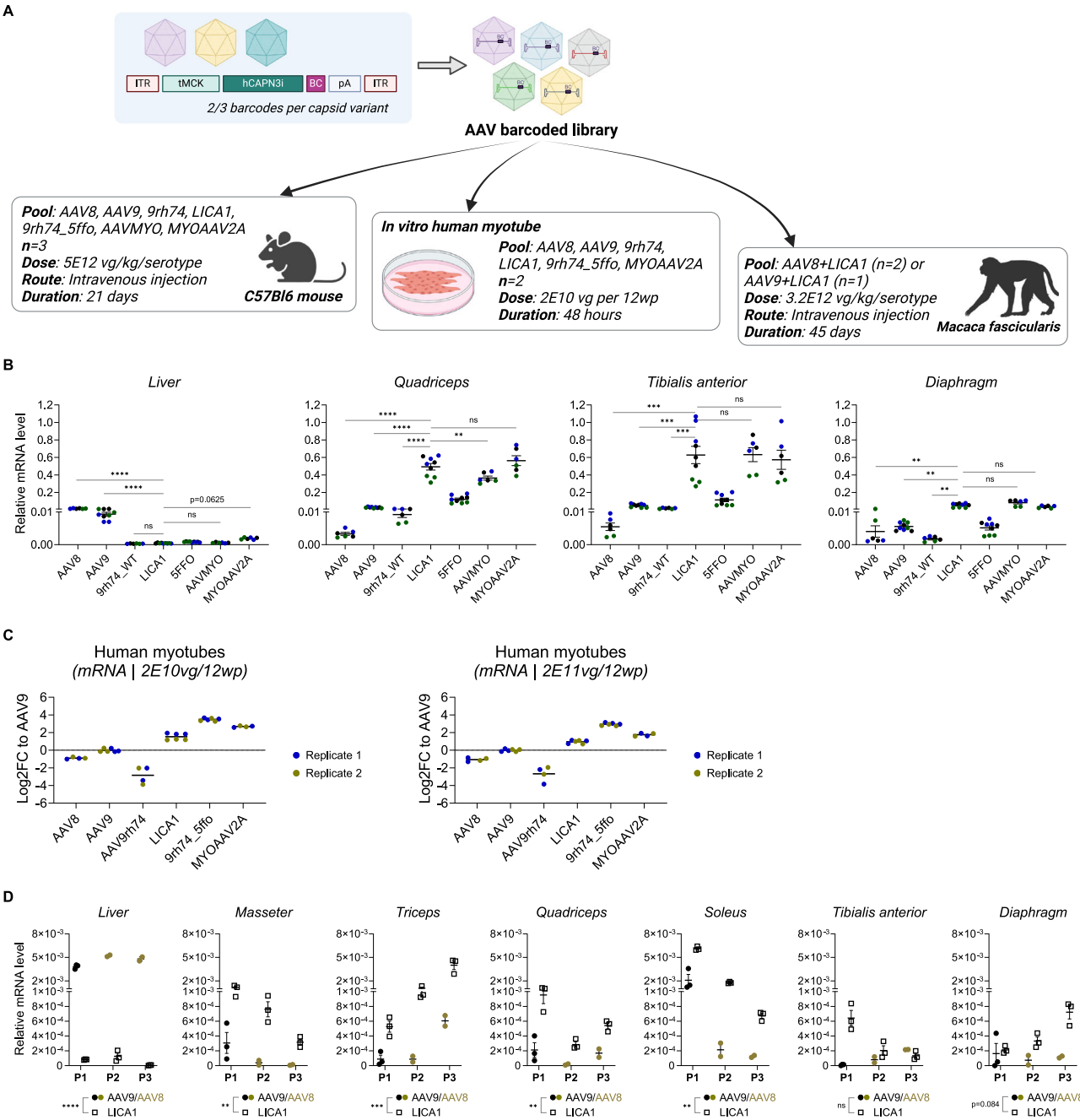


Fig. 8 | A LICA1 showed cross-species enhanced transduction in the skeletal muscle. Scheme of comparative study between multiple AAV capsids in multiple species. Transgene expression cassette is ITR_tMCK_hCAPN3-inactive_BC_SV40pA_ITR. Barcodes (BCs) allow quantitative measurements of transgene expression mRNA level. 2/3 barcodes were used for each capsid to minimize the sequence bias. AAV production of each capsid variant was done separately, before pooled together at the equimolar amount before AAV transduction in vitro (**C**, human myotube, $n = 2$ biological replicates, dose: 2E10/2E11 vg per 12wp well, duration: 48 h), or in vivo injection in C57Bl6 mice (**B**, $n = 3$ mice, dose: 5E12 vg/kg per capsid variant) and *Macaca fascicularis* NHP (**D**, $n = 3$ NHPs, dose: 3.2E12 vg/kg per capsid variant). Created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en). **B** The mRNA enrichment (BC_mRNA/Rplp0/ BC_AAV) of different capsid variants in the liver and three skeletal muscles in C57Bl6 mice measured by RT-qPCR. Each dot correponds to the individual barcode. The color of each barcode corresponds to the individual mouse in Fig. S6C. The statistics were performed on the average on different barcodes (using for the same capsid variant in each mouse) by using two-way ANOVA (- mice * capsids) followed by FDR correction. **C** The mRNA enrichment, measured by NGS, of different capsid variants in in vitro human myotubes at two AAV concentration. The log2FC compared to AAV9 of all variants is presented. Each dot correponds to the individual barcode. The colors correpsond to different biological replicates. **D** The mRNA enrichment (BC_mRNA/Rplp0/BC_AAV) of different capsid variants in the liver and skeletal muscles in NHP measured by RT-qPCR. Each dot correponds to the individual barcode. The data was presented for individual primates (P1-3). The statistics were to compare the mRNA level of LICA1 with natural capsids, AAV8/AAV9, performed on the fold change of LICA1 to AAV8/AAV9 (averages of barcodes using for the same capsid variant in each primate were used to calculate the fold-change) by using two-tailed one-sample *t*-test. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$; ns not significant. Data in (**B**) and (**D**) are presented as mean ± SEM. Source data are provided as a Source Data file.

cells type 2 (AT2, lung), colon enterocytes (colon) and high bulk expression also in kidney, esophagus, and bladder (Fig. S1B, C). There was very low AAV transduction observed in both kidney and lung, two organs with highest ITGB6 expression. Poor kidney and lung transduction despite high expression of targeted receptor in these two organs was also observed with other neurotropic engineered AAVs[52,53]. Route of administration in targeting kidney might explain the low transduction of AAV_ITGs in the kidney where local delivery via the renal vein or the renal pelvis showed much better transduction than IV injection for multiple AAV serotypes[54]. On the other hand, AT2 cells in the lung comprise only <5% of the alveolar surface area[55], therefore they have limited contact with pulmonary capillary blood, which might explain poor transduction efficiency in this organ. Natural AAVs were previously shown to transduce the gastrointestinal tract[56] and pancreas[57], suggesting a potential transduction of AAV_ITGs in these organs. Further validation of these tissues is therefore required in future studies to confirm AAV_ITGs' specificity.

Mechanism of enhanced transduction by αVβ6 binding remains unclear. In most cases, the improved transduction of AAV_ITGs and other mAAVs was evident at the VCN level, indicating better cell entry via αVβ6 binding. In some cases, for example observed in Dia and Heart of dystrophic mice, the improvement is only seen at the transgene expression level, suggesting the role of αVβ6 beyond cellular internalization. Indeed, integrins are constantly trafficked in cells, initially internalized by either clathrin-mediated or clathrin-independent endocytosis to enter the endocytic–exocytic pathway before being recycled[58]. The foot-and-mouth disease virus, which contains αVβ6-binding RGD motif to infect cells, also requires EEA1/Rab5-positive vesicles for intracellular trafficking[59,60], suggesting a similar trafficking mechanism with αVβ6-targeting AAVs. Future studies are necessary to better understand the αVβ6's role in intracellular myotropic AAV trafficking within endocytic system.

In addition, we conducted a study comparing LICA1 and five other published myotropic AAVs. The LICA1's skeletal muscle transduction level was equivalent to that of AAVMYO[15], MYOAAV-1A/−2A[17], and significantly better than AAVMYO2/3[16]. The liver infection rate was significantly lower with LICA1 compared to the AAVMYO, MYOAAV-1A/−2A, and of similar level with AAVMYO2/3. It is presumably due to the use of a liver-detargeted backbone and the low expression level of αVβ6 in liver. As a result, the LICA1 vector exhibited the highest muscle/liver transduction ratio among all AAVs tested, suggesting increased specificity towards skeletal muscle. This finding highlights the importance of selecting an appropriate targeting receptor for rational design and further supports αVβ6 as a promising candidate for targeting skeletal muscle.

LICA1 is a potential vector for muscular diseases

AAV gene therapy in muscle diseases typically requires very high doses (≥1E14 vg/kg) for functional benefits[45,61], yet can result in severe and even fatal adverse events[7]. In this study, we explored low dose (5E12 vg/kg) treatment using the LICA1 vector in two MD mouse models, DMD and LGMDR3. Of note is that this dose is at least 20 times lower than the doses currently used in clinical trials for neuromuscular diseases[3]. In both models, LICA1 was significantly better than AAV9 in delivering and expressing therapeutic transgenes, consequently restoring better histological dystrophic phenotypes. In TA and Qua, LICA1 was able to transduce >80% of fibers. It was still a challenge to effectively transduce diaphragm muscle at this dose, yet more than 50% of Dia fibers were positive for transgene expression with LICA1 in both models while AAV9 transduced very poorly. This improvement in transgene expression translates directly into improved histological restoration, including centronucleation index and fibrosis level. In particular, with only more than 50% successfully transduced fibers, LICA1 was able to reduce diaphragm fibrosis by 42.8–47.0% (mdx and SGCA$^{-/-}$ models respectively), whereas no change was seen in AAV9-treated groups. The biomarker for muscle damage level, MYOM3, was reduced by 57.5–67.2% (mdx and SGCA$^{-/-}$ models respectively) by LICA1 and significantly greater than AAV9. Similarly, global muscle force was significantly restored to a higher level with LICA1 than with AAV9 in SGCA-KO mice. These data confirmed superior muscle transduction by LICA1 and resulting therapeutic benefits were obtained even at low-dose treatment in two MD models.

However, treatment efficacy varies between two disease models at molecular levels. We profiled transcriptomic changes in Qua following AAV treatment in both MD models. Despite similar transduction efficiency of LICA1 in the two models, restoration of dystrophic transcriptional changes in SGCA-KO was significantly greater. It is noteworthy that μDys is an incomplete form of dystrophin. The μDys used in the present study lacks several functional domains, including multiple spectrin-like repeats that bind to nNOS, F-actin, sarcomeric lipid and microtubules, and a dystrobrevin- and syntrophin-binding C-terminus[45]. This might explain the inadequate efficacy in restoring global gene expression in μDys gene therapy trials, in spite of highly effective gene transfer. Similarly, despite excellent functional restoration by microdystrophin gene transfer in various animal models, outcomes from these clinical trials are unsatisfactory[6]. Therefore, careful assessment of molecular restoration should be included for evaluating gene therapy efficacy.

In summary, this study presents an alternative computational method that aids rational AAV design and ensures high-production AAV variants. The proof-of-concept design targeting skeletal muscle resulted in a high-productivity cross-species myotropic AAV, thereby effectively delivering therapeutic transgenes and restoring dystrophic phenotypes in two MD mouse models at a low dose. This work contributes to the ongoing efforts to reduce AAV treatment doses and further advance AAV engineering, paving the way for more effective and accessible gene therapies in the future.

Methods

Animal care and use

All animals were handled according to French and European guidelines for human care and the use of experimental animals.

All procedures on mice were performed at the Centre d'exploration et de recherche fonctionnelle expérimentale (CERFE; Evry) and approved by the local ethics committee and the regulatory affairs of the French Ministry of Research (MESRI) under the numbers 2018-024-B #19736, 2022-004 #35896. C57Bl/6, B6Ros.Cg-Dmdmdx-4Cv/J mice were obtained from the Jackson Laboratory. A knockout mouse model of *α*-sarcoglycan (C57Bl/6 background) was obtained from the Kevin Campbell laboratory (University of Iowa, USA)[48]. Mice were housed in an SPF barrier facility with 12 h/12 h light/dark cycles, a temperature of 20–23 °C, and humidity of ≥ 40%. Mice were provided with food according to the manufacturer's instructions (ref: SAFE 150-SP-25) and water *ad libitum*. Only male mice were used in the present study. Well-being and weights of the animals were monitored for the duration of the study. The animals were anesthetized with a mix of ketamine (100 mg/kg) and xylazine (10 mg/kg), or with isoflurane (4%) for blood samples. For AAV intravenous injections, a maximum volume of 150 μl containing AAV vectors was injected via the sinus route after the animals had been anesthetized with isoflurane. The AAV intravenous doses used in the present study were 5E12 or 1E13 vg/kg. At the end of the studies, mice were euthanized by physical cervical dislocation and animals' death was confirmed before organ collection.

The NHP study was reviewed by the Animal Welfare Body of Cynbiose and the Ethics Committee of VetAgro-Sup, France and approved under number 2305 (MESRI number: 2023010517363397). Two females and one male Cynomolgus monkeys (*Macaca fascicularis*) of 29–34 month-old were used in this study. NHPs were first screened for their Nab levels against multiple AAV capsids (Table S3). No NHP was Nab-negative for all serotypes tested. All NHPs were housed in

isosexual social groups during the acclimation period of 3 weeks and the in vivo experimental phase. NHPs were exposed to a 12-h light:dark cycle, controlled temperature of 22 ± 2 °C, offered *ad libitum* access to tap water and fed daily with Specific Primate Diet (Ssniff ref: V3944-000) in appropriate amounts fitting the animals' size (100 g for animals under 5 kg), supplemented with fresh fruits and vegetables as part of their behavioral enrichment program. A pool of AAVs (AAV8 + LICA1 or AAV9 + LICA1) were injected intravenously through the saphenous veins at the total dose of 1.3E13 vg/kg (0.1625 mL/kg). On Day 46, after at least 8 h of fasting, all animals were sedated with ketamine (5–10 mg/kg) and midazolam (0.5 mg/kg), then euthanized by a lethal intravenous injection of a veterinary euthanizing agent, and their deaths were confirmed. Following euthanasia, the animals underwent dissection, exsanguination, and a 0.9% NaCl flush via the heart's left ventricle, before organ collection for tissue analysis.

Cell culture and in vitro study

Adherent HEK293-T (293_WT, ATCC Ref CRL-11268) cells and other derivatives (293_αVβ6, 293_α7β1, 293_AAVR-KO) were maintained in the proliferating medium containing DMEM (Thermo Fisher Scientific), supplied with 10% fetal bovine serum and 1X gentamycin at 37 °C, 5% CO2. Human immortalized myoblasts (AB1190 cell line) were maintained in Skeletal Muscle Cell Growth Medium (PromoCell, C23060) and differentiated in Skeletal Muscle Differentiation Medium (PromoCell, C23061).

In vitro AAV infection was performed by directly adding AAV into culture medium at the dose of 1E9 or 1E10 vg per 24-well plate well. After 48 h post-infection, cells were washed and subjected to VCN and gene expression analysis. To inhibit AAV infection, AAVs were incubated with recombinant hITGAV-hITGB6 protein (Bio-Techne, 3817-AV-050) at 37 °C for 30 min, at a concentration of 1 μg protein per 5E9vg AAV before addition to the cells (1E4 vg per cell). The same condition treated with recombinant hSGCA protein served as a control for the comparison.

Structural prediction using AlphaFold2 and backbone relaxation using the Rosetta Relax protocol

The predicted 3D prediction of the hybrid capsid Cap9rh74 (aa 219–737) was obtained with LocalColabFold (https://github.com/YoshitakaMo/localcolabfold.git)[35]. The weight used for monomer prediction was *alphafold2_ptm*, the prediction was recycled 3 times, template was used. The highest pLDDT unrelaxed structure was then relaxed using the Rosetta Relax protocol[62]. The lowest total estimated energy structure was then used for next design step. The arguments used for relaxation protocol include use_input_sc, constrain_relax_to_start_coords, ignore_unrecognized_res, relax = fast, nstruct = 50.

Extraction of αVβ6-binding motifs

Peptides at the binding interface of TGF-β1 (PDB_id: 5FFO) or TGF-β3 (PDB_id: 4UM9) with αVβ6 were isolated to be grafted into relaxed Cap9rh74. Isolated peptides both have the XRGDLXXL/I motif and bind to αVβ6 at a very similar position (Fig. S2E). The minimized number of amino acids with αVβ6 contact was selected to minimize the size of the capsid VR loop after modification, as well as to facilitate the grafting step (a bigger motif will be more difficult and computationally expensive for sequence sampling).

Grafting αVβ6-binding motifs into the Cap9rh74 capsid

The predicted VR-IV loop of Cap9rh74, aa 453–459 (corresponding to aa 235–241 in PDB numbering), was substituted with a novel loop incorporating αVβ6-binding motifs using the RosettaRemodel protocol for motif-grafting (Fig. 1A, and Fig. S1F)[38]. The redesigned loops incorporated the binding motif and were flanked by two-amino acid linkers on either end. During the design process, the conformation of the binding motif was preserved, while the sequences and structures of the linkers were sampled. The length of the linker was selected to balance computational efficiency and flexibility in structural grafting. The lowest total estimated energy structure was subsequently subjected to experimental validation and retrospective docking against αVβ6. The arguments used for motif-grafting includes: ex1, ex2, use_input_sc, num_trajectory = 300, use_clusters = false, find_neighbors.

Retrospective docking of the designed capsid to αVβ6

Prior to docking, the top-scoring capsid structures obtained from the Remodel step were subjected to relaxation using the Rosetta Relax protocol. Additionally, the αVβ6 structure (PDB_id: 4UM9, chain: A-B) was repacked using the PackRotamersMover in Rosetta. The initial docking position was established by positioning the designed capsid structure such that its new VR4 region was in close proximity to the natural binding motif, thereby reducing the search space for potential binding sites. Custom Rosetta scripts were employed for sampling and analyzing the docking positions. Structures were visualized using UCSF ChimeraX 1.2.5.

AAV production

All AAVs in the present study were produced in HEK293-T cells cultured in suspension using the triple transfection protocol. After 3 days, cells were chemically lysed with Triton-X100 (dilution 1/200, 2.5 h, 37 °C) and lysates were clarified by centrifugation (2000g, 15 min, 4 °C). AAV was precipitated from lysates by adding 40% PEG-8000 at the final concentration of 8% and incubated overnight. The mixture was spun at 3500 g for 30 min at 4 °C, the pellet was collected and resuspended in 20 ml TMS buffer. The material was then subjected to cesium chloride gradient purification, concentrated using an Amicon spin concentrator (Milipore), and stored at −80 °C. The AAV titer was determined by qPCR after AAV DNA was extracted in triplicate using a MagNA Pure 96. Primers used for titration are listed in Table S1.

Creation of stable overexpression cell lines using the PiggyBac system

The HEK293T cells were transfected with a mix of 3 plasmids at the molecular ratio of 2:2:1 using Lipofectamine 3000. These plasmids include pITR_CBh_hITGAV_CMV_Puro, containing expression cassettes of human *Itgav* genes, and puromycin-resistant gene, pITR_CBh_hITGB6_CMV_Bsd, containing expression cassettes of human *Itgb6* genes and blasticidin-resistant gene, and transposase-containing plasmid (Fig. S2B). Antibiotic selection of cells with double integration was carried out 48 h post-transfection by adding 2 μg/ml of puromycin (Gibco, 12122530) and 5 μg/ml of blasticidin (Sigma, 15205). The cells were then subjected to a second round of antibiotic selection 5 days after transfection. Western blot confirmed the overexpression of 2 proteins (Fig. S2C). A similar approach was used to create the 293_α7β1 cell line, where *Itga7* and *Itgb1* are stably overexpressed in the 293 cell.

Creation of AAVR-KO cell lines

Three sgRNAs (5′-AGCAAGUAUAAAACAGGUAC-3′, 5′-GGAUUUUAUCAGGAUAUUAU-3′, 5′-UCUGGAGGGCCAUGGAGAAG-3′) targeting exon 2 of the human KIAA0319L genes were incubated with SpCas9 protein at the sgRNA/Cas9 ratio of 1.5/1 for 10 min, before transfected into HEK293T cells using Immortalized Cell Lipofection protocol (Synthego). After 24 h, serial dilutions were done to isolate the single clones. After 2-week clonal expansion, the single clones were phenotyped by western blot, and the clone with complete absence of AAVR protein at expected size was subsequently used in the AAV transduction experiment.

Quantification of the MYOM3 protein fragment in the serum

Serum MYOM3 quantification was performed by a custom sandwich ELISA assay. First, a polyclonal MYOM3 antibody (Proteintech, 17692-1-

AP, dilution: 1/10) was coated on a 96-well plate and incubated overnight at 4 °C. After washing three times with PBST and saturated with saturation solution (3% BSA in PBS), dilutions of serum samples were added and incubated for 2 h at room temperature. Next, a monoclonal antibody (Proteogenix, REF: 51-H1-B4) coupled with SULFO-TAG (MSD) was used for detection by incubating for 2 h at room temperature. After being washed three times with wash buffer (0.05% Tween-20 in PBS). MESO Quickplex SQ 120 (MSD) was then used to measure the absorbance of SULFO-TAG. A set of concentration-defined MYOM3 peptides (His-tagged, Proteogenix) was included in the same experiment and subsequently used to calculate the MYOM3 concentration in the serum samples.

Quantification of anti-capsid IgG in the serum

Concentration of IgG against AAV capsids was measured with an ELISA assay. A 96-well plate was coated with AAV to a final concentration of 1 μg/ml. A standard curve of purified mouse IgG was included in parallel. After blocking, multiple dilutions of serum samples were added in duplicate and incubated for 1 h at 37 °C. A horseradish peroxidase-conjugated monoclonal antibody specific for mouse IgM (Southern Biotech) was added to the plates and incubated for 1 h at 37 °C. A substrate solution containing 3,3′,5,5′-tetramethylbenzidine (Becton Dickinson, REF: 555214) was then added to the reaction before it was stopped with H2SO4 3 M solution. Measurements were performed at 450 nm using an ENSPIRE microplate reader (PerkinElmer). Anti-AAV IgG concentration was determined using the standard curve.

Quantification of neutralizing AAV antibodies in the non-human primate serum

The in vitro AAV vector neutralizing antibody assay was performed as previously described[63]. The 293_E4 cells were seeded at 2×10^4 cells/100 μl/96-wp well, supplemented with Ponasterone A (Invitrogen, REF: 450478) at 1 μg/ml concentration in DMEM + 10%FBS, and incubated at 37 °C for 24 h. In the next day, the serum samples from non-human primates were first heat-inactivated for 30 min at 56 °C, and serially diluted in the factor of 3.16. 20 μl of diluted serum was mixed with 20 μl 2×10^9 vg/ml AAVs, which express GFP/Luc fusion transgene as in Fig. 3, and incubated for 1 h at 37 °C before added into 293_E4 cells. The condition incubated with 20 μl FBS was used as reference for the inhibition calculation (100% transduction). After 24 h, luciferase activities were measured. The neutralizing titer, defined as the highest serum dilution reducing luciferase activity by 50% compared to the reference condition.

DNA extraction and quantification of vector copy number

Total DNA was extracted from frozen muscles or cells using a KingFisher (Thermo Fisher Scientific). Absolute quantification of viral genome (DNA_{AAV}) was performed using qPCR on the LightCycler480 system (Roche) using Taqman Gene Expression Assays, in parallel with known DNA concentration standards. Genomic DNA (DNA_g) was measured via the Rplp0 gene. Vector copy number (VCN) was determined as the number of viral genomes per nucleus by normalizing DNA_{AAV} to twice the DNA_g copies.

RNA extraction and expression analysis

Total RNA was extracted from frozen muscles or cells using an IDEAL32 extraction robot. DNA contamination was removed from the RNA samples with the TURBO DNA-free kit (Thermo Fisher Scientific). Gene expression was measured with 1000 ng of total RNA, reverse-transcribed with a mixture of random oligonucleotides and oligo-dT and the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific), and quantified by qPCR on the LightCycler 480 system (Roche) using Taqman Gene Expression Assays (Thermo Fisher Scientific). Each PCR reaction was performed in duplicate, and data was normalized across samples using Rplp0.

For multiplex quantification using RT-qPCR, similarly, total RNA from different tissues was extracted, treated with DNase, and reverse-transcribed using both random hexamers and oligo-dT. The primers used for qPCR (barcode primers) includes a common forward primer in the hC3i ORF and unique reverse primers covering the barcodes (Table S1, length = 19–22 nt, Tm = 56–59 °C). The *Ct* values of barcode primer were adjusted to *Ct* values of primers in the common SV40 polyA, by the linear regression built from qPCR result of a series of barcode plasmids' dilution. Similar amplifications were observed between all barcode primers. The adjusted values were then normalized across samples to Rplp0 using $2^{\Delta Ct}$ method and the barcode ratio for given capsid (1/3 for capsids produced with 3 barcodes and 1/2 for capsids produced with 2 barcodes).

For multiplex quantification using NGS, cDNAs were obtained as described above. The cDNAs were then PCR-amplified using Q5 High-Fidelity 2X Master Mix (NEB M0492) for 25 cycles using primers upstream and downstream of barcode sequences. A series of synthesized DNA oligoes with known sequences and known concentrations was spiked into cDNA samples before the PCR, to assure the linearity after PCR amplification and NGS library preparation. The sequencing libraries were prepared following the Illumina protocol, and sequenced on MiSeq instrument (Genewiz), resulting in ~60,000 2 x 250bp reads per sample. A custom Python code using biopython package (v1.81) was used to count high-quality reads (Phred score ≥ 25) for all barcodes and spike-in DNAs. The barcode counts remained in the linear phage, according to the spike-in DNA counts. The barcode data was then normalized to the barcode ratio for given capsid (1/3 for capsids produced with 3 barcodes and 1/2 for capsids produced with 2 barcodes).

RNA sequencing and transcriptomic analysis

The RNA quality of samples was initially checked using the Bioanalyzer 2100 (Agilent). The samples that had an RNA integrity number > 8 were then used for RNA sequencing (Genewiz). The Stranded Total RNA Library Prep Kit (Illumina) was used to create the sequencing libraries, which were sequenced on the NovaSeq instrument (Illumina) following the Illumina protocol, resulting in ~20 million paired-end reads per library. The paired-end reads were filtered and subjected to quality control using fastp[64]. They were then mapped onto the GRCm38/mm10 genome using HISAT2 (v2.2.1)[65] and count tables were generated using htseq-count (2.0.4)[66]. Differentially expressed genes (DEGs) were identified using the DESeq2 (v1.38.3) R package[67]. Pathway analysis was carried out in R-Studio (v4.2.2) using either over-representation methods with Gene Ontology or functional class scoring with Gene Set Enrichment Analysis.

Analysis of single-nucleus RNA sequencing

The single-nucleus RNA sequencing dataset was reanalyzed[29]. Raw read counts, barcodes, and features (gene names) were retrieved from Gene Expression Omnibus with accession number GSE156498. The same parameters of filtering, normalization and clustering in the original publication were used, except analysis was carried out with Python packages scanpy (1.9.4). Comparison between the gene expression of WT and D51 groups was done by Wilcoxon Rank Sum test using Python scipy (v1.11.2) package. Python packages seaborn (v0.12.2) and matplotlib (3.7.2) were used for visualization.

Western blotting

Proteins were extracted from tissues or cells with RIPA buffer (Thermo Fisher Scientific) supplemented with Protease Inhibitor Cocktails (Roche) and Benzonase (1/1000, Millipore). The protein concentration was measured with the Pierce BCA Protein Assay (Thermo Fisher Scientific). Equal amounts of protein were loaded and separated on a precast 4–12% Bis-Tris polyacrylamide gel (Thermo Fisher Scientific) before being transferred to a nitrocellulose membrane with the iBlot2

Table 1 | Data collection and refinement statistics

	αVβ6 + TGF-β3 peptide (PDB: 4UM9)[37]	αVβ6 + proTGF-β1 (PDB: 5FFO)[36]
Data collection		
Space group	C2	P1
Cell dimensions		
a, *b*, *c* (Å)	184.4, 170.0, 102.4	81.4, 91.4, 131.0
α, β, γ (°)	90, 98.7, 90	90.0, 86.3, 89.9
Resolution (Å)	50.0–2.60 (2.69–2.60)*	50.0–3.50 (3.64–3.50)*
R_{merge}	12.9 (198)	16.8 (178.8)
I / σI	4.9 (0.5)	4.6 (0.4)
Completeness (%)	96.4 (96.6)	94.6 (97.1)
Redundancy	2.4 (2.4)	1.8 (1.8)
Refinement		
Resolution (Å)	50.0–2.60	50.0-3.50
No. reflections	91,038 (9,101)	45,279 (3,450)
R_{work} / R_{free}	21.4 (37.1) / 25.9 (37.3)	23.2 (30.2)/27.9 (33.9)
No. atoms		
Protein	16,626	22945
Ligand/ion	166/14	936
Water	230	5
B-factors		
Protein	76.0	216.70
Ligand/ion	84.7/71.9	248.40
Water	51.1	142.30
R.m.s. deviations		
Bond lengths (Å)	0.009	0.007
Bond angles (°)	1.3	0.8

*Values in parentheses are for highest-resolution shell.

Dry Blotting system (Thermo Fisher Scientific). The membrane was blocked with Odyssey Blocking Buffer (LI-COR) for 1 h at room temperature and incubated with primary antibodies (diluted in 50% Odyssey Blocking Buffer) overnight at 4 °C. After washing with TBST, the membrane was incubated with secondary antibodies (1/5000 diluted in 50% Odyssey Blocking Buffer) for 1 h at room temperature. The membrane was then washed again and the blotting signals were acquired with the Odyssey Infrared Imaging system. Total protein was used as a loading control for quantification (LI-COR).

Measurement of luciferase activity

Cells or tissues were homogenized in a lysis buffer (25 mM Tris-phosphate, 15% glycerol, 1 mM DTT, 1 mM EDTA, 8 mM MgCl2, 0.2% Triton X-100, protease inhibitor), followed by 3 freezing/thawing cycles and clarification by centrifugation (5 min, 10,000 g, 4 °C). Luciferase activity in homogenates was measured as a luminescence signal after addition of D-luciferin substrate by using EnSpire Plate Reader (PerkinElmer). Protein quantities measured by BCA assay in the same homogenates were used as normalization for comparison across samples.

Immunofluorescence

Dissected tissues were frozen in isopentane in liquid nitrogen. Transverse 8 μm cryosections of the muscles were air-dried and stored at −80 °C. Slices were rehydrated for 5 min in PBS, fixed for 5 min in 3.7% ice-cold methanol-free PFA, permeabilized in 0.1% Triton-X100 for 5 min, and blocked for 30 min in 10% goat serum in PBS (blocking buffer) and for an additional 30 min in Mouse-on-Mouse IgG Blocking Solution (fixation and permeabilization steps were only applied for dystrophin staining). Sections were incubated overnight at 4 °C with primary antibodies diluted in 10% blocking buffer. Samples were washed twice in PBS and incubated for 1 h at room temperature (RT) with Alexa-conjugated secondary antibodies (dilution 1/1000) in 10% blocking solution in a dark humid chamber. Samples were washed three times in PBS and mounted in DAPI Fluoromount-G (Southern Biotech, ref: 0100-20). Images were digitized using an Axioscan Z1 slide scanner (Zeiss) under a Plan-Apochromat 10×/0.45 M27 dry objective (Zeiss) and using a digital CMOS camera ORCA-Flash4.0 (Hamamatsu). The size of a pixel from scanned images is 0.65 μm. Tile scan images were reconstructed with ZEN software (Zeiss).

Quantification of fibrosis in muscle transverse sections

Sirius Red stained transverse sections were used for fibrosis quantification using QuPath (v0.3.2). For each muscle scan, a small artificial neural network was trained to classify positive and negative fibrotic pixels, and subsequently used to quantify fibrotic areas.

Myofiber segmentation

Muscle sections were stained with laminin or wheat germ agglutinin (WGA) for membrane labeling. The Cellpose2 cyto2 model[68] was fine-tuned on manually myofiber-labeled images based on laminin or WGA staining (hyperparameters: n_epochs = 200, learning_rate = 0.05, weight_decay = 0.0001). The labeled dataset used in fine-tuning was prepared in such a way that the model can simultaneously segment myofibers and ignore low-quality staining areas. Fine-tuned models were then used to extract myofiber masks. Reconstruction of myofiber masks of whole-scan images was done using the cellpose package[68]. Reconstructed masks were then converted into Regions of Interest (ROIs) for subsequent quantification (each ROI corresponds to an individual myofiber) using the *Labels_To_Rois.py* FIJI plugin[69].

Quantification of centronucleated myofibers

Muscle sections were co-labeled with laminin/WGA for membrane and DAPI for nucleus. ROIs for each myofiber were generated as described and utilized for the detection and quantification of centronucleated myofibers using a custom FIJI macro. Briefly, centronucleated myofibers were identified as ROIs that are positive for DAPI staining after exclusion of all nuclei that overlapped with the myofiber membrane.

Quantification of SGCA-positive or μDys-positive myofibers

Muscle sections were co-labeled with laminin/WGA for membrane and SGCA/dystrophin. Similar to above, ROIs were generated based on membrane labeling and used for subsequent quantification of the SGCA/dystrophin signal using FIJI macro.

Statistical analysis

Results are presented as mean ± SEM, unless noted otherwise. Significance of differences in multiple pairwise comparisons of more than two groups was determined by one-way ANOVA. The significance of differences in pairwise comparisons of multiple groups with multiple treatments was determined by two-way ANOVA. To account for multiple testing, Benjamini-Hochberg (BH) procedure with an FDR threshold of 0.05 to control the false discovery rate (FDR) or post hoc tests were used. The significance of differences using fold change values was determined by two-tailed one-sample *t*-test with hypothetical value of 1. Statistical tests were performed using GraphPad Prism 9.5.1, R 4.2.2, or Python 3.10. Results were considered significant when *p*-values or adjusted *p*-values were < 0.05.

Reporting summary

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

Data availability

All data associated with this study are present in the paper or the Supplementary Materials and Source data file. The transcriptomic data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) database under accession code PRJNA1039498. Processed data including differential gene expression analysis are provided in the Supplementary Data 1, 2. The DNA and amino acid sequences of AAV capsid variants generated in this study are provided in the Supplementary Note. Crystal structures of αVβ6 and its binding motifs used in this study are re-used from the Protein Data Bank with accession codes 4UM9[37] and 5FFO[36] (Table 1). The single-nucleus RNAseq dataset of TA muscles from WT and Dmd Exon 51 Knockout mice is re-used from the GEO database under accession code GSE156498. The plasmid constructs and reagents generated as part of this study are available under a material transfer agreement from the corresponding authors. Source data are provided with this paper.

References

1. Wang, D., Tai, P. W. L. & Gao, G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.* **18**, 358–378 (2019).
2. Pupo, A. et al. AAV vectors: the rubik's cube of human gene therapy. *Mol. Ther. J. Am. Soc. Gene Ther.* **30**, 3515–3541 (2022).
3. Crudele, J. M. & Chamberlain, J. S. AAV-based gene therapies for the muscular dystrophies. *Hum. Mol. Genet* **28**, R102–R107 (2019).
4. Duan, D. Systemic AAV micro-dystrophin gene therapy for duchenne muscular dystrophy. *Mol. Ther. J. Am. Soc. Gene Ther.* **26**, 2337–2356 (2018).
5. Mack, D. L. et al. Systemic AAV8-mediated gene therapy drives whole-body correction of myotubular myopathy in dogs. *Mol. Ther. J. Am. Soc. Gene Ther.* **25**, 839–854 (2017).
6. Mercuri, E., Bonnemann, C. G. & Muntoni, F. Muscular dystrophies. *Lancet* **394**, 2025–2038 (2019).
7. Ertl, H. C. J. Immunogenicity and toxicity of AAV gene therapy. *Front. Immunol.* **13**, 975803 (2022).
8. Verdera, H. C., Kuranda, K. & Mingozzi, F. AAV vector immunogenicity in humans: a long journey to successful gene transfer. *Mol. Ther. J. Am. Soc. Gene Ther.* **28**, 723–746 (2020).
9. High-dose AAV gene therapy deaths. *Nat. Biotechnol.* **38**, 910 (2020).
10. Ogden, P. J., Kelsic, E. D., Sinai, S. & Church, G. M. Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science* **366**, 1139–1143 (2019).
11. DiMattia, M. A. et al. Structural insight into the unique properties of adeno-associated virus serotype 9. *J. Virol.* **86**, 6947–6958 (2012).
12. Li, C. & Samulski, R. J. Engineering adeno-associated virus vectors for gene therapy. *Nat. Rev. Genet.* **21**, 255–272 (2020).
13. Tseng, Y. S. & Agbandje-McKenna, M. Mapping the AAV capsid host antibody response toward the development of second generation gene delivery vectors. *Front. Immunol.* **5**, 9 (2014).
14. Buning, H. & Srivastava, A. Capsid Modifications for targeting and improving the efficacy of AAV vectors. *Mol. Ther. Methods Clin. Dev.* **12**, 248–265 (2019).
15. Weinmann, J. et al. Identification of a myotropic AAV by massively parallel in vivo evaluation of barcoded capsid variants. *Nat. Commun.* **11**, 5432 (2020).
16. El Andari, J. et al. Semirational bioengineering of AAV vectors with increased potency and specificity for systemic gene therapy of muscle disorders. *Sci. Adv.* **8**, eabn4704 (2022).
17. 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 (2021).
18. Ruoslahti, E. & Pierschbacher, M. D. Arg-Gly-Asp: a versatile cell recognition signal. *Cell* **44**, 517–518 (1986).
19. Bauer, A. et al. Molecular signature of astrocytes for gene delivery by the synthetic adeno-Aassociated viral vector rAAV9P1. *Adv. Sci. (Weinh)* **9**, e2104979 (2022).
20. Zolotukhin, S., Trivedi, P. D., Corti, M. & Byrne, B. J. Scratching the surface of RGD-directed AAV capsid engineering. *Mol. Ther. J. Am. Soc. Gene Ther.* **29**, 3099–3100 (2021).
21. Hynes, R. O. Integrins: A family of cell surface receptors. *Cell* **48**, 549–554 (1987).
22. Hussein, H. A. et al. Beyond RGD: virus interactions with integrins. *Arch. Virol.* **160**, 2669–2681 (2015).
23. Asokan, A., Hamra, J. B., Govindasamy, L., Agbandje-McKenna, M. & Samulski, R. J. Adeno-associated virus type 2 contains an integrin alpha5beta1 binding domain essential for viral cell entry. *J. Virol.* **80**, 8961–8969 (2006).
24. Summerford, C., Bartlett, J. S. & Samulski, R. J. AlphaVbeta5 integrin: a co-receptor for adeno-associated virus type 2 infection. *Nat. Med.* **5**, 78–82 (1999).
25. Stewart, P. L. & Nemerow, G. R. Cell integrins: commonly used receptors for diverse viral pathogens. *Trends Microbiol.* **15**, 500–507 (2007).
26. Strauch, E. M. et al. Computational design of trimeric influenza-neutralizing proteins targeting the hemagglutinin receptor binding site. *Nat. Biotechnol.* **35**, 667–671 (2017).
27. Cao, L. et al. Design of protein-binding proteins from the target structure alone. *Nature* **605**, 551–560 (2022).
28. Ruoslahti, E. RGD and other recognition sequences for integrins. *Annu. Rev. Dev. Biol.* **12**, 697–715 (1996).
29. Chemello, F. et al. Degenerative and regenerative pathways underlying Duchenne muscular dystrophy revealed by single-nucleus RNA sequencing. *Proc. Natl Acad. Sci. USA.* **117**, 29691–29701 (2020).
30. Pang, X. et al. Targeting integrin pathways: mechanisms and advances in therapy. *Signal Transduct. Target Ther.* **8**, 1 (2023).
31. Dusart, P. et al. A tissue centric atlas of cell type transcriptome enrichment signatures. *bioRxiv* https://doi.org/10.1101/2023.01.10.520698 (2023).
32. Shen, X., Storm, T. & Kay, M. A. Characterization of the relationship of AAV capsid domain swapping to liver transduction efficiency. *Mol. Ther. J. Am. Soc. Gene Ther.* **15**, 1955–1962 (2007).
33. Ballon, D. J. et al. Quantitative whole-body imaging of I-124-labeled adeno-associated viral vector biodistribution in nonhuman primates. *Hum. Gene Ther.* **31**, 1237–1259 (2020).
34. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
35. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
36. Dong, X. et al. Force interacts with macromolecular structure in activation of TGF-beta. *Nature* **542**, 55–59 (2017).
37. Dong, X., Hudson, N. E., Lu, C. & Springer, T. A. Structural determinants of integrin beta-subunit specificity for latent TGF-beta. *Nat. Struct. Mol. Biol.* **21**, 1091–1096 (2014).
38. Huang, P. S. et al. RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS One* **6**, e24109 (2011).
39. Alford, R. F. et al. The rosetta all-atom energy function for macromolecular modeling and design. *J. Chem. Theor. Comput.* **13**, 3031–3048 (2017).
40. Pillay, S. et al. An essential receptor for adeno-associated virus infection. *Nature* **530**, 108–112 (2016).
41. Zhang, R. et al. Divergent engagements between adeno-associated viruses with their cellular receptor AAVR. *Nat. Commun.* **10**, 3760 (2019).
42. Xu, G. et al. Structural basis for the neurotropic AAV9 and the engineered AAVPHP.eB recognition with cellular receptors. *Mol. Ther. Methods Clin. Dev.* **26**, 52–60 (2022).

43. Duan, D., Goemans, N., Takeda, S., Mercuri, E. & Aartsma-Rus, A. Duchenne muscular dystrophy. *Nat. Rev. Dis. Primers* **7**, 13 (2021).
44. Stedman, H. H. et al. The mdx mouse diaphragm reproduces the degenerative changes of Duchenne muscular dystrophy. *Nature* **352**, 536–539 (1991).
45. Bourg, N. et al. Co-administration of simvastatin does not potentiate the benefit of gene therapy in the mdx mouse model for duchenne muscular dystrophy. *Int. J. Mol. Sci.* **23**, 2016 (2022).
46. Rouillon, J. et al. Serum proteomic profiling reveals fragments of MYOM3 as potential biomarkers for monitoring the outcome of therapeutic interventions in muscular dystrophies. *Hum. Mol. Genet.* **24**, 4916–4932 (2015).
47. Eymard, B. et al. Primary adhalinopathy (alpha-sarcoglycanopathy): clinical, pathologic, and genetic correlation in 20 patients with autosomal recessive muscular dystrophy. *Neurology* **48**, 1227–1234 (1997).
48. Duclos, F. et al. Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. *J. Cell Biol.* **142**, 1461–1471 (1998).
49. Wang, B. et al. Construction and analysis of compact muscle-specific promoters for AAV vectors. *Gene Ther.* **15**, 1489–1499 (2008).
50. de Alencastro, G. et al. Tracking adeno-associated virus capsid evolution by high-throughput sequencing. *Hum. Gene Ther.* **31**, 553–564 (2020).
51. Bryant, D. H. et al. Deep diversification of an AAV capsid protein by machine learning. *Nat. Biotechnol.* **39**, 691–696 (2021).
52. Shay, T. F. et al. Primate-conserved carbonic anhydrase IV and murine-restricted LY6C1 enable blood-brain barrier crossing by engineered viral vectors. *Sci. Adv.* **9**, eadg6618 (2023).
53. Deverman, B. E. et al. Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat. Biotechnol.* **34**, 204–209 (2016).
54. Furusho, T. et al. Enhancing gene transfer to renal tubules and podocytes by context-dependent selection of AAV capsids. *bioRxiv* https://doi.org/10.1101/2023.07.28.548760 (2023).
55. Han, S., Budinger, G. R. S. & Gottardi, C. J. Alveolar epithelial regeneration in the aging lung. *J. Clin. Invest.* **133**, e170504 (2023).
56. Gombash, S. E. et al. Systemic gene delivery transduces the enteric nervous system of guinea pigs and cynomolgus macaques. *Gene Ther.* **24**, 640–648 (2017).
57. Chen, M. et al. Efficient gene delivery and expression in pancreas and pancreatic tumors by capsid-optimized AAV8 vectors. *Hum. Gene Ther. Methods* **28**, 49–59 (2017).
58. Moreno-Layseca, P., Icha, J., Hamidi, H. & Ivaska, J. Integrin trafficking in cells and tissues. *Nat. Cell Biol.* **21**, 122–132 (2019).
59. Kotecha, A. et al. Rules of engagement between alphavbeta6 integrin and foot-and-mouth disease virus. *Nat. Commun.* **8**, 15408 (2017).
60. Jackson, T., Sheppard, D., Denyer, M., Blakemore, W. & King, A. M. The epithelial integrin alphavbeta6 is a receptor for foot-and-mouth disease virus. *J. Virol.* **74**, 4949–4956 (2000).
61. Israeli, D. et al. An AAV-SGCG dose-response study in a gamma-sarcoglycanopathy mouse model in the context of mechanical stress. *Mol. Ther. Methods Clin. Dev.* **13**, 494–502 (2019).
62. Conway, P., Tyka, M. D., DiMaio, F., Konerding, D. E. & Baker, D. Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Sci.* **23**, 47–55 (2014).
63. Meliani, A. et al. Determination of anti-adeno-associated virus vector neutralizing antibody titer with an in vitro reporter system. *Hum. Gene Ther. Methods* **26**, 45–53 (2015).
64. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
65. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
66. Anders, S., Pyl, P. T. & Huber, W. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
67. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
68. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
69. Waisman, A., Norris, A. M., Elias Costa, M. & Kopinke, D. Automatic and unbiased segmentation and quantification of myofibers in skeletal muscle. *Sci. Rep.* **11**, 11793 (2021).

Acknowledgements

This work was supported by the "Association Française contre les Myopathies" (AFM), and "Institut National de la Santé Et de la Recherche Médicale" (INSERM, FranceRelance N°221513A10). The authors are Genopole's members, first French biocluster dedicated to genetic, biotechnologies and biotherapies. We are grateful to the "Imaging and Cytometry Core Facility" and to the in vivo evaluation, services of Genethon for technical support, to Ile-de-France Region, to Conseil Départemental de l'Essonne (ASTRE), INSERM and GIP Genopole, Evry for the purchase of the equipment. We would like to acknowledge the technical help of Carolina Pacheco Algalan, Alejandro Arco Hierves, and Louise Philidet. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

Author contributions

The project was conceptualized by A.V.H. and I.R. A.V.H., L.S., E.P., A.D., P.R.L.B., N.G., P.V., and J.P. conducted experiments and performed data analysis. Funding supporting this project was obtained by I.R. A.V.H. and I.R. supervised the project. The manuscript was written by A.V.H. and I.R.

Competing interests

A.V.H. and I.R. are inventors on PCT application EP2023/065499 for the integration of RGDLxxL/I motif in AAV capsid for enhanced muscle transduction efficiency. I.R. is a part-time employee of Atamyo Therapeutics. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-52002-4.

Correspondence and requests for materials should be addressed to Ai Vu Hong or Isabelle Richard.

Peer review information *Nature Communications* thanks Renzhi Han, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.