Pervasive Horizontal Transfer of Adeno-Associated Virus Capsid Genes

Robert J. Gifford [1,2]

[1] MRC-University of Glasgow Centre for Virus Research, University of Glasgow, Glasgow, UK, [2] Centre for Epidemic Response and Innovation, Stellenbosch University, Stellenbosch, South Africa

Abstract

Adeno-associated viruses (AAVs) are small, generally apathogenic DNA viruses with exceptional gene delivery capabilities, making them indispensable in gene therapy and biomedical research. Despite their growing therapeutic use, key aspects of AAV natural biology remain poorly understood, complicating efforts to explain occasional AAV-associated disease and optimize gene therapy vectors. By analyzing sequence data from virus isolates and endogenous viral elements (EVEs), I reveal a striking evolutionary pattern: AAV sub-lineages, defined by the replication-associated (*rep*) gene, have broadly co-diverged with host groups over millions of years, whereas capsid (*cap*) diversity has been shaped by extensive recombination. Moreover, one capsid lineage, *Mammalian-wide* (*M-wide*), has spread horizontally across diverse *rep* lineages and host taxa via multiple recombination events. Furthermore, some AAVs with *M-wide* capsids - including AAV-4, AAV-12, and bovine AAV (BAAV) - originate from historical adenovirus (Ad) stocks, raising the possibility that laboratory conditions may have contributed to capsid transfer. Distinguishing natural from laboratory-driven recombination is essential for understanding the ecology of AAV infection and its implications for gene therapy. A systematic sequencing effort in human and primate populations is needed to assess the extent of recombinant capsid acquisition, determine whether laboratory-driven recombination has shaped naturally circulating AAV diversity, and track ongoing recombination events that could impact vector safety and efficacy.

Main Text

Introduction

Adeno-associated viruses (AAVs) are small, non-enveloped, single-stranded DNA viruses of the *Dependoparvovirus* genus (family *Parvoviridae*) [1]. They require co-infection with helper viruses, such as adenoviruses or herpesviruses, for productive replication [2]. The AAV genome encodes two primary genes: *rep* (essential for replication and genome packaging) and *cap* (encoding capsid proteins that determine host range and tissue specificity). The genome, flanked by inverted terminal repeats (ITRs), employs rolling hairpin replication mediated by Rep proteins [3, 4].

AAVs are widely used in gene therapy, where viral genes are replaced with therapeutic transgenes while retaining the ITRs and supplying *rep* and *cap* genes in *trans* [5]. Recombinant AAV (rAAV) vectors are central in treating genetic disorders, with regulatory approval for conditions like hemophilia and inherited retinal diseases, and transformative potential for treating many other genetic diseases [6]. The capsid proteins, which determine host range and tissue specificity, are key targets for modification to improve gene delivery, transduction efficiency, and tissue tropism, as well as to evade neutralizing antibodies that arise from prior exposure to wild-type AAVs [7].

Despite decades of research on rAAVs, the natural biology of AAVs remains poorly understood. Like other parvoviruses, AAVs are thought to be capable of evolving rapidly due to high recombination and mutation rates [8]. However, their extensive genomic 'fossil record'—comprising EVEs—suggests stable, long-term host-virus associations [9]. Sequence and serological evidence indicate that diverse AAVs circulate in humans and non-human primates [10], but virus-host interaction dynamics remain unclear. And while AAVs are typically considered apathogenic [2], or even beneficial [11], reports of AAV-2 linked to unexplained childhood hepatitis [12], along with emerging evidence of potential pathogenicity in other contexts, underscore the need for deeper insights into AAV ecology and host interactions [13].

A comprehensive understanding of AAV evolutionary dynamics is essential both for optimizing AAV-based gene therapies and for elucidating the relationships between AAV infection and

disease. Here, we reveal that AAV capsids have undergone extensive horizontal transfer, a process that may influence both viral evolution and gene therapy applications. Furthermore, our findings prompt critical questions about the influence of laboratory-driven recombination on AAV diversity.

Results

AAV Sub-lineages Show Host-Virus Co-divergence in rep

Phylogenetic analysis of *rep* genes from AAV isolates and EVEs establishes that the AAVs form a distinct, robustly supported subclade in the genus *Dependoparvovirus* (**Fig. 1a**). Rooted AAV *rep* trees, with high bootstrap support (>80%), mirror vertebrate phylogeny, suggesting long-term host-virus co-divergence. Among AAVs isolated from placental mammals, clades align with mammalian orders (including primates, rodents, bats, and artiodactyls) indicating stable, long-term host-virus associations. EVE data independently support these relationships: the '*dependo.13-cercopithecidae*' locus, an AAV *rep* pseudogene orthologous across multiple cercopithecine monkey species, indicates a >20 million-year (My) primate AAV-host relationship. Similarly, the '*dependo.46-gliridae*' locus, clusters with AAVs from mice and voles, suggesting an AAV-rodent association exceeding 30 My.

Primate AAV clades also show structured host associations (**Fig. 1b**). A well-supported subclade includes AAV serotypes that are generally associated with Asian primates and humans, while taxa outside this subclade exhibit more mixed host associations, including African primates (**Table 1**). While this biogeographic pattern is evident, it is not absolute: serotypes from the Asian clade have occasionally been detected in African primates, and AAVs associated with both Asian and African primates are known to circulate in humans [10]. However, the full extent of these associations remains unclear due to uncertain host linkages and uneven sampling.

Extensive Horizontal Transfer of AAV Capsids

Phylogenies constructed using *cap* genes reveal a markedly different topology compared to *rep*-based phylogenies, highlighting the distinct evolutionary trajectories of these genomic regions. Within the AAV clade, *cap* lineages broadly align with mammalian host orders, mirroring

patterns observed in *rep*. However, one notable exception emerges: a single *cap* lineage—here termed '*Mammalian-wide-1*' (*M-wide*)—is widely distributed across multiple *rep* lineages and diverse host taxa, providing strong evidence that recombination has enabled pervasive horizontal transfer of AAV capsid genes (**Fig. 2**).

As the tanglegram clearly illustrates, several *M-wide*-bearing taxa—including BAAV, AAV-4, AAV-11, AAV-12, and *Myodes* AAV type 2 (MgAAV-2)—are embedded in *rep* phylogenies alongside viruses whose capsids align with their respective host groups (artiodactyls, primates, and rodents). The most parsimonious explanation for this pattern is that multiple *rep* lineages independently acquired *M-wide* capsids via recombination.

AAVs were first identified as contaminants in early adenovirus (Ad) stocks that were derived from primate and cattle tissue samples in the 1950s and propagated in culture. Notably, several AAVs encoding *M-wide* capsids —including AAV-4, AAV-12 and bovine AAV (BAAV), were originally isolated as contaminants of these stocks (**Table 1**). Phylogenetic analyses show that the *M-wide* capsids of these viruses form a distinct, highly supported subclade (**Fig. 2**), suggesting a potential shared history linked to laboratory culture.

Notably, one virus in this 'laboratory-associated' clade (AAV-11) was not isolated from Ad stocks but was obtained from macaque tissue in 2004. However, its close similarity (>99% nucleotide identity) in rep to AAV-10 (which encodes the standard primate AAV capsid) suggests a recent divergence, possibly as recently as the past few decades (Extended Methods). This timeframe suggests AAV-11 could have acquired its *M-wide* capsid subsequent to culture-associated recombination events.

Discussion

This study leverages the growing availability of AAV and EVE sequence data to characterize AAV evolutionary dynamics. Phylogenetic analysis reveals that, over long evolutionary timescales, AAV *rep* lineages have broadly co-diverged with their hosts. However, comparative analysis of *rep* and *cap* phylogenies demonstrates that capsid diversity has been shaped by extensive horizontal transfer, with a single lineage (*M-wide*) spreading widely across diverse AAV *rep* lineages and host taxa via recombination (**Fig. 2**). The basal placement of rodent-associated

AAVs within the *M-wide* clade suggests that rodents may represent the ancestral source of *M-wide* capsids.

Notably, the properties of *M-wide* capsids encoded by primate AAVs have been examined in depth, revealing striking functional diversity. AAV-4 preferentially transduces epithelial cells, particularly in the retinal pigmented epithelium and airway, due to its O-linked sialic acid binding. AAV-11, in contrast, shows affinity for intestinal tissues and lymphoid organs, with emerging evidence of neuronal tropism. AAV-12, uniquely, does not use conventional AAV glycan receptors and instead displays a strong preference for nasal and salivary gland tissues. Such variability among closely related capsids underscores the potential of capsid recombination to generate functionally distinct AAVs.

Beyond the broad host-virus co-divergence observed at higher taxonomic levels (**Fig 1a**), phylogenetic analysis of *rep* also suggests finer-scale structuring of AAV diversity. In *rep* phylogenies, primate AAVs form a well-supported subclade, with lineages primarily linked to Asian primates clustering separately from those associated with African primates and their endogenous counterparts (**Fig. 1b**). Notably, AAVs from both groups circulate in humans, a raising key question: Do these patterns reflect long-standing host-virus associations, or have they been shaped by recent shifts in transmission dynamics or human activity?

This question gains additional significance considering the presence of *M-wide* capsids in bovine and primate AAVs that were historically propagated as adventitious agents in adenovirus stocks. extensive horizontal transfer observed in AAV capsid genes. While in many cases the acquisition of *M-wide* likely reflects events in the distant evolutionary past, the history of these isolates suggests a possible link to laboratory-associated recombination. Adenoviruses, isolated in the 1950s, were among the first mammalian viruses to be maintained *in vitro*, at a time when few viruses could be cultured outside live hosts. Given the limited virological tools available, AAVs likely remained undetected in these early adenovirus cultures for years. These stocks, often derived from biological materials spanning multiple species, may have provided ideal conditions for AAV co-infection and recombination. The presence of helper viruses, high viral loads, passage in permissive cell types, and the potential for inadvertent serial passage in research settings could have further facilitated capsid gene exchange.

Since the 2000s, however, AAVs encoding *M-wide* capsids have been detected in primate centers [14] and dairy cattle [15] (**Table 1**). Thus, if they did indeed originate in lab cultures, these viruses seem to have since established themselves in host populations. Alternatively, *M-wide* may be naturally widespread, with its presence in Ad stocks reflecting only the pre-existing diversity of bovine and primate AAVs. Disentangling these scenarios is essential to developing a fuller understanding of AAV ecology and evolution. Although no evidence links *M-wide* acquisition to disease, determining whether this recombination occurred naturally or through anthropogenic factors is crucial for developing an ecosystems-level understanding of AAV ecology and evolution [16].

With the power and affordability of modern sequencing, broad-scale characterization and surveillance of AAV diversity is entirely feasible. Such efforts would not only clarify the natural ecology and evolution of AAVs but also inform gene therapy applications by identifying circulating diversity and potential immune interactions. Moreover, given that rAAV production can potentially lead to encapsidated contaminant sequences [17], the possibility of therapy-associated recombination as a long-term risk warrants careful evaluation.

Materials and Methods

Sequence Data and Classification

Adeno-associated virus (AAV) genomic sequences were obtained from NCBI GenBank and curated in AAV-Atlas, a GLUE-based database integrating published AAV genomic data with standardized metadata, alignments, and serotype classifications {Singer, 2018 #420}. To focus on naturally circulating AAV diversity, sequences were filtered to exclude patent-associated clones and laboratory-derived constructs. AAV sequences were classified using Maximum Likelihood Clade Assignment (MLCA) as implemented in GLUE {Singer, 2018 #420}. Only wild-type isolates with complete or near-complete *rep* and *cap* gene coverage were included.

AAV-derived endogenous viral elements (EVEs) were sourced from Parvovirus-GLUE-EVE, a repository cataloging identified loci with genomic context and orthology relationships {Campbell, 2022 #156}. EVE orthology across host species was confirmed using comparative genomic analyses of flanking regions via BLASTn and whole-genome alignment where available.

Gene Identification and Alignment

The rep and cap coding regions were extracted based on reference genome annotations in AAV-Atlas. Coding regions were aligned using a codon-aware protocol in GLUE {Singer, 2018 #420}, with virtual translation ensuring reading frame integrity for EVE sequences, followed by back-translation for codon-level alignment. Multiple-sequence alignments were generated using MUSCLE {Edgar, 2004 #422}, and manual curation in Se-Al {Rambaut, 1997 #424} was performed to confirm the presence of essential functional motifs.

Phylogenetic Reconstruction

Maximum likelihood (ML) phylogenies were inferred for rep and cap genes separately using RAxML (v8.2.12) {Stamatakis, 2014 #421}, with optimized partitioning to maximize alignment coverage and taxon representation. The rep-based tree was rooted using mammalian Dependoparvovirus outgroups, whereas the cap-based tree was midpoint-rooted, producing a congruent topology. Bootstrap support (1,000 replicates) was computed using the GTR+G model for nucleotide sequences and JTT model for amino acid sequences, selected via likelihood ratio testing.

Estimation of Divergence Time

To estimate the divergence time between AAV-10 and AAV-11, pairwise sequence divergence was calculated within the rep gene. Across 1,869 nucleotides, the two viruses differed at 9 positions, corresponding to 0.482% sequence divergence. Using a molecular clock model with published substitution rates for single-stranded DNA viruses {Duffy, 2008 #427}, ranging from 1×10^{-5} to 2×10^{-4} substitutions/site/year, we estimated divergence occurred within the last 12 to 241 years, suggesting a relatively recent evolutionary split.

Phylogenetic Incongruence Testing

Phylogenetic incongruence between rep and cap genes was assessed using the Shimodaira-Hasegawa (SH) test, performed in CONSEL {Shimodaira, 2001 #423}. ML trees were inferred separately from cap and rep partitions using RAxML (v8.2.12) {Stamatakis, 2014 #421}, and per-site log-likelihoods were computed with the -f g option on a concatenated alignment containing only high-confidence sites. The Approximately Unbiased (AU) test strongly rejected the rep tree (AU = 0.006, SH = 0.012) in favor of the cap tree (AU = 0.994, SH = 0.988), providing strong evidence of independent evolutionary histories shaped by recombination.

Data Availability

All sequence alignments, phylogenetic trees, and analysis pipelines used in this study are publicly available through the following GLUE-based repositories:

- AAV-Atlas: https://github.com/giffordlabcvr/AAV-Atlas
- Parvovirus-GLUE-EVE: https://github.com/giffordlabcvr/Parvovirus-GLUE-EVE

These resources provide open-source, platform-independent tools and datasets for full reproducibility of the analyses described in this study.

Figure 1. Phylogeny of AAV *rep* genes supports host-virus co-divergence at higher taxonomic levels.

(a) **AAV genome structure.** A schematic representation of the 4.7 kilobase (Kb) AAV genome, which encodes two primary genes: *rep* and *cap*, flanked by inverted terminal repeats (ITRs).

(b) **Phylogeny of AAV and AAV-derived EVEs based on *rep*.** A rooted maximum-likelihood phylogeny depicting evolutionary relationships among AAVs and AAV-derived endogenous viral elements (EVEs) based on *rep* polypeptide sequences. Brackets indicate avian, marsupial, and placental mammal AAVs clustering with their respective hosts, reflecting host phylogeny at higher taxonomic levels. The inset phylometric tree (bottom right) illustrates the timescale of divergence among host taxa. Additionally, long-term associations between AAV lineages and placental mammal orders are evident. Boxes next to EVE taxa indicate approximate minimum ages inferred from orthology. Branch lengths represent evolutionary change (amino acid substitutions per site) as indicated by the scale bar. Asterisks denote bootstrap values >70% (1,000 replicates). Values for phylogenies reconstructed using amino acids are shown above branches, while nucleotide-based values are shown below. The primate AAV clade (collapsed) is expanded in panel (c).

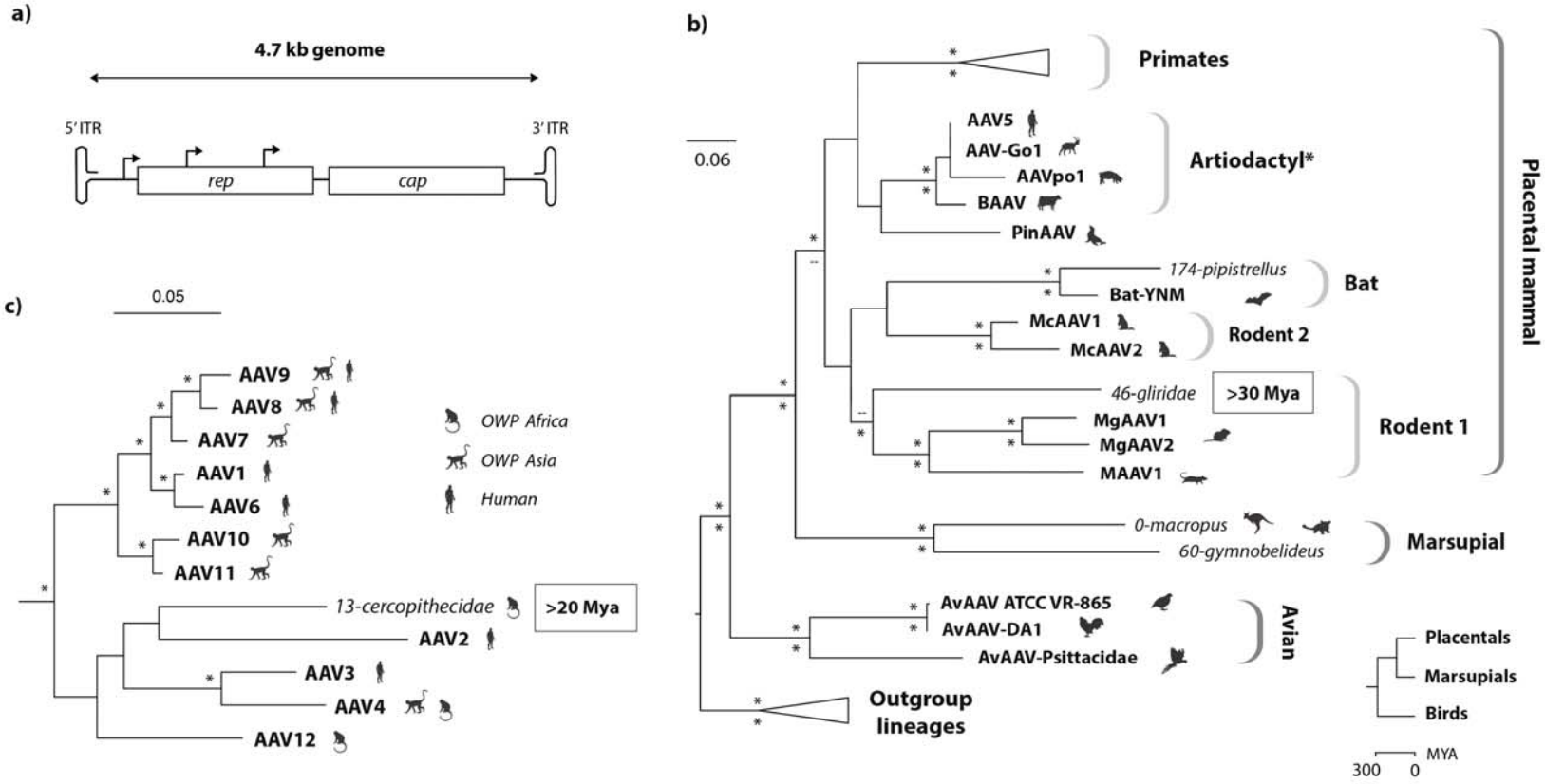
(c) **Phylogeny of primate AAVs and EVEs based on *rep*.** A rooted maximum-likelihood phylogeny showing evolutionary relationships among primate AAVs and AAV-derived EVEs based on *rep* nucleotide sequences. Taxon associations with humans, non-human primates from Asia, and non-human primates from Africa follow the key to the right of the phylogeny. As in (b), boxes next to EVE taxa indicate approximate minimum ages based on orthology. Asterisks denote bootstrap values >70% (1,000 replicates). Branch lengths represent evolutionary change (nucleotide substitutions per site) as indicated by the scale bar.

Abbreviations: MYA, million years ago.

Figure 2. Tanglegram showing *rep/cap* phylogenetic discordance and M-wide spread. A tanglegram comparing phylogenies of *rep* (left) and *cap* (right), highlighting recombination signals. The *rep*-based tree is rooted on *Dependoparvovirus* outgroups, while the *cap*-based tree is midpoint-rooted. Shaded clades correspond to host taxonomic groups, while the M-wide

capsid lineage (red) spans multiple mammalian orders. Connecting lines link taxa in both trees, illustrating discordant *rep/cap* phylogenies and independent acquisitions of M-wide. Taxa potentially acquiring M-wide are indicated in bold. A darker shaded region within the M-wide clade highlights a sub-lineage associated with M-wide-encoding AAV isolates obtained from adenovirus stocks (see Table 1). Pinniped AAV (PinAAV) may have acquired M-wide from carnivore AAVs, but available data are insufficient to confirm this. Asterisks denote bootstrap support >70%.

Abbreviations: AAV, adeno-associated virus; EPV, endogenous parvoviral element.


a)
4.7 kb genome
5' ITR
3' ITR
rep
cap
b)
0.06
Primates
AAV5
AAV-Go1
AAVpo1
BAAV
PinAAV
Artiodactyl*
174-pipistrellus
Bat-YNM
Bat
McAAV1
McAAV2
Rodent 2
46-gliridae
>30 Mya
MgAAV1
MgAAV2
MAAV1
Rodent 1
Placental mammal
0-macropus
60-gymnobelideus
Marsupial
AvAAV ATCC VR-865
AvAAV-DA1
AvAAV-Psittacidae
Avian
Outgroup lineages
Placentals
Marsupials
Birds
MYA
300
0
c)
0.05
AAV9
AAV8
AAV7
AAV1
AAV6
AAV10
AAV11
OWP Africa
OWP Asia
Human
13-cercopithecidae
>20 Mya
AAV2
AAV3
AAV4
AAV12

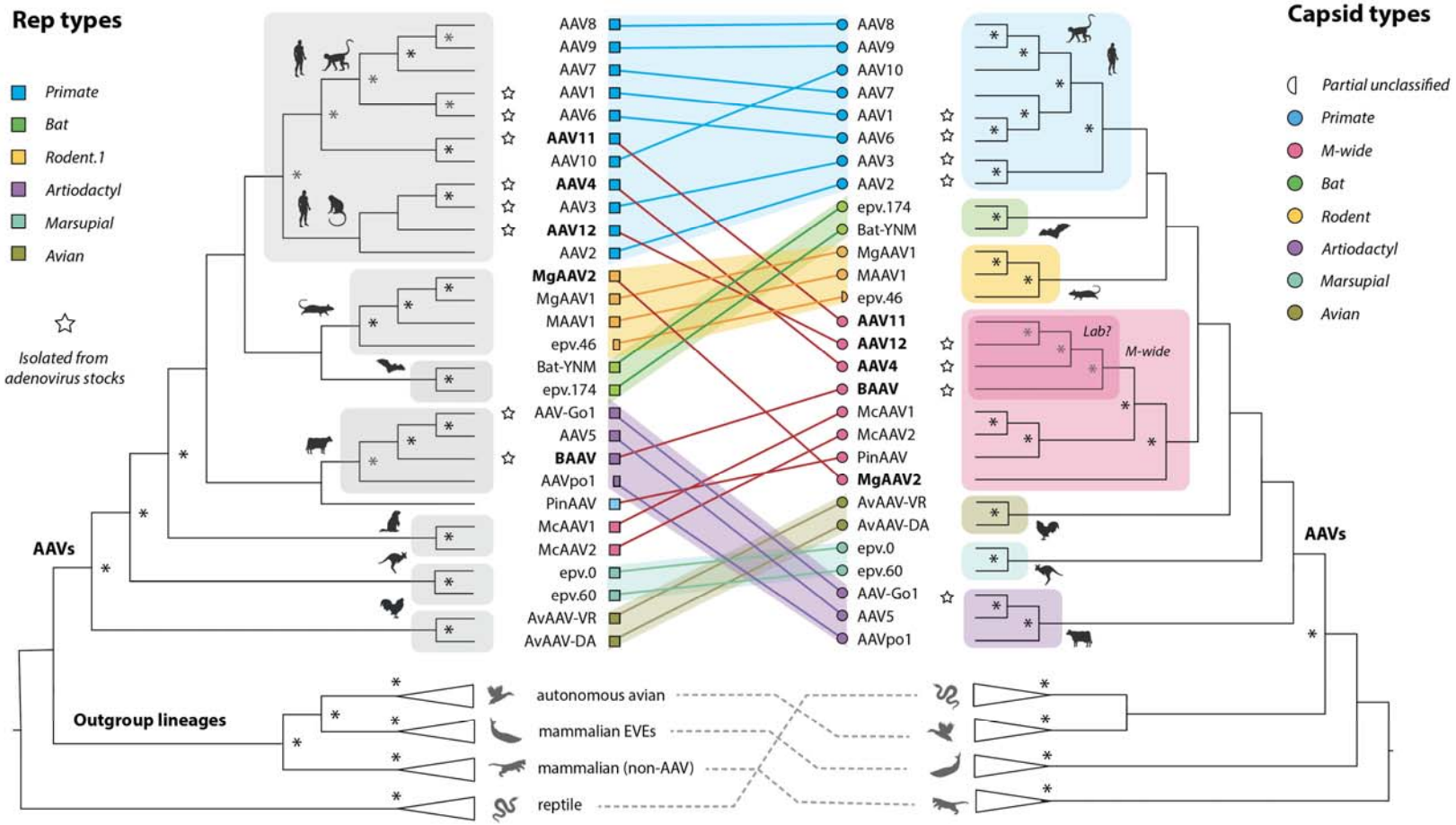

Rep types
Primate
Bat
Rodent.1
Artiodactyl
Marsupial
Avian
Isolated from adenovirus stocks
AAVs
Outgroup lineages
AAV8
AAV9
AAV7
AAV1
AAV6
AAV11
AAV10
AAV4
AAV3
AAV12
AAV2
MgAAV2
MgAAV1
MAAV1
epv.46
Bat-YNM
epv.174
AAV-Go1
AAV5
BAAV
AAVpo1
PinAAV
McAAV1
McAAV2
epv.0
epv.60
AvAAV-VR
AvAAV-DA
autonomous avian
mammalian EVEs
mammalian (non-AAV)
reptile
Lab?
M-wide
Capsid types
Partial unclassified
Primate
M-wide
Bat
Rodent
Artiodactyl
Marsupial
Avian
AAVs

Table 1. Isolation History & Host Associations of Primate and Artiodactyl Adeno-Associated Viruses

AAV Isolation history			Ad Isolation history					AAV-Host Associations			
Name[1]	Isolation Source	Year	Ad Strain(s)	Iso. Source	Cell culture	Year[2]	Ad Host	Primary Association	Seq.[3]	Serology[4]	
										Human	NHP
AAV-1	Ad stocks	1965	Ad7	Human	MK, Human	1953	*H. sapiens*	Humans	Yes	++++	+++
AAV-2	Ad stocks	1965	Ad12, Ad7	Human	Human	1953	*H. sapiens*	Humans	Yes	+++++	++
AAV-3	Ad stocks	1965	Ad7[19]	Human	AGMK	1953	*H. sapiens*	Humans?	Yes	+++	++
AAV-4*	Ad stocks	1968	SV15	Macaque	MK, AGMK	1956	*M. mulatta*	OWM Africa	No	+	+++++
AAV-5	Hu. Tissue (Culture)	1982	*n/a*					Humans?	No	+++	++
AAV-6	Ad stocks	1998	Ad5	Human	Human	1953	*H. sapiens*	Humans	No	++++	++
AAV-7	OWM Tissue (PCR)	2004						OWM Asia	Yes	++	+++++
AAV-8	OWM Tissue (PCR)	2004						OWM Asia	Yes	++	+++++
AAV-9	Hu. Tissue (PCR)	2004						OWM Asia	Yes	+++	+++++
AAV-10	OWM Tissue (PCR)	2004						OWM Asia	Yes	+++	+++++
AAV-11*	OWM Tissue (PCR)	2004						OWM Asia	Yes	*n/d*	++
AAV-12*	Ad stocks	2008	SV18	Vervet	n/k	*n/k*	*C. aethiops*	OWM Africa	No	-	*n/d*
BAAV*	Ad stocks	1970	Bovine Ad	Cow	Bovine	1957	*B. taurus*	Cattle	Yes	*n/d*	*n/d*
AAV-Go.1	Ad stocks	2004	Caprine Ad	Goat	Caprine	2004	*C. hircus*	Goat	Yes	*n/d*	*n/d*

Footnote:

1. Asterisks indicate isolates with M-wide capsids.
2. Where multiple adenovirus (Ad) stocks are implicated, the year of the earliest obtained is shown.
3. Independent sequence evidence for presence in the presumed primary host.
4. Serological prevalence in humans and non-human primates is categorized as follows:
    - **+**: Very low
    - **++**: Low
    - **+++**: Moderate
    - **++++**: High
    - **+++++**: Very high
    - -: Absent
    - *n/d*: Not determined

5. Serological information was compiled from the following sources [11, 20-29].

Abbreviations: AAV = adeno-associated virus; PCR = polymerase chain reaction; OWM = Old World Monkeys; MK = monkey kidney cells; AGMK = African green monkey kidney cells; Hu. = Human; n/a = not applicable; n/k = not known.

Acknowledgments

I am grateful to Professor Robert Kotin for his valuable feedback on the manuscript.

References

1. Cotmore, S.F., et al., *ICTV Virus Taxonomy Profile: Parvoviridae.* J Gen Virol, 2019. **100**(3): p. 367-368.
2. Meier, A.F., C. Fraefel, and M. Seyffert, *The Interplay between Adeno-Associated Virus and its Helper Viruses.* Viruses, 2020. **12**(6).
3. Tattersall, P. and D.C. Ward, *Rolling hairpin model for replication of parvovirus and linear chromosomal DNA.* Nature, 1976. **263**(5573): p. 106-9.
4. Im, D.S. and N. Muzyczka, *The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity.* Cell, 1990. **61**(3): p. 447-57.
5. Hastie, E. and R.J. Samulski, *Adeno-associated virus at 50: a golden anniversary of discovery, research, and gene therapy success--a personal perspective.* Human gene therapy, 2015. **26**(5): p. 257-265.
6. Wang, J.H., et al., *Adeno-associated virus as a delivery vector for gene therapy of human diseases.* Signal Transduct Target Ther, 2024. **9**(1): p. 78.
7. Zolotukhin, S. and L.H. Vandenberghe, *AAV capsid design: A Goldilocks challenge.* Trends Mol Med, 2022. **28**(3): p. 183-193.
8. Duffy, S., L.A. Shackelton, and E.C. Holmes, *Rates of evolutionary change in viruses: patterns and determinants.* Nat Rev Genet, 2008. **9**(4): p. 267-76.
9. Campbell, M.A., et al., *Comparative analysis reveals the long-term coevolutionary history of parvoviruses and vertebrates.* PLoS Biol, 2022. **20**(11): p. e3001867.
10. Gao, G., et al., *Clades of Adeno-associated viruses are widely disseminated in human tissues.* J Virol, 2004. **78**(12): p. 6381-8.
11. Mayor, H.D., et al., *Antibodies to adeno-associated satellite virus and herpes simplex in sera from cancer patients and normal adults.* Am J Obstet Gynecol, 1976. **126**(1): p. 100-4.
12. Tacke, F., *Severe hepatitis outbreak in children linked to AAV2 virus.* Nature, 2023. **617**: p. 471-472.
13. Sant'Anna, T.B. and N.M. Araujo, *Adeno-associated virus infection and its impact in human health: an overview.* Virology Journal, 2022. **19**(1): p. 173.
14. Mori, S., et al., *Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsid protein.* Virology, 2004. **330**(2): p. 375-83.
15. Ng, T.F., et al., *A metagenomics and case-control study to identify viruses associated with bovine respiratory disease.* J Virol, 2015. **89**(10): p. 5340-9.
16. French, R.K. and E.C. Holmes, *An Ecosystems Perspective on Virus Evolution and Emergence.* Trends Microbiol, 2020. **28**(3): p. 165-175.
17. Brimble, M.A., S.M. Winston, and A.M. Davidoff, *Stowaways in the cargo: Contaminating nucleic acids in rAAV preparations for gene therapy.* Mol Ther, 2023. **31**(10): p. 2826-2838.
18. Singer, J.B., et al., *GLUE: a flexible software system for virus sequence data.* BMC Bioinformatics, 2018. **19**(1): p. 532.

19. Benyesh-Melnick, M. and H.S. Rosenberg, *The Isolation of Adenovirus Type 7 from a Fatal Case of Pneumonia and Disseminated Disease.* J Pediatr, 1964. **64**: p. 83-7.
20. Sprecher-Goldberger, S., et al., *Complement-fixation antibodies to adenovirus-associated viruses, cytomegaloviruses and herpes simplex viruses in patients with tumors and in control individuals.* Am J Epidemiol, 1971. **94**(4): p. 351-8.
21. Klamroth, R., et al., *Global Seroprevalence of Pre-existing Immunity Against AAV5 and Other AAV Serotypes in People with Hemophilia A.* Hum Gene Ther, 2022. **33**(7-8): p. 432-441.
22. Hüser, D., et al., *High Prevalence of Infectious Adeno-associated Virus (AAV) in Human Peripheral Blood Mononuclear Cells Indicative of T Lymphocytes as Sites of AAV Persistence.* J Virol, 2017. **91**(4).
23. Georg-Fries, B., et al., *Analysis of proteins, helper dependence, and seroepidemiology of a new human parvovirus.* Virology, 1984. **134**(1): p. 64-71.
24. Erles, K., P. Sebökovà, and J.R. Schlehofer, *Update on the prevalence of serum antibodies (IgG and IgM) to adeno-associated virus (AAV).* J Med Virol, 1999. **59**(3): p. 406-11.
25. Coria, M.F. and H.D. Lehmkuhl, *Isolation and identification of a bovine adenovirus type 3 with an adenovirus-associated virus.* Am J Vet Res, 1978. **39**(12): p. 1904-6.
26. Calcedo, R. and J.M. Wilson, *Humoral Immune Response to AAV.* Front Immunol, 2013. **4**: p. 341.
27. Boutin, S., et al., *Prevalence of serum IgG and neutralizing factors against adeno-associated virus (AAV) types 1, 2, 5, 6, 8, and 9 in the healthy population: implications for gene therapy using AAV vectors.* Hum Gene Ther, 2010. **21**(6): p. 704-12.
28. Blacklow, N.R., et al., *A seroepidemiologic study of adenovirus-associated virus infection in infants and children.* Am J Epidemiol, 1971. **94**(4): p. 359-66.
29. Blacklow, N.R., M.D. Hoggan, and W.P. Rowe, *Serologic evidence for human infection with adenovirus-associated viruses.* J Natl Cancer Inst, 1968. **40**(2): p. 319-27.