

## Molecular Characterization of Adeno-Associated Viruses Infecting Children

Chun-Liang Chen,<sup>†</sup> Ryan L. Jensen,<sup>†‡</sup> Bruce C. Schnepf,<sup>‡</sup> Mary J. Connell,<sup>‡</sup>  
 Richard Shell, Thomas J. Sferra, Jeffrey S. Bartlett, K. Reed Clark,  
 and Philip R. Johnson\*

*Center for Gene Therapy, Columbus Children's Research Institute, Columbus Children's Hospital,  
 and Department of Pediatrics, College of Medicine and Public Health, The Ohio State University,  
 Columbus, Ohio 43205*

Received 20 June 2005/Accepted 7 September 2005

**Although adeno-associated virus (AAV) infection is common in humans, the biology of natural infection is poorly understood. Since it is likely that many primary AAV infections occur during childhood, we set out to characterize the frequency and complexity of circulating AAV isolates in fresh and archived frozen human pediatric tissues. Total cellular DNA was isolated from 175 tissue samples including freshly collected tonsils ( $n = 101$ ) and archived frozen samples representing spleen ( $n = 21$ ), lung ( $n = 16$ ), muscle ( $n = 15$ ), liver ( $n = 19$ ), and heart ( $n = 3$ ). Samples were screened for the presence of AAV and adenovirus sequences by PCR using degenerate primers. AAV DNA was detected in 7 of 101 (7%) tonsil samples and two of 74 other tissues (one spleen and one lung). Adenovirus sequences were identified in 19 of 101 tonsils (19%), but not in any other tissues. Complete capsid gene sequences were recovered from all nine AAV-positive tissues. Sequence analyses showed that eight of the capsid sequences were AAV2-like (~98% amino acid identity), while the single spleen isolate was intermediate between serotypes 2 and 3. Comparison to the available AAV2 crystal structure revealed that the majority of the amino acid substitutions mapped to surface-exposed hypervariable domains. To further characterize the AAV capsid structure in these samples, we used a novel linear rolling-circle amplification method to amplify episomal AAV DNA and isolate infectious molecular clones from several human tissues. Serotype 2-like viruses were generated from these DNA clones and interestingly, failed to bind to a heparin sulfate column. Inspection of the capsid sequence from these two clones (and the other six AAV2-like isolates) revealed that they lacked arginine residues at positions 585 and 588 of the capsid protein, which are thought to be essential for interaction with the heparin sulfate proteoglycan coreceptor. These data provide a framework with which to explore wild-type AAV persistence in vivo and provide additional tools to further define the biodistribution and form of AAV in human tissues.**

Infection with wild-type adeno-associated virus (AAV) is common in humans and occurs without apparent pathogenic consequence (2, 4, 5, 10). The bulk of our understanding about naturally acquired AAV infection comes from observations made over 30 years ago (3–5). However, the emergence of human gene transfer vectors based on recombinant AAV has rekindled an interest in the biology of wild-type AAV (26). Past epidemiologic data notwithstanding, almost nothing is known about the life cycle of wild-type AAV in a permissive host like humans. For example, we lack clear information about the portal of entry, the role of helper viruses in primary infection, cellular receptors for viral attachment, sites of primary or secondary replication, sites of latency, and in vivo molecular forms of viral DNA. A better understanding of these issues will undoubtedly influence the use of recombinant AAV vectors for gene transfer in humans.

Recently, when viral genomes of unexpected diversity were recovered from a surprising array of tissues and organs, it

became clear that AAV infection of primates (human and nonhuman) is much more complex than previously appreciated (11, 13). The genetic diversity observed in these studies suggested that multiple genotypes (many more than the historically accepted five serotypes) of AAV circulate in humans and monkeys. Furthermore, the data implied that whatever the portal of entry, AAV can become widely disseminated following primary infection. The implications of these findings for recombinant AAV gene transfer vectors are unclear, but the immunologic and genetic influence of prior infection on recombinant AAV-mediated gene transfer must be considered.

To extend our understanding of wild-type AAV infection in humans, we set out to characterize AAV genomes directly out of freshly acquired human tissues. Because our assumption was that many (if not all) AAV infections begin in the respiratory tract in children as they are concurrently infected with adenoviruses, we collected tonsils and adenoids from pediatric subjects undergoing surgical excision in an outpatient surgery center. We demonstrated that 7% of these samples contained wild-type AAV DNA. In a follow-on set of experiments, we obtained a range of archived, frozen normal tissues from a repository and showed that 3% of these samples also contained wild-type AAV DNA. Analysis of the *cap* gene sequences from all samples revealed that most of the isolates were closely related to AAV serotype 2; a single isolate shared significant

\* Corresponding author. Present address: Children's Hospital of Philadelphia, 3615 Civic Center Boulevard, 1216B Abramson Research Center, Philadelphia, PA 19104-4318. Phone: 267-426-0351. Fax: 267-426-0363. E-mail: johnsonphi@chop.edu.

<sup>†</sup> C.-L.C. and R.L.J. contributed equally to this work.

<sup>‡</sup> Present address: Children's Hospital of Philadelphia, Philadelphia, PA 19104.

homology with serotypes 2 and 3. Interestingly, none of the isolates in our study were predicted to bind heparin sulfate, suggesting that this receptor is not necessary for wild-type infection in humans.

## MATERIALS AND METHODS

**Human tissues.** All samples were acquired after approval from the Columbus Children's Hospital Institutional Review Board, and where required, informed consent was obtained. Fresh human tonsil and adenoid specimens ( $n = 101$ ) were collected from children aged 2 to 13 years undergoing elective surgical excision at Columbus Children's Hospital. Additional human tissues ( $n = 74$ ) representing normal liver, spleen, muscle, heart, and lung were obtained from the Cooperative Human Tissue Network based at Columbus Children's Hospital. The ages of the individual subjects ranged from 0 to 30 years. Subjects under 6 months of age ( $n = 49$ ) were analyzed but were considered unlikely to have been exposed to adenovirus and AAV infection. In addition, for 3 of the 74 samples, the age of the subjects could not be determined. Thus, there were 22 samples from individuals aged 6 months to 30 years available for analyses, and 6 of the 22 were from individuals older than 14 years.

**Adenovirus isolation from tonsils and adenoids.** Freshly collected tissue from each subject ( $n = 101$ ) was minced, mixed, and divided into three aliquots. Two aliquots were frozen for DNA isolation (see below) while the third was further minced and sieved through a cell strainer. The sieved material was adsorbed onto A459 cells cultured in Dulbecco's modified Eagle's medium containing 20% fetal calf serum. Cultures were maintained at 37°C in 5% CO<sub>2</sub> for 14 days and then blind passed into fresh A549 cells and monitored for an additional 14 days. Cultures were scored positive or negative solely on the appearance of characteristic cytopathic effects.

**DNA isolation from human tissue samples.** Freshly thawed tissue (0.2 to 0.5 g) was digested for 15 h in 3 ml of digestion buffer (10 mM Tris, pH 8; 100 mM NaCl; 0.5% sodium dodecyl sulfate; 25 mM EDTA) supplemented with 2 mg/ml proteinase K at 55°C with constant agitation. DNA was extracted twice with phenol-chloroform/isoamyl alcohol (25:24:1) and RNA removed by DNase-free RNase A digestion (20 µg/ml) (QIAGEN Inc.) at 37°C for 30 min. RNase A was removed by two sequential phenol-chloroform-isoamyl alcohol extractions. A final chloroform extraction was performed, followed by DNA ethanol precipitation using 0.3 M sodium acetate (pH 5.2) and 2 volumes of ethanol. The DNA pellet was air-dried and suspended in 1 ml of 10 mM Tris, pH 8.0, and DNA concentration determined by A<sub>260</sub>.

**AAV capsid PCR.** Nested PCR was used to amplify a 255-bp conserved region of the AAV *cap* gene. An initial round of PCR was performed on genomic DNA (100 ng) using a conserved degenerate primer set (CapSS2978: 5'-GGYGCC GAYGGAGTGGGYARTKCC-3' and Cap 18S: 5'-GAWKCCCARTWGTT GTTRATGAGTC-3'), where Y is C or T, R is A or G, K is G or T, W is A or T, M is A or C, S is C or G, I is inosine. One µl of the first-round PCR served as the template for the second round using nested, degenerate PCR primers (Cap 19S: 5'-GYARTKCCRCGGWRATTGGCA-3' and CapSS3189: 5'-GATGAGTKYTGCCAGTCWCGKGG-3'). Reaction components for both rounds were 400 nM of each primer, 400 nM deoxynucleoside triphosphates, 0.5 unit SureStart *Taq* polymerase (Stratagene), 1X SureStart reaction buffer in a final volume of 25 µl.

PCR cycling conditions were: 1 cycle at 94°C for 12 min; 36 cycles at 94°C for 30 seconds, 52.5°C for 30 seconds, and 72°C for 1 min; followed by a 4-min extension step at 72°C. To confirm PCR amplicon identity, the AAV nested capsid PCR (or Ad PCR products-see below) were resolved on 0.8% agarose gels and in-gel Southern blot hybridization was performed using AAV2 *cap* or adenovirus hexon sequences as the [ $\alpha$ -<sup>32</sup>P]dCTP-labeled hybridization probes. DNA samples identified as containing AAV sequences were subjected to further PCR to isolate the complete AAV capsid coding region.

Due to the length of the *cap* gene product, a dual PCR approach was employed whereby the 5' half of the *cap* gene was amplified as a 1.8-kb PCR product, while the 3' half was amplified as a 1.5-kb PCR product. Two different forward 1.8-kb primers were alternatively utilized to amplify these PCR fragments depending on which primer yielded the greatest amplification efficiency (AAV2-1.8F1: 5'-AAC ATGTGCGCCGTGATTGACGGG-3' or AAV2-1.8F2, 5'-GACCGGATGTTT AAATTTGAAGTC-3'). Similarly, 2 different reverse 1.5-kb primers were alternatively employed for amplification (AAVCap3' Rev, 5'-TCGTTTCAGTTGAAC TTTGGTCTCTGCG-3' or AAVCap3' RevDeg: 5'-CARWRTTYWACTGAM ACGAAT-3'). PCR conditions and primer concentrations were the same as for the 255-bp conserved capsid region. Amplified PCR products were agarose gel purified and cloned into pCR2.1-TOPO vector using TOPO TA cloning kit

according to manufacturer's instructions (Invitrogen, Inc.). DNA clones were sequenced using BigDye terminator chemistry and an ABI 727 capillary electrophoresis automatic sequencer (PE Applied BioSystems Inc.) by the Columbus Children's Research Institute Core Sequencing Laboratory.

**Adenovirus hexon PCR.** For detection of adenovirus sequences, total cellular DNA (100 ng) was subjected to identical PCR conditions as that used for AAV capsid PCR. A primer set targeting a conserved region (300 bp) of the hexon gene was employed (9): hex1, 5'-GCCSCARTGGKWCATACGCACATC-3', and hex2, 5'-CAGCACSCCICGRATGTCAAA-3'.

**Quantitative PCR.** AAV genome copy number in tissue samples was quantified using real-time TaqMan PCR analysis (ABI 7700, PE Applied BioSystems). The primers and probe set were selected following alignment of 255-bp *cap* DNA sequences (ForCAPSS: 5'-AACGACAACCACCTACTTTGGC-3' (50 nM); RevCAPSS: 5'-AAGTGGCAGTGGAACTCTGTC-3' (900 nM); probe, [6-FAM]5'-CTACAGCACCCCTGGGGATTTTGA-3') [6-carboxyfluorescein (FAM)-tetramethylrhodamine] (270 nM). PCR conditions were: 50°C 2 min, 95°C 10 min, 40 cycles of 95°C 15 seconds, and 60°C 1 min using 250 ng of human total cellular DNA in 1X Taqman PCR master mix.

**Sequence analyses.** The DNA and putative protein sequences were aligned and analyzed using the Clustal W method implemented in MegAlign software in DNASTAR (DNASTAR, Inc). The phylogenetic relationship of all AAV DNA sequences and corresponding putative protein sequences were carried out using Neighbor-Joining method with Kimura two-parameter model implemented in MEGA2 package (19). Similarity percentages between AAV2 and the new AAV sequences were determined using one-pair alignment according to the Lipman-Pearson method. Recombination analysis was performed by using the Similarity Plot method as implemented in the SimPlot software (available at <http://www.med.jhu.edu/deptmed/sray/>) (21, 22).

**Heparin binding analysis.** To assess the ability of different AAV capsids to bind heparin, preparations of infectious AAV (harboring the capsid of interest) were subjected to iodixanol density gradient purification (33) and then applied to a POROS HE-20 heparin column (1.7 ml bed volume) using a Biocad Sprint high-pressure liquid chromatography apparatus as previously described (7). Virus was eluted using a linear salt gradient (0.1 to 1 M NaCl). Flowthrough, wash, and eluate (1 ml fractions) were collected for AAV DNase-resistant particle determination.

**AAV serology.** Subjects with the diagnosis of cystic fibrosis were recruited from Columbus Children's Hospital's Cystic Fibrosis Clinic and were eligible for the study if blood was to be drawn or a centrally placed catheter accessed for a clinical indication. Consent for participation was obtained from subjects (or their legal guardian if they were under 18 years of age) after protocol approval from the Columbus Children's Hospital Institutional Review Board.

To detect antibodies reactive with AAV2 capsid proteins, an enzyme-linked immunosorbent assay (ELISA) was performed. Polystyrene 96-well plates (Nunc Immuno Plate/Maxisorp Surface, Nalge Nunc International) were coated with 50 ng of viral capsid protein. AAV2 capsids were affinity purified as previously described (7) from lysates of 293 cells infected with a recombinant adenovirus type 5 carrying the AAV type 2 capsid coding sequences driven by a standard human cytomegalovirus promoter. After plates were blocked (1% normal sheep serum in phosphate-buffered saline containing 5% dry skim milk) for 2 h at room temperature, serum samples diluted 1:100 in phosphate-buffered saline/5% dry skim milk were added to the test wells and incubated for 2 h at room temperature. After washing 3 times (phosphate-buffered saline containing 0.05% Tween), horseradish peroxidase-linked sheep anti-human immunoglobulin G (Sigma Chemical Co., St. Louis, MO) was added, and the plates incubated for an additional 2 h at room temperature. After three washes, o-phenylenediamine dihydrochloride (Sigma) was added to each well and developed in the dark. Optical density at 450 nm (OD<sub>450</sub>) was recorded after 30 min (HTS 7000 Bio Assay Reader; Perkin-Elmer Corp.). The OD<sub>450</sub> reported was the difference between capsid and mock-coated wells. To show that the AAV capsid preparation was not contaminated with adenovirus proteins, AAV2 coated wells were also developed with adenovirus type 5 immune rabbit serum, and were shown to be at background levels. The relationship between age and OD value was estimated using the Pearson correlation.

Serum neutralization activity directed toward AAV type 2 was determined as previously described (20). Briefly, dilutions of sera were incubated at 57°C for 15 min to inactivate complement and were then mixed for 1 h at 37°C with 1,000 infectious units of a recombinant AAV2 vector that expressed *Escherichia coli* β-galactosidase (rAAV/β-gal). The samples were applied onto monolayers of the C12 cells (8) and incubated at 37°C for 4 h. Subsequently, the monolayers were infected with adenovirus type 5 at a multiplicity of infection of 20. At 48 h after rAAV/β-gal transduction, cell monolayers were stained for β-galactosidase activity using the substrate 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

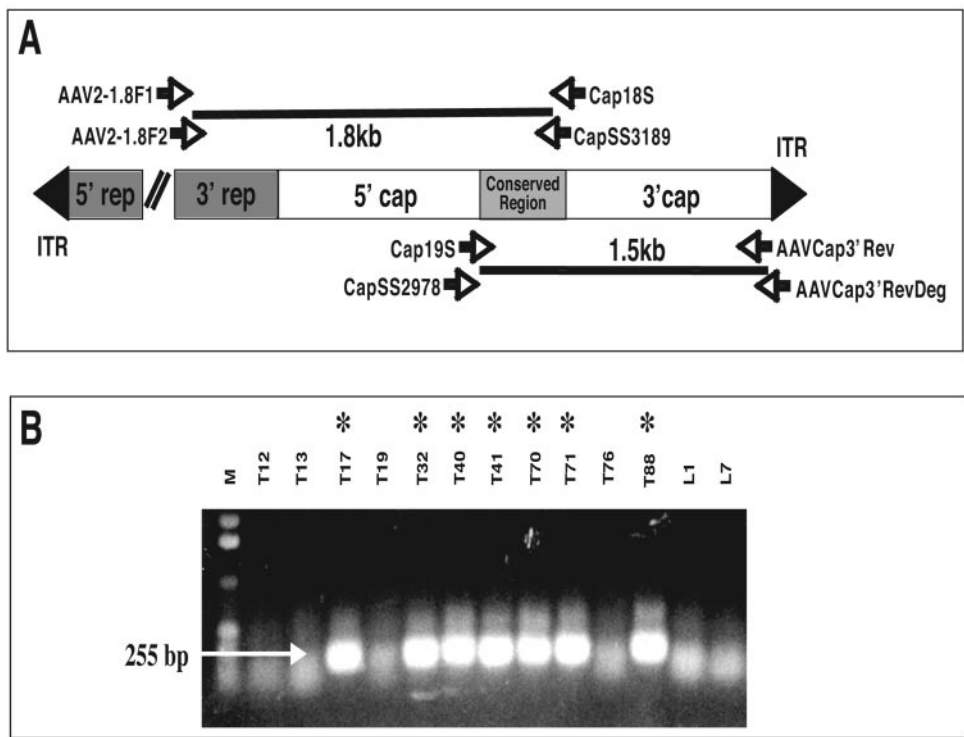


FIG. 1. PCR schematic for amplification of the complete AAV capsid coding region. (A) The diagram depicts the relative location of degenerate primers (given in Materials and Methods) used to amplify the AAV *cap* gene. Initially samples were screened with degenerate nested primers (Cap18S, Cap19S, CapSS3189, CapSS2978) to two conserved regions that flank the HVR3 coding region (gray box). To amplify the complete capsid gene, another set of nested primers were constructed (AAV2-1.8F1, AAV2-1.8F2, AAVCap3' Rev, AAVCap3' RevDeg) that bind to 3' regions of *rep* and *cap* and amplify 1.8- and 1.5-kb DNA amplicons. (B) Representative amplification of the 255-bp conserved AAV sequence from human tissue DNA (100 ng) following nested PCR (see Materials and Methods for reaction conditions). Asterisks indicate the samples that are positive for AAV amplification.

(X-Gal). All serum samples were screened at a dilution of 1:10, and those found to have >90% neutralization (reduction in the number of infectious units compared to controls) were assayed again over a range of twofold dilutions (1:100 to 1:1,600).

**GenBank accession numbers.** The nucleotide sequences described in this paper have been submitted to GenBank. The accession numbers are AY695370, AY695371, AY695372, AY695373, AY695374, AY695375, AY695376, AY695377, and AY695378.

**RESULTS**

**AAV in tonsils and adenoids.** It is likely that many primary AAV infections occur during childhood when the incidence of primary adenovirus infection increases with age (3, 4, 6). Moreover, the probable portal of entry for most of these primary infections is the oropharynx, which serves as the gateway to the respiratory and gastrointestinal tracts. Since lymphoid tissues in the oropharynx (tonsils and adenoids) are known targets for adenovirus replication, we set out to characterize the frequency and complexity of AAV in freshly excised tonsils and adenoids obtained from children ages 2 to 13 years.

Using standard virus culture techniques, including blind passage of cellular lysates after 2 weeks in culture, we were unable to isolate replicating adenovirus from any of the 101 tonsil and adenoid samples. Because replicating adenoviruses were not isolated, we made no further attempt to identify AAV in these cultures. Instead, we prepared total cellular DNA from the 101 samples and then screened for the presence of AAV DNA

sequences by PCR. Using a previously identified region within the AAV *cap* gene as our target (13), we designed and validated a nested, degenerate primer combination that readily detected the published AAV serotypes (Fig. 1A). Control reactions were performed on human cellular DNA that was spiked with plasmid DNA representing *cap* genes of AAV serotypes 1 to 5. Using the optimized primer sets, we achieved a sensitivity of 15 copies in a background of 100 ng of total cellular DNA (data not shown).

With this approach, we identified AAV DNA sequences in 7 of the 101 (7%) tonsil and adenoid samples (Fig. 1B and Table 1). To determine the AAV copy number in these samples, quantitative PCR was also performed on the same total

TABLE 1. Summary of AAV and adenovirus sequence detection in human tissues<sup>a</sup>

Viral DNA found	No. of samples positive					
	T+A (n = 101)	Liver (n = 19)	Spleen (n = 21)	Muscle (n = 15)	Heart (n = 3)	Lung (n = 3)
AAV positive	7	0	1	0	0	1
AdV positive	19	0	0	0	0	0
AAV and AdV positive	2 <sup>b</sup>	0	0	0	0	0

<sup>a</sup> T+A = tonsils and adenoids; AdV = adenovirus.  
<sup>b</sup> Two samples (T17 and T32) contained both AAV and adenovirus sequences.

TABLE 2. AAV sequence relatedness and DNA copy number in pediatric tissues

Sample (age, yr) <sup>a</sup>	Amino acid identity <sup>b</sup> (%)	Nucleotide identity <sup>c</sup> (%)	AAV copies/ $\mu$ g <sup>d</sup>
T17 (2.3)	AAV2 (98.1)	AAV2 (96.0)	560
T32 (5.5)	AAV2 (98.1)	AAV2 (96.0)	210
T40 (2.9)	AAV2 (98.0)	AAV2 (96.1)	6,550
T41 (2.6)	AAV2 (98.0)	AAV2 (96.0)	590
T70 (3.2)	AAV2 (98.4)	AAV2 (96.8)	7,800
T71 (5.8)	AAV2 (98.0)	AAV2 (96.0)	7,600
T88 (4.7)	AAV2 (98.0)	AAV2 (96.5)	33,000
S17 (8)	AAV3 (92.7)	AAV2 (90.0)	200
LG15 (1)	AAV2 (98.4)	AAV2 (96.8)	80

<sup>a</sup> T = tonsil/adenoid; S = spleen; LG = lung. The age of the subject in years is in parentheses.

<sup>b</sup> Percent amino acid identity to the indicated AAV serotype is given in parentheses.

<sup>c</sup> Percent nucleotide identity to the indicated AAV serotype is given in parentheses.

<sup>d</sup> AAV DNA copy number determined using Q-PCR (see Materials and Methods). Values shown are the average of two separate determinations.

cellular DNA. Copy numbers ranged from 210 to 33,000 copies/ $\mu$ g cellular DNA (Table 2).

The same DNA samples were also screened for adenovirus sequences using a PCR primer pair specific for a conserved region in the adenovirus hexon gene. This hexon-based PCR strategy has been shown to detect representative serotypes of all of the known adenovirus subgenera A to F (9). Not surprisingly, 19 tonsil and adenoid samples (19%) scored positive for adenovirus hexon sequences (Table 1). Analysis of the hexon PCR amplicon (301 bp) revealed significant identity (95% to 99%) to human adenovirus serotypes 2 and 5 for the majority (17 of 19) of the positive samples. Two tonsil and adenoid samples (T17 and T32) contained both AAV and adenovirus sequences.

**AAV DNA in other tissues.** To extend the findings described above, we obtained 74 frozen, archived normal tissues from individuals aged 0 to 30 years. The range of tissues and number of samples analyzed are shown in Table 1. Subjects under 0.5 years of age ( $n = 49$ ) were analyzed but were considered unlikely to have been exposed to adenovirus and AAV. In addition, for three of the 74 samples, the age of the subjects could not be determined. Thus, there were 22 samples from individuals aged 0.5 to 30 years available for analysis, and 16 of the 22 were from individuals between the ages of 0.5 and 14 years. Using the same PCR scheme described for tonsils and adenoids, we found AAV DNA in 2 of the 16 (12.5%) samples from children aged 0.5 to 14 years. One was lung from a 1-year-old subject and the other was spleen from an 8-year-old subject. DNA copy numbers were lower in these two samples than in the tonsils and adenoids (Table 2). None of the other 72 samples contained detectable AAV DNA and none of the 74 samples contained adenovirus DNA (Table 1).

**AAV *cap* and *rep* gene sequences from human tissues.** DNA sequence analysis of the 255-bp amplified *cap* gene sequence from the nine positive samples revealed significant homology with the corresponding region of AAV2. To isolate complete AAV capsid genes, we synthesized additional degenerate primers at the 3' ends of the *rep* and *cap* genes (Fig. 1A). These primers were combined with the forward and reverse con-

served sequence primers to amplify 1.8-kb and 1.5-kb PCR products representing the 5' and 3' halves of the *cap* gene. Individual PCR products were cloned and sequenced; a minimum of 4 individual clones were analyzed for each 1.8-kb and 1.5-kb PCR product (intraclone variation was  $\sim$ 0.1%). After a single contiguous sequence was assembled for each tissue isolate, nucleotide sequence alignments revealed significant identity with AAV2 (96 to 97%) for eight sequences. The spleen sequence (S17) was the most divergent and intermediate between serotypes 2 and 3 (Table 2). The complete *rep* coding sequences were subsequently determined for six of the nine isolates (T32, T40, T70, T71, T88, and S17) and all shared  $>99\%$  amino acid identity with AAV2. Three conservative mutations in Rep (for all six isolates) were identified (T183A, V508A, and F619S).

The capsid gene amino acid translation and alignment for all nine clones is shown in Fig. 2. Consistent with the nucleotide sequence analysis, eight of the amino acid sequences shared 98% identity with AAV2. Moreover, the majority of the observed amino acid substitutions (relative to AAV2) found in the seven of the tonsil sequences and the lung sequence were conserved among the individual samples. This suggests that a specific virus isolate was circulating in the local population during the time period (winter 2002 to 2003) of tissue procurement. The majority of the observed amino acid substitutions were located in previously identified hypervariable regions (HVRs) 5 to 7, 9, and 10 (11), all of which were predicted to be exposed on the surface of the virion. Two isolates possessed identical sequences (T41 and T71), and two others (T17 and T32) were nearly identical (two amino acid differences). The deduced phylogenetic relationship among the nine *cap* gene sequences is depicted in Fig. 3, along with previously identified clade B and C viruses (11–13). As expected, seven of the eight sequences clustered closely with each other within AAV2-like clade B.

The spleen isolate (S17) appeared to be intermediate between AAV2 and AAV3 and shared significant identity to the recently described clade C isolates (Fig. 3) (12). Additional sequence analyses (Simplot) revealed two discreet AAV3-like domains (Fig. 4), suggesting that sequential recombination events between AAV2 and AAV3 serotypes generated the S17 isolate. The first AAV3-like region corresponded to amino acids 180 to 235 and overlapped HVR 2' (Fig. 2). The second AAV3-like region encompassed amino acids 466 to 673 and covered HVRs 5 to 11. S17 possessed several surface exposed amino acid clusters in these HVRs that were identical to AAV3.

**Most circulating AAV2-like viruses are not predicted to bind heparin sulfate.** Careful inspection of the AAV capsid sequences described above revealed that none of the eight AAV2-like sequences retained arginine residues at positions 585 and 588. These residues have been shown to be critical for heparin sulfate proteoglycan (HSPG) receptor binding (18, 24). In each of our sequences, R585S and R588T substitutions were observed. To further experimentally define the heparin binding capacity of AAV capsids with these substitutions, we directly measured the ability of two AAV2 capsids derived from tonsil and adenoids (T70 and T88) to bind heparin.

Infectious AAV preparations representing T70 and T88 were generated from full-length molecular clones derived di-

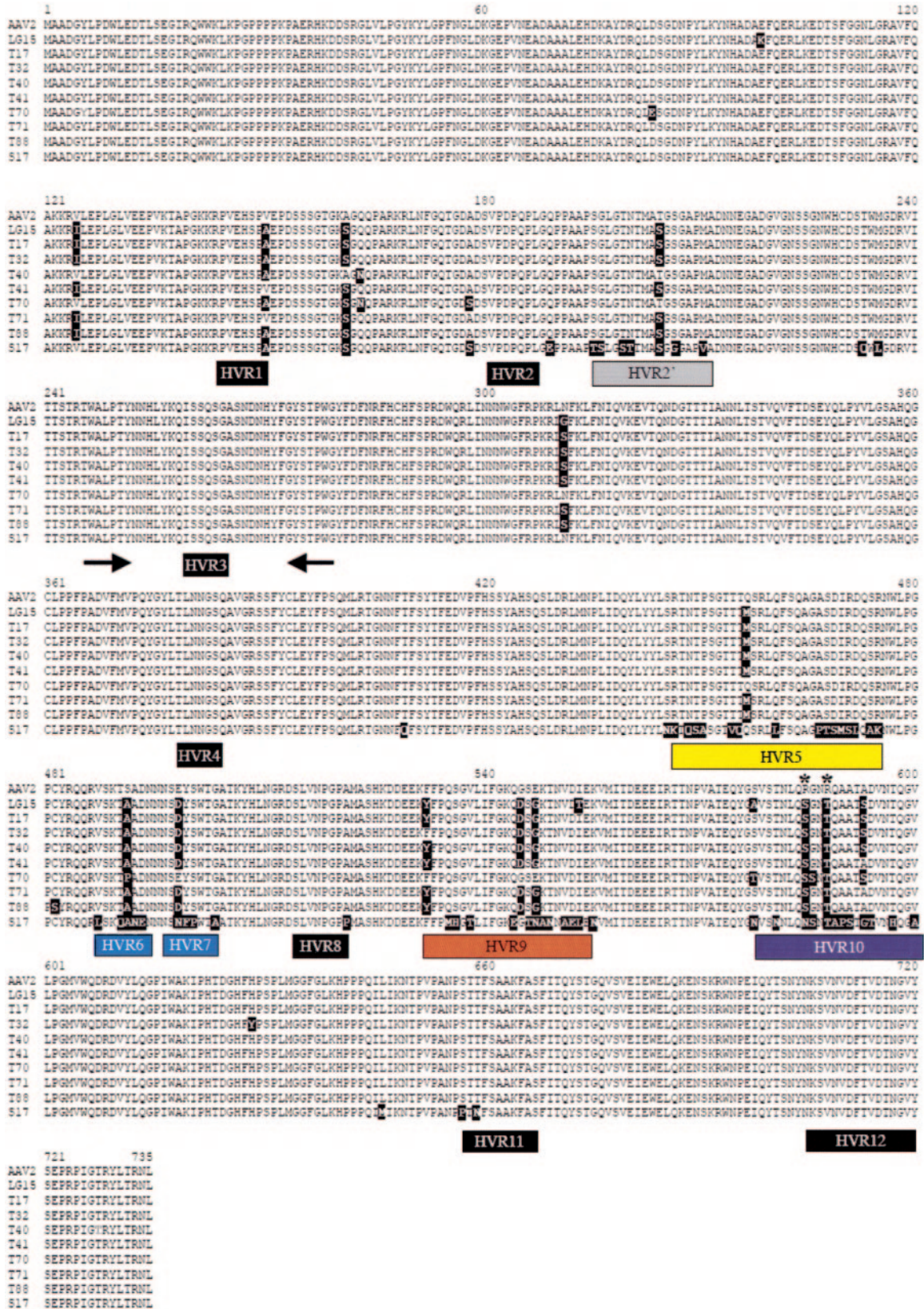


FIG. 2. Predicted VP1 capsid amino acid alignment of AAV2 and novel human AAVs. Diagram shows sequence alignment using the CLUSTAL W program. Black boxes designate amino acid substitutions compared to the AAV2 sequence. The locations of previously identified HVR regions (11) are labeled (HVR 1 to 12), as is an additional region (HVR 2') that possesses several substitutions. Several HRV regions (5 to 7, 9, and 10) are colored to facilitate visualization of these regions onto the known atomic structure of AAV2, while invariant HVRs are labeled with black boxes. The locations of R585S and R588T are starred, and arrows denote the approximate locations of nested primers used to amplify the 255-bp HVR3 fragment.

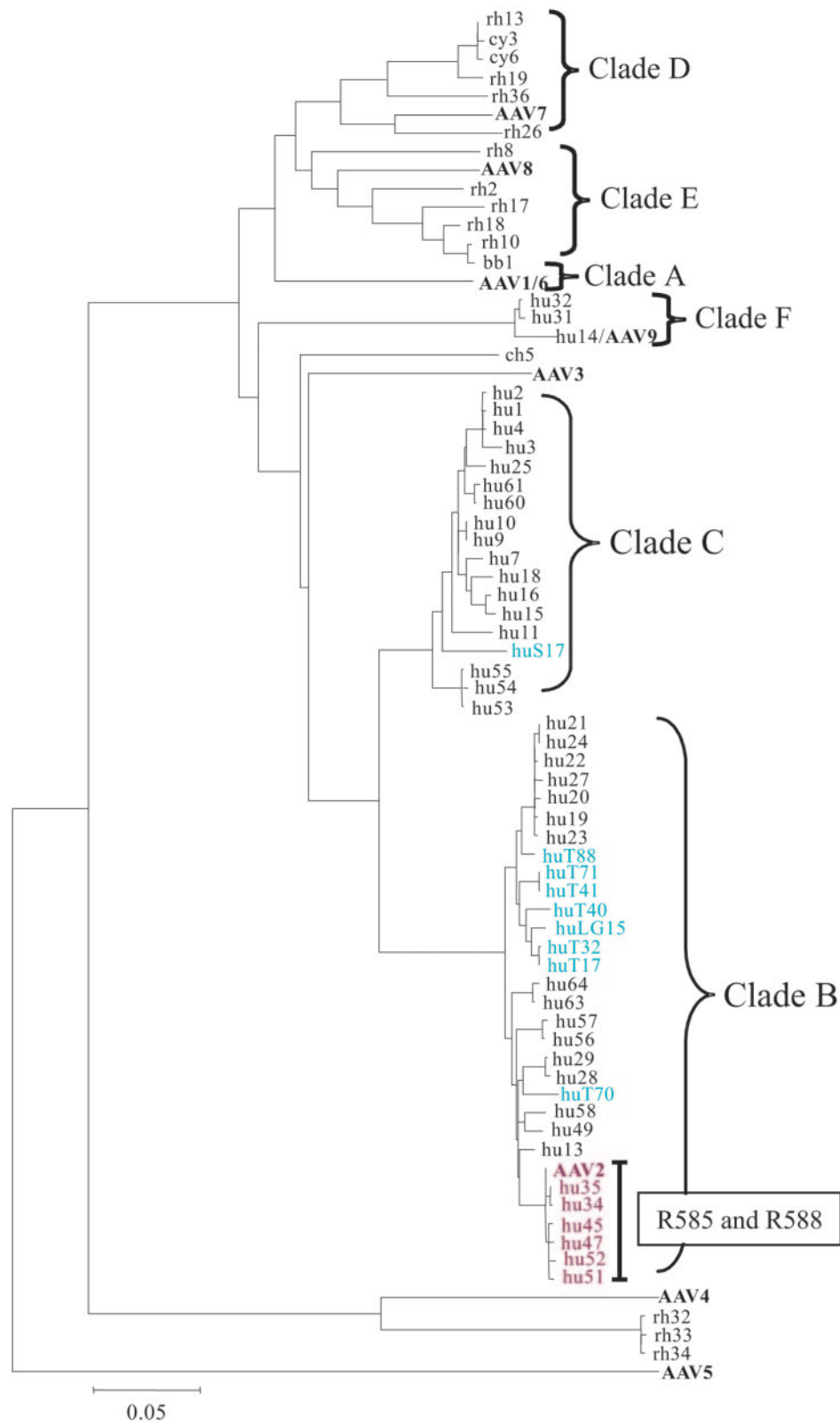


FIG. 3. Phylogenetic analysis of VP1 capsid nucleotide sequences. A neighbor-joining program with a Kimura two-parameter setting was used to derive phylogenetic distances based on 2,200 bp of VP1 sequence. Recently described AAV clade nomenclature (12) was adopted and organized by vertical brackets. The human isolates identified herein are designated in teal type. Due to space restrictions, only a few representative isolates from clades A, D, and E are shown. Sequence isolates are labeled with reference to the source species (bb, baboon; ch, chimpanzee; cy, cynomolgus macaque; hu, human; rh, rhesus macaque). Clade B sequences possessing R585 and R588 amino acids and predicted to bind HSPG efficiently are labeled in red type. The scale for genetic distance is indicated in the bottom left corner.

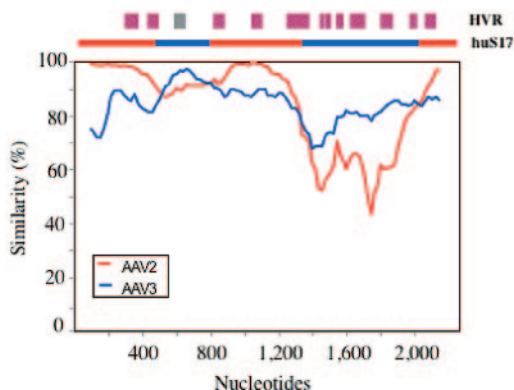


FIG. 4. S17 sequence homology comparison with AAV2 and AAV3. Simplot analyses of similarity percentages of S17 VP1 versus AAV 2 (red) and AAV3 (blue) are shown. Data were plotted within a sliding window of 200 bp, centered on the position plotted, with a step size between data points of 20 bp. Positions containing gaps were excluded from the comparison. The bar on the top shows the predicted composition of the S17 capsid gene. The corresponding positions of the HVRs are labeled as magenta boxes (HVR 2' in gray).

rectly from the tonsil and adenoid tissue. A full description of these and other infectious clones derived from tissues is presented elsewhere (Schnepf et al., in preparation). DNA sequences of the clones were identical to the original *cap* gene nucleotide sequences generated directly from tissues. When applied to a standard heparin chromatography column (HE20-POROS) under low-salt conditions, 93.5% of the total DNase-resistant particles applied to the column were found in the flowthrough or wash buffers (Table 3). In contrast, prototype AAV2 readily bound to the column, with on average 86% of the DNase-resistant particles eluting at 300 mM NaCl (7). Following chromatography, the identity of each virus isolate was further confirmed by amplifying and sequencing a 600-bp portion of viral DNA that was collected from either the

TABLE 3. Heparin sulfate column binding of AAV2 virus preparations

Virus	Nonbound DRP <sup>a</sup>	Bound DRP <sup>b</sup>	Percent Bound
AAV2	$2.4 \times 10^9$	$1.5 \times 10^{10}$	0.86
T70-43	$2.0 \times 10^{10}$	$1.0 \times 10^9$	0.05
T88-41	$1.7 \times 10^{10}$	$1.4 \times 10^9$	0.08

<sup>a</sup> DNase-resistant particle (DRP) copy number was determined using Taqman Q-PCR. Primers and probe were homologous to a conserved cap sequence (see Materials and Methods). Nonbound represent total virus found in wash and flowthrough material. Data shown are the average of two experiments.

<sup>b</sup> Copy number present in pooled 1-ml gradient fractions. Data shown are the average of two experiments.

flowthrough or peak fraction. In each case, the predicted sequence was recovered (data not shown).

In support of the experimental observations noted above, a computer generated electrostatic potential map (Fig. 5) of the VP3 trimer was produced using the recently described atomic structure of AAV2 VP3 (32). AAV2 VP3 monomers contain large regions of strong positive charge (Fig. 5A, circled blue regions) grouped at the threefold axis of symmetry in the VP3 trimer (17). These regions have been implicated in HSPG coreceptor binding due to the collective involvement of 5 basic amino acids (R484, R487, K532, R585, and R588) that map to this region. As a consequence of the serine and threonine substitutions in our sequences at positions 585 and 588, respectively, the model predicted a significant net reduction in overall positive charge (Fig. 5B). Thus, the electrostatic potential mapping data was consistent with failure of AAV derived from T70 and T88 to bind to the heparin column.

We also examined the ability of the S17 isolate (AAV2/3-like) to bind heparin based on the known heparin-binding properties of prototype AAV3 (18). To accomplish this analysis, we exchanged the *cap* coding region in the T70 molecular clone for the S17 *cap* gene. Infectious particles were generated and analyzed for heparin binding as described above. As ob-

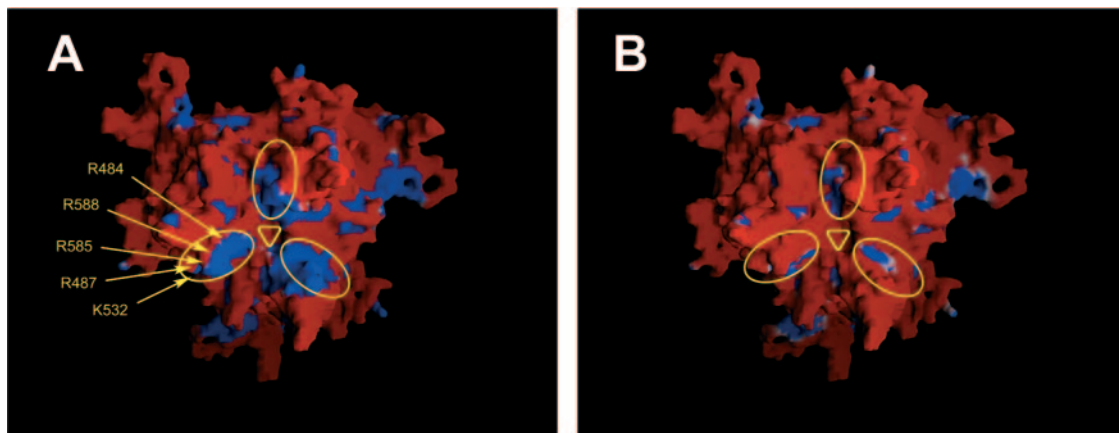


FIG. 5. Surface diagrams of AAV2 trimer atomic models. (A) Electrostatic surface potential of the VP3 AAV2 trimer viewed down the threefold axis (yellow triangle) calculated with GRASP (23) running from negative (red) to positive (blue). Labeled arrows indicate the positions of residues implicated in HSPG binding. (B) Predicted electrostatic surface potential of AAV2 VP3 trimer as a result of R585S and R588T substitutions. Amino acid substitutions were modeled using energy minimization simulations with Quanta (Accelrys, San Diego, CA) prior to generating the electrostatic potential map in GRASP. The surface electrostatic potential scale is the same as depicted in panel A. Highlighted regions denote predicted HSPG coreceptor engagement domains in the VP3 trimer.

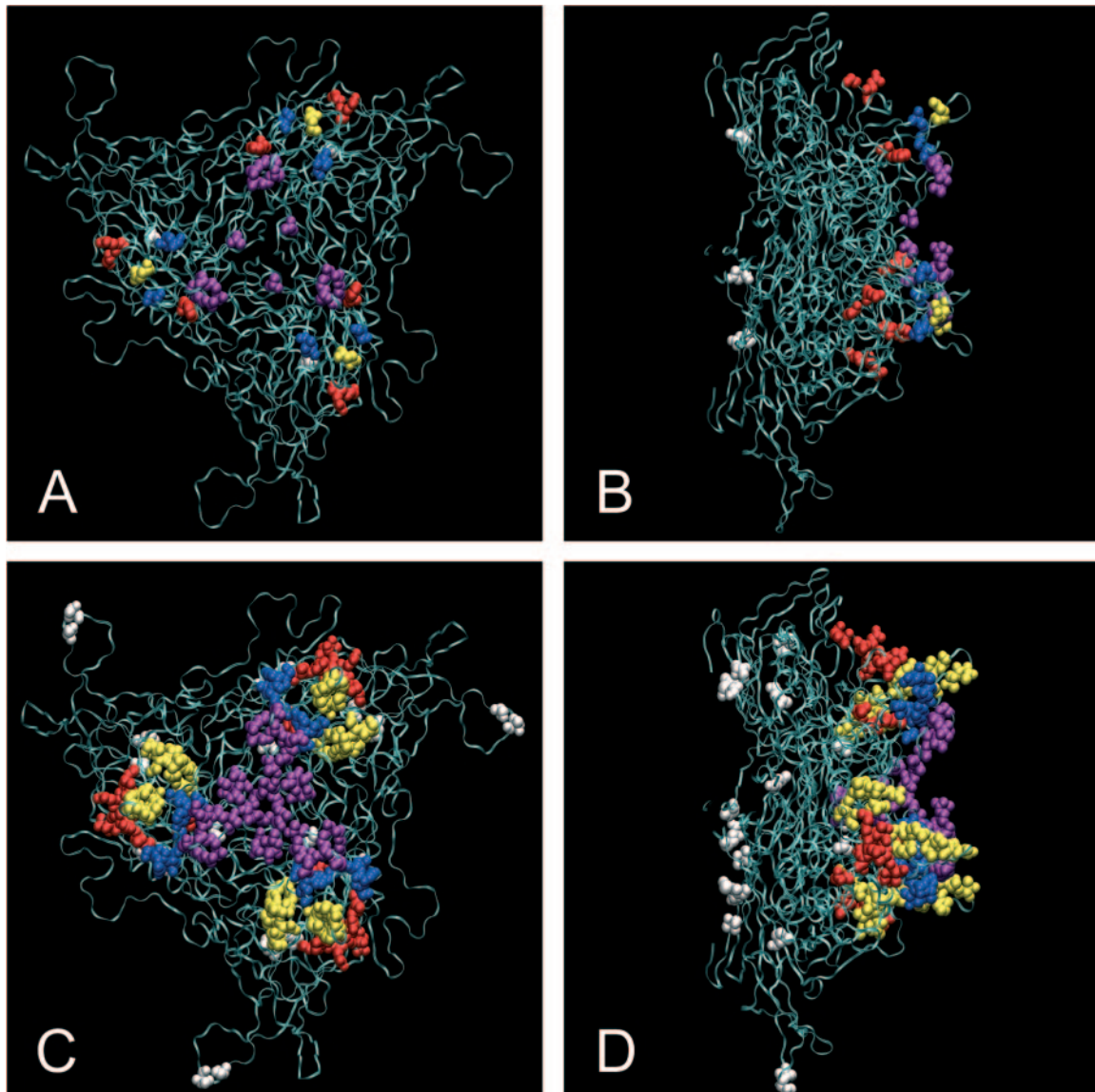


FIG. 6. Ribbon diagrams of atomic models of AAV2 VP3 trimers showing the location of predicted amino acid substitutions in the human AAV isolates. (A) Ribbon drawing viewed down the threefold axis of symmetry of the AAV2 VP3 trimer.  $C_{\alpha}$  backbones for the three VP3 monomers are rendered as teal ribbons. Predicted locations of the observed amino acid substitutions present within the eight AAV2-like sequences are color coded to reflect HVR location (HVR 5 to 7, 9, and 10) within the primary sequence (Fig. 2). White space-filling amino acid substitutions mapped outside the known HVRs. (B) Side view of the predicted location of the observed amino acid substitution demonstrating surface display (right side). (C) Superimposition of observed S17 amino acid substitutions relative to the AAV2 VP3 trimer atomic structure viewed down the threefold axis. (D) Side view of the predicted location of the observed amino acid substitutions in isolate S17 (surface display oriented on right side). Images were generated in NAMD/VMD (UIUC Theoretical Biophysics Group) and rendered using Raster3D.

served for T70 and T88, the S17 capsid also failed to bind to a heparin column (<4% bound) (data not shown).

**Sequence variation in the AAV capsid maps to surface exposed regions.** Given the recently described atomic ribbon structure of AAV2 (25), we were able to map the observed amino acid substitutions in the 8 AAV2-like sequences onto the capsid surface (Fig. 6A and 6B). The majority of amino acid substitutions were located in areas of the capsid predicted to be surface exposed and previously identified as hypervariable regions (HVRs 5 to 7, 9, and 10) (Fig. 3). In contrast, regions predicted to encode core  $\beta$ -barrel domains responsible

for structural integrity were almost invariant. Similarly, the more divergent S17 capsid amino acid sequence was modeled onto the same atomic structure (Fig. 6C and 6D). The majority of the S17 amino acid substitutions also mapped to surface exposed HVR regions 5 to 7, 9, and 10. Interestingly, several HVR 10 substitutions (purple shaded region) were located near the central pore complex that is thought to directly interact with the viral encapsidation complex.

Because of the predilection for surface substitutions, we analyzed synonymous versus nonsynonymous nucleotide substitutions in the *cap* and *rep* gene coding regions for all nine

TABLE 4. Summary of synonymous and nonsynonymous nucleotide substitutions in *rep* and *cap*

Isolate	Substitutions <sup>a</sup>											
	<i>cap</i> HVR (400 bp)				<i>cap</i> non-HVR (1,800 bp)				<i>3' rep</i> (600 bp)			
	s	ns	s/ns	Rate/nt	s	ns	s/ns	Rate/nt	s	ns	s/ns	Rate/nt
LG15	20	11	1.8	7.5	57	6	9.5	3.5	9	4	2.3	2.1
S17	48	58	0.8	25.8	92	9	10.2	5.6	11	2	5.5	2.3
T17	17	10	1.7	6.6	62	4	15.5	3.7	7	3	2.3	1.4
T32	17	10	1.7	6.6	62	4	15.5	3.7	6	3	2.0	1.2
T40	12	10	1.2	5.4	54	3	18.0	3.2	9	3	3.0	1.9
T41	15	8	1.9	5.6	47	4	11.8	2.8	4	3	1.3	1.1
T70	10	6	1.7	3.9	48	6	8.0	3.0	17	4	4.3	2.6
T71	15	8	1.9	5.6	47	4	11.8	2.8	3	3	1.0	0.9
T88	18	9	2.0	6.6	45	5	9	2.8	13	4	3.3	2.4
Avg ± SD			1.6 ± 0.27	8.2 ± 3.9 <sup>b</sup>			12.1 ± 2.8 <sup>c</sup>	3.5 ± 0.6			2.8 ± 1.1	1.8 ± 0.6

<sup>a</sup> s is the number of observed synonymous substitutions in the indicated coding region; ns is the number of observed nonsynonymous substitutions in the indicated coding region; s/ns is the calculated ratio. The rate/nt value represents the number of observed nucleotide substitutions per 100 bp of sequence.

<sup>b</sup> Substitution rate in HVR region was greater than that observed for non-HVR and *3' rep* region,  $P = 0.04$  and  $0.02$ , respectively, using paired  $t$  test.

<sup>c</sup> Ds/Da ratio in non-HVR region was greater than that observed for HVR and *3' rep* region,  $P = 0.00002$  and  $0.0001$ , respectively, using paired  $t$  test.

sequences (Table 4). The synonymous/nonsynonymous ratios were calculated for nucleotides corresponding to (i) surface-exposed HVR, (ii) non-HVR, and (iii) approximately 600 bp of the *3'* end of the *rep* gene. Statistically significant differences in the synonymous/nonsynonymous Ds/Da ratios were observed when we compared the non-HVR *cap* region ratio (average 12.1, synonymous/nonsynonymous) to either the HVR (average 1.6, synonymous/nonsynonymous) or Rep (2.8 average synonymous/nonsynonymous) ratios ( $P = 0.00002$  and  $0.0001$ , respectively). These data indicated a strong bias for conservative synonymous mutations within the  $\beta$ -barrel domains in non-HVR regions.

This conclusion was consistent with scanning mutagenesis experiments that demonstrated severe functional constraints on particle assembly and stability in these capsid core regions (25, 27, 30). We also observed a significantly greater nucleotide substitution rate in the HVR (average of 8.2 substitutions per 100 nucleotides) relative to the non-HVR (average of 3.5 substitutions per 100 nucleotides) and *rep* coding regions (average of 1.8 substitutions per 100 nucleotides) ( $P = 0.04$  and  $0.02$ , respectively). These rates likely reflect the inherent flexibility for amino acid substitutions that each region possesses in that surface-exposed loop domains (that overlap HVR domains) are known to tolerate multiple amino substitutions and remain infectious, while much of the non-HVR regions and *rep* coding regions are functionally constrained.

**Seroreactivity to AAV2.** To further characterize AAV infection in children, we would have liked to evaluate the AAV2 serologic status of our 175 subjects. Unfortunately, we were unable to collect or acquire serum from these individuals. Instead, we screened a separate cohort of 68 individuals (ages 3 to 39) drawn from a local cystic fibrosis clinic. Sera were assayed both for binding antibodies (ELISA) to the AAV2 capsid and for the ability to neutralize AAV2 in vitro (Fig. 7).

If one restricts attention to those individuals aged 14 and younger (same ages as our tonsil and adenoid cohort), the seropositivity rate was 11.5% (3 of 26). Interestingly, one of the three positive samples had a low titer (1:200) in vitro neutralizing activity but did not register as positive in the ELISA. This rate compares favorably with the 7% rate observed when

judged by detection of AAV DNA in tonsils and adenoids. Although the numbers are somewhat smaller (2 of 16), the 11.5% seropositive rate also compares favorably with the 12.5% rate of AAV DNA found in tissues outside the oropharynx. The rates estimated from DNA detection would represent a minimum based on two assumptions: (i) not every infected individual would harbor AAV in the oropharynx or other organs and (ii) there was a reasonable chance of sampling error.

After the age of 14 years, the number of seropositive samples rose dramatically to 76% (32 of 42). The correlation with age was significant ( $r = 0.464$ ,  $P < 0.001$  by Pearson correlation). Only 24% (10 of 42) of the sera from those older than 14 years possessed significant in vitro neutralizing activity. Considering all ages, only 16% (11 of 68) of the samples had significant (titer  $> 1:100$ ) in vitro neutralizing activity. There were six samples from those over 14 years that had significant ELISA activity (OD  $> 0.4$ ) that did not mediate significant in vitro neutralization. Such ELISA reactivity probably represented antibodies that bound to the capsid but did not neutralize the virus.

## DISCUSSION

Renewed interest in the biology of wild-type AAV infection prompted us to pursue prospective molecular characterization of AAV infection in a cohort of children undergoing elective surgical excision of tonsils and adenoids during the winter season. Although we were unable to recover infectious adenovirus using standard in vitro culture techniques, we did identify AAV DNA sequences in 7% (7 of 101) of the samples analyzed. We also found adenovirus DNA in 19% of the samples, and in two cases, AAV and adenovirus DNA were present in the same tissue. One might speculate that these latter two samples were recovered from individuals recently infected with AAV and adenovirus, although there are clearly other possibilities to explain these observations.

**Molecular epidemiology of AAV infection in children.** Our discovery of AAV2 DNA in tonsils and adenoids from children was not unexpected. AAV2 appears to be the most prevalent of the human AAV serotypes (3, 4, 17, 31), and importantly, our

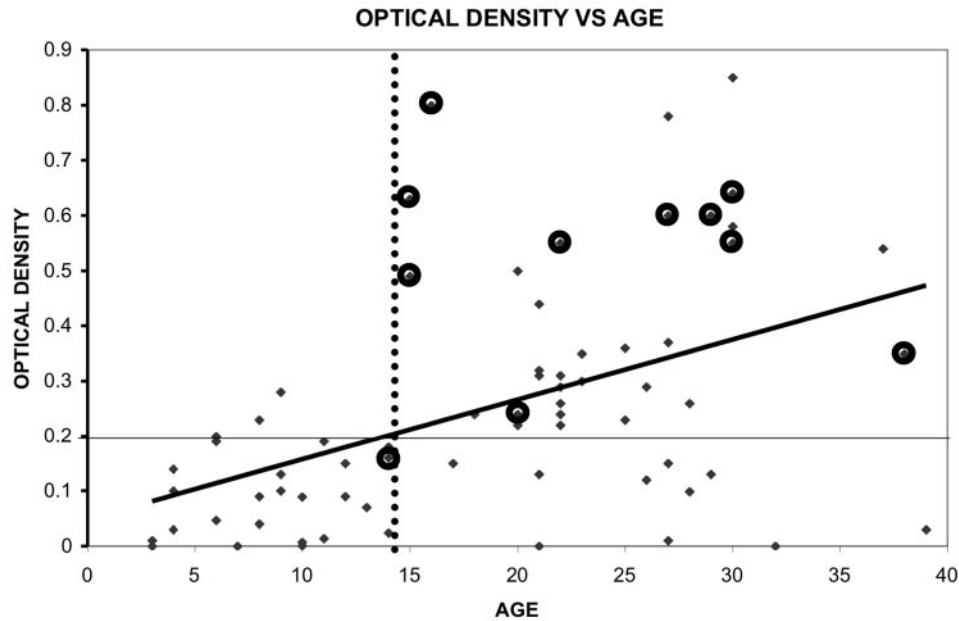


FIG. 7. Serology of AAV infection as a function of age.  $OD_{450}$  values from a standard ELISA (see Materials and Methods) are plotted versus the age of the subject. Sera were tested at a 1:100 dilution. OD values below 0.2 (thin solid line) were considered negative. The same sera were tested for neutralization activity against AAV2 (see Materials and Methods). Data points that are circled represent samples that had neutralization titers of  $>1:100$ .

cohort was temporally (winter) and geographically (central Ohio) restricted. Moreover, the portal of entry for AAV would be expected to parallel that of adenovirus. Thus, the oropharynx would be a fertile field from which to harvest AAV (and adenovirus).

While the DNA sequences recovered from these children were similar to the prototype AAV2 and to each other, two significant observations emerged from the sequence analyses. First, the majority of the observed amino acid substitutions were located in previously defined hypervariable regions (11) predicted to be exposed on the surface of the virion (Fig. 6A and 6B). This observation also extended to the more distantly related S17 (AAV2/3) capsid sequence (Fig. 6C and 6D). These data, considered together with the observed nucleotide substitution patterns, suggest that the HVR domains are structurally flexible and possess the capacity to evolve, perhaps in response to host immune pressure. The fact that *rep* gene sequences from the same tissues were nearly invariant further supported the idea of functional constraints. The identification of the S17 AAV2/3 chimeric sequence confirmed earlier published observations (12), and suggested another mechanism (intermolecular recombination) involved in AAV evolution.

Second, and perhaps most intriguing, was the discovery that none of the seven AAV2-like sequences derived from tonsils and adenoids were predicted to bind HSPG. This was also true for the AAV2-like sequence recovered from lung tissue. To confirm that other compensatory mutations that would restore HSPG binding in the capsid had not occurred, we generated infectious molecular clones from 2 of the tissue samples and demonstrated they indeed did not bind HSPG (Table 3). To extend this observation, we examined other available AAV2-like sequences (12) and noted that 77% of the 31 currently

identified clade B AAV2-like sequences lack R585 and R588. Thus, these data suggest that preponderance of AAV2-like isolates do not bind HSPG, and that HSPG is not required for natural infection in humans. This notion is consistent with the fact that most other serotypes of AAV do not bind HSPG.

**Seroepidemiology of AAV2 infection in children.** To see if the rates of AAV infection as judged by DNA isolation (7 to 12%) were consistent with serological estimates, we analyzed a set of sera from children in the same geographic locale. Ideally, we would have also analyzed sera from the cohort of children who donated tissue samples, but we were unable to obtain serum from these children. In those children 14 years and younger (same ages as our tonsil and adenoid cohort), the seropositivity rate was roughly 12% and thereby confirmed our estimates derived from AAV DNA detection in tissues.

**AAV disseminates beyond the portal of entry.** Previously published work has described the presence of AAV DNA in multiple adult human tissues (12). We have now extended these findings to children with the discovery of AAV DNA in tissues within and beyond the probable portal of entry (oropharynx). In children ages 0.5 to 14 years, we found AAV DNA in 2 (lung and spleen) of the 16 nonoropharyngeal tissue samples (12.5%) available for analysis. Although the number of samples analyzed was smaller, the percentage containing AAV DNA compared favorably (12.5% versus 18%) with data from adults (12).

To extend beyond the portal entry, infectious agents generally use one of 3 pathways: direct (contiguous) spread, lymphatic, or bloodstream. In our samples, AAV could have easily reached the lung by contiguous spread from the oropharynx, either with or without adenovirus. To reach the spleen, however, the route of viral spread was almost certainly hematog-

enous, in the form of free or cell-associated (perhaps leukocytes) virus.

**Biology of wild-type AAV infection.** The now emerging picture is that AAV infection of humans is more complex than previously appreciated. It has been known for decades that AAV is not associated with disease or pathology, and that infection incidence in humans generally parallels that of adenovirus. Not surprisingly, infectious AAV has been isolated from sites where adenovirus is traditionally recovered, including the gastrointestinal tract (1). More recently, it has been appreciated that AAV DNA can be found in many human tissues, including liver, muscle, lymph nodes, leukocytes, kidney, and cervical tissues (11, 12, 15, 16, 28, 29), and now tonsils and adenoids.

Considered together, these data suggest the following biologic scenario. AAV most likely enters the body through the oropharynx in association with adenovirus. Replication ensues (with adenovirus help) and new AAV particles are formed and released from infected cells in the oropharynx. Secondary rounds of replication in newly infected cells follow, again creating new waves of AAV particles. Such rounds of replication would continue until the host immune system responds and blunts the infectious process. By the time replication is controlled, AAV has had the opportunity to spread to the lungs (contiguous) and through the bloodstream to more distant sites. This scenario does not address the rare AAV5 serotype, which apparently enters the body through the genital tract (1, 14). However, there remains only a single isolate of AAV5, and recent studies of human tissues have failed to find AAV5-like DNA.

While the events envisioned above are entirely plausible, many important questions regarding the *in vivo* biology of natural AAV infection remain unanswered. For example, while AAV and adenovirus appear to be linked early in the infectious process, there appears to be an unlinking sometime during and following dissemination, allowing AAV DNA (but not adenovirus) to persist in organs and tissues outside the oropharynx. It is formally possible that AAV DNA persists by integrating in target cells on chromosome 19 (AAVS1), but this has not been demonstrated *in vivo*. Moreover, the target cells for AAV replication and persistence have not been identified, nor have the specific cellular receptors for viral attachment been defined.

It should be remembered that even with widespread dissemination, AAV has not been found to cause disease or pathology. Nonetheless, the effect of prior (or subsequent) wild-type AAV infection on gene transfer mediated by recombinant AAV vectors is unknown, and a more thorough understanding of the natural AAV infectious process is needed.

#### ACKNOWLEDGMENTS

We thank Janet Weslow-Schmidt, Xiaobin Li, and Will Ray for expert technical and computer modeling assistance. We also thank Greg Wiet, Sue Hammond, and Janet Berry for assistance with tonsil tissue procurement, Steve Qualman and the Cooperative Human Tissue Network, and the CCRI Viral Vector and Sequencing Core Laboratories for help with DNA sequencing and virus production.

This work was funded by National Institutes of Health (NIAID/DAIDS) grant 2-PO1-AI56354.

#### REFERENCES

- Bantel-Schaal, U., and H. zur Hausen. 1984. Characterization of the DNA of a defective human parvovirus isolated from a genital site. *Virology* **134**:52–63.
- Blacklow, N. R., M. D. Hoggan, A. Z. Kapikian, J. B. Austin, and W. P. Rowe. 1968. Epidemiology of adenovirus-associated virus infection in a nursery population. *Am. J. Epidemiol.* **88**:368–378.
- Blacklow, N. R., M. D. Hoggan, and W. P. Rowe. 1967. Isolation of adenovirus-associated viruses from man. *Proc. Natl. Acad. Sci. USA* **58**:1410–1415.
- Blacklow, N. R., M. D. Hoggan, and W. P. Rowe. 1968. Serologic evidence for human infection with adenovirus-associated viruses. *J. Natl. Cancer Inst.* **40**:319–327.
- Blacklow, N. R., M. D. Hoggan, M. S. Sereno, C. D. Brandt, H. W. Kim, R. H. Parrott, and R. M. Chanock. 1971. A seroepidemiologic study of adenovirus-associated virus infection in infants and children. *Am. J. Epidemiol.* **94**:359–366.
- Brandt, C. D., H. W. Kim, A. J. Vargosko, B. C. Jeffries, J. O. Arrobio, B. Rindge, R. H. Parrott, and R. M. Chanock. 1969. Infections in 18,000 infants and children in a controlled study of respiratory tract disease. I. Adenovirus pathogenicity in relation to serologic type and illness syndrome. *Am. J. Epidemiol.* **90**:484–500.
- Clark, K. R., X. Liu, J. P. McGrath, and P. R. Johnson. 1999. Highly purified recombinant adeno-associated virus vectors are biologically active and free of detectable helper and wild-type viruses. *Hum. Gene Ther.* **10**:1031–1039.
- Clark, K. R., F. Voulgaropoulou, and P. R. Johnson. 1996. A stable cell line carrying adenovirus-inducible rep and cap genes allows for infectivity titration of adeno-associated virus vectors. *Gene Ther.* **3**:1124–1132.
- Elnifro, E. M., R. J. Cooper, P. E. Klapper, and A. S. Bailey. 2000. PCR and restriction endonuclease analysis for rapid identification of human adenovirus subgenera. *J. Clin. Microbiol.* **38**:2055–2061.
- Eries, K., P. Sebokova, and J. R. Schlehofer. 1999. Update on the prevalence of serum antibodies (IgG and IgM) to adeno-associated virus (AAV). *J. Med. Virol.* **59**:406–411.
- Gao, G., M. R. Alvira, S. Somanathan, Y. Lu, L. H. Vandenberghe, J. J. Rux, R. Calcedo, J. Sanmiguel, Z. Abbas, and J. M. Wilson. 2003. Adeno-associated viruses undergo substantial evolution in primates during natural infections. *Proc. Natl. Acad. Sci. USA* **100**:6081–6086.
- Gao, G., L. H. Vandenberghe, M. R. Alvira, Y. Lu, R. Calcedo, X. Zhou, and J. M. Wilson. 2004. Clades of Adeno-associated viruses are widely disseminated in human tissues. *J. Virol.* **78**:6381–6388.
- Gao, G. P., M. R. Alvira, L. Wang, R. Calcedo, J. Johnston, and J. M. Wilson. 2002. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl. Acad. Sci. USA* **99**:11854–11859.
- Georg-Fries, B., S. Biederlack, J. Wolf, and H. zur Hausen. 1984. Analysis of proteins, helper dependence, and seroepidemiology of a new human parvovirus. *Virology* **134**:64–71.
- Grossman, Z., E. Mendelson, F. Brok-Simoni, F. Mileguir, Y. Leitner, G. Rechavi, and B. Ramot. 1992. Detection of adeno-associated virus type 2 in human peripheral blood cells. *J. Gen. Virol.* **73**:961–966.
- Han, L., T. H. Parmley, S. Keith, K. J. Kozlowski, L. J. Smith, and P. L. Hermonat. 1996. High prevalence of adeno-associated virus (AAV) type 2 rep DNA in cervical materials: AAV may be sexually transmitted. *Virus Genes* **12**:47–52.
- Hildinger, M., A. Auricchio, G. Gao, L. Wang, N. Chirmule, and J. M. Wilson. 2001. Hybrid vectors based on adeno-associated virus serotypes 2 and 5 for muscle-directed gene transfer. *J. Virol.* **75**:6199–6203.
- Kern, A., K. Schmidt, C. Leder, O. J. Muller, C. E. Wobus, K. Bettinger, C. W. Von der Lieth, J. A. King, and J. A. Kleinschmidt. 2003. Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. *J. Virol.* **77**:11072–11081.
- Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*. **17**:1244–1245.
- Lo, W. D., G. Qu, T. J. Sferra, R. Clark, R. Chen, and P. R. Johnson. 1999. Adeno-associated virus-mediated gene transfer to the brain: duration and modulation of expression. *Hum. Gene Ther.* **10**:201–213.
- Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* **73**:152–160.
- Lukashov, V. V., and J. Goudsmit. 2001. Evolutionary relationships among parvoviruses: virus-host coevolution among autonomous primate parvoviruses and links between adeno-associated and avian parvoviruses. *J. Virol.* **75**:2729–2740.
- Nicholls, A., K. A. Sharp, and B. Honig. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**:281–296.
- Opie, S. R., K. H. Warrington, Jr., M. Agbandje-McKenna, S. Zolotukhin, and N. Muzyczka. 2003. Identification of amino acid residues in the capsid proteins of adeno-associated virus type 2 that contribute to heparan sulfate proteoglycan binding. *J. Virol.* **77**:6995–7006.

25. **Rabinowitz, J. E., W. Xiao, and R. J. Samulski.** 1999. Insertional mutagenesis of AAV2 capsid and the production of recombinant virus. *Virology* **265**:274–285.
26. **Samulski, R. J.** 2003. AAV vectors, the future workhorse of human gene therapy. *Ernst Schering Res. Found Workshop* **43**:25–40.
27. **Shi, W., G. S. Arnold, and J. S. Bartlett.** 2001. Insertional mutagenesis of the adeno-associated virus type 2 (AAV2) capsid gene and generation of AAV2 vectors targeted to alternative cell-surface receptors. *Hum. Gene Ther.* **12**: 1697–1711.
28. **Tobiasch, E., M. Rabreau, K. Geletneky, S. Larue-Charlus, F. Severin, N. Becker, and J. R. Schlehofer.** 1994. Detection of adeno-associated virus DNA in human genital tissue and in material from spontaneous abortion. *J. Med. Virol.* **44**:215–222.
29. **Walz, C. M., T. R. Anisi, J. R. Schlehofer, L. Gissmann, A. Schneider, and M. Muller.** 1998. Detection of infectious adeno-associated virus particles in human cervical biopsies. *Virology* **247**:97–105.
30. **Wu, P., W. Xiao, T. Conlon, J. Hughes, M. Agbandje-McKenna, T. Ferkol, T. Flotte, and N. Muzyczka.** 2000. Mutational analysis of the adeno-associated virus type 2 (AAV2) capsid gene and construction of AAV2 vectors with altered tropism. *J. Virol.* **74**:8635–8647.
31. **Xiao, W., N. Chirmule, S. C. Berta, B. McCullough, G. Gao, and J. M. Wilson.** 1999. Gene therapy vectors based on adeno-associated virus type 1. *J. Virol.* **73**:3994–4003.
32. **Xie, Q., W. Bu, S. Bhatia, J. Hare, T. Somasundaram, A. Azzi, and M. S. Chapman.** 2002. The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc. Natl. Acad. Sci. USA* **99**:10405–10410.
33. **Zolotukhin, S., B. J. Byrne, E. Mason, I. Zolotukhin, M. Potter, K. Chesnut, C. Summerford, R. J. Samulski, and N. Muzyczka.** 1999. Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther.* **6**:973–985.