

# Cloning of an Avian Adeno-Associated Virus (AAAV) and Generation of Recombinant AAV Particles

Ioannis Bossis and John A. Chiorini\*

*Gene Therapy and Therapeutics Branch, NIDCR, NIH, Bethesda, Maryland 20892*

Received 2 December 2002/Accepted 28 March 2003

Recent studies have proposed that adeno-associated viruses (AAVs) are not evolutionarily linked to other mammalian autonomous parvoviruses but are more closely linked to the autonomous parvoviruses of birds. To better understand the relationship between primate and avian AAVs (AAAVs), we cloned and sequenced the genome of an AAAV (ATCC VR-865) and generated recombinant AAAV particles. The genome of AAAV is 4,694 nucleotides in length and has organization similar to that of other AAVs. The entire genome of AAAV displays 56 to 65% identity at the nucleotide level with the other known AAVs. The AAAV genome has inverted terminal repeats of 142 nucleotides, with the first 122 forming the characteristic T-shaped palindromic structure. The putative Rep-binding element consists of a tandem (GAGY)<sub>4</sub> repeat, and the putative terminal resolution site (*trs*), CCGGT/CG, contains a single nucleotide substitution relative to the AAV<sub>2</sub> *trs*. The Rep open reading frame of AAAV displays 50 to 54% identity at the amino acid level with the other AAVs, with most of the diversity clustered at the carboxyl and amino termini. Comparison of the capsid proteins of AAAV and the primate dependoviruses indicate that divergent regions are localized to surface-exposed loops. Despite these sequence differences, we were able to produce recombinant AAAV particles carrying a *lacZ* reporter gene by cotransfection in 293T cells and were able to examine transduction efficiency in both chicken primary cells and several cell lines. Our findings indicate that AAAV is the most divergent AAV described to date but maintains all the characteristics unique to the genera of dependovirus.

Adeno-associated viruses (AAVs) are small, nonpathogenic parvoviruses that require coinfection with a helper virus, such as adenovirus or herpesvirus, for productive infection (2). To date, eight AAV isolates (AAV types 1 to 8 [AAV1 to -8]) have been characterized and sequenced (2, 4, 19, 20, 25, 32, 51, 56), with AAV2 having been the most extensively studied.

AAV virions are approximately 20 to 25 nm in diameter and are composed of a mixture of assembled proteins (VPs) that encapsidate a linear ~4.7-kb single-stranded DNA (ssDNA) of plus or minus polarity (7, 43). The genome of AAVs is flanked by inverted terminal repeats (ITRs), which in the case of AAV<sub>2</sub> are 145 nucleotides in length. The ITR is organized as three interrupted palindromes that can fold in an energetically favored T-shaped hairpin structure, which can exist in two orientations, termed flip and flop (42). The ITRs serve as origin of replication and contain *cis*-acting elements required for rescue, integration, excision from cloning vectors, and packaging (41, 42, 49, 58).

The genetic map of the AAVs has been derived primarily from studies of AAV2 but is conserved in all serotypes (26, 27, 29, 36, 42, 45, 46, 58, 60, 64). Two major open reading frames (*rep* and *cap* ORFs) and three transcriptional active promoters (*P*<sub>5</sub>, *P*<sub>19</sub>, and *P*<sub>40</sub>) have been identified in the genome of AAV2. The *P*<sub>5</sub> and *P*<sub>19</sub> promoters encode the nonstructural replication proteins Rep78 and Rep68 and Rep52 and Rep40, respectively. Due to differential splicing, Rep78 and Rep52 have different C termini from Rep68 and Rep40. Transcription initiation from two promoters results in Rep78 and Rep68 having

different N termini from Rep52 and Rep40. The *P*<sub>40</sub> promoter transcribes two alternatively spliced mRNAs. The major mRNA species encodes the major capsid protein VP3 from a conventional AUG codon and the minor capsid protein VP2 from an upstream in-frame ACG codon. The minor mRNA species encodes the entire *cap* ORF to produce the minor capsid protein VP1 (47). VP1, VP2, and VP3 are found in a ratio of 1 to 1 to 10, respectively, and this stoichiometry is generated by the high abundance of one of the mRNA species and the low translation efficiency from an ACG codon in the case of VP2 (14, 47, 55). Previous studies have indicated that VP2 and VP3 are sufficient for particle formation and accumulation of encapsidated ssDNA progeny, while VP1 is required for assembly of highly infectious particles (63, 64).

All four Rep proteins possess NTP binding activity, DNA helicase activity, and nuclear localization sequences; however, only Rep78 and Rep68 possess DNA binding ability (33, 34, 66). Mutant AAVs defective for the synthesis of the small Rep proteins (Rep52 and Rep40) are able to replicate DNA, but no ssDNA progeny is encapsidated (16). The ability of Rep78 and Rep68 to bind and nick DNA in a sequence- and strand-specific manner inside the ITR is essential in every phase of the AAV life cycle, namely, DNA replication, AAV gene expression, rescue from the integrated state, and self-excision from cloning vectors (29, 35, 44). Nicking of the DNA within the ITR at the terminal resolution site (*trs*) requires binding of Rep78 and Rep68 proteins to a motif composed of tandem repeats of GAGY.

Among AAV serotypes, AAV1, AAV4, AAV7, and AAV8 are believed to be of simian origin, while AAV2, AAV3, and AAV5 are from humans. AAV6 was found in a human adenovirus preparation and is very similar to AAV1. AAVs have

\* Corresponding author. Mailing address: NIH 10/1N113, 10 Center Dr., MSC 1190, Bethesda, MD 20892-1190. Phone: (301) 496-4279. Fax: (301) 402-1228. E-mail: Jchiorini@dir.nidcr.nih.gov.

also been reported in other mammalian species including canines, bovines, ovines, and equines (8). An avian AAV (AAAV) was first isolated from the Olson strain of quail bronchitis adenovirus (68). It was later found that 50% of adenoviral field isolates from chickens in the United States and Ireland contained AAVs serologically indistinguishable from the initial isolate (24). The AAV was found to be 20 nm in diameter, was serologically distinct from AAV1–4, did not agglutinate erythrocytes from several species tested, and required adenovirus or herpesvirus for replication (5, 68). In addition, AAV was found to inhibit replication of several avian adenoviruses and herpesviruses (5, 52, 53). Physicochemical studies revealed that the capsid of AAV consists of three VP proteins similar to those of other AAVs. The buoyant density of AAV in CsCl gradients (1.39 to 1.44 g/cm<sup>3</sup>) is similar to what has been reported for all AAVs (6, 30, 68). One study (30) also provided a limited restriction endonuclease map of AAV.

The ability of AAV vectors to infect dividing and nondividing cells and establish long-term transgene expression and the lack of pathogenicity have made them attractive for use in gene therapy applications. Recent evidence has indicated lack of cross competition in binding experiments, suggesting that each AAV serotype may have a distinct mechanism of cell entry. Comparison of the *cap* ORFs from different serotypes has identified blocks of conserved and divergent sequences, with most of the latter residing on the exterior of the virion, thus explaining the altered tissue tropism among serotypes (19–21, 48, 56). Vectors based on new AAV serotypes may have different host ranges and different immunological properties, thus allowing for most efficient transduction in certain cell types. In addition, characterization of new serotypes will aid in identifying viral elements required for altered tissue tropism.

Serological studies have provided evidence of AAV infection in humans (69). Six percent of an unselected adult population was found positive for antibody to AAV by agar gel precipitation (AGP), and 15.6% was positive by virus neutralization. Fourteen percent of poultry workers (industry or research) were positive for AAV antibody by AGP and 66% were positive by virus neutralization. In the same studies, no cross-reaction was noted by AGP when antiserum to AAV was reacted against primate antigens of serotypes 1 to 4 or when antiserum to AAV serotypes 1 to 4 were reacted against AAV antigen. In addition, antiserum prepared against primate AAV1–4 did not neutralize the AAV. These results suggest that AAV is a distinct serotype and that infections are not restricted to avian species but are found in the adult human population.

Based on the genome organization and sequence homology among insect densovirus, rodent parvovirus, and human dependovirus, it has been previously proposed that these viruses may have diverged from a common ancestor and evolved strictly in their hosts (3). However, the high sequence homology between avian autonomous parvovirus and primate AAVs and the epidemiological documentation of AAV transmission to humans provide evidence for host-independent evolution of at least some parvovirus genera. To better understand the relationship between the avian and the primate AAVs, the complete viral genome of AAV was cloned and sequenced and used to generate recombinant viral particles.

## MATERIALS AND METHODS

**Cell culture and virus propagation.** 293T and COS cells were maintained in Iscove's modified Eagle medium and AMEM, respectively DF1 cells (spontaneously immortalized chicken embryonic fibroblasts), QNR cells (quail neuroretinal cells), A549 cells, and primary chicken embryonic fibroblasts (CEF) were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Primary chicken embryonic kidney cells were maintained in  $\beta$ -mercaptoethanol supplemented with 10% fetal bovine serum (FBS). Primary chicken pituitary cells were maintained in DMEM supplemented with 5% horse serum. QT6 cells (quail fibrosarcoma) were maintained in Ham's F12K supplemented with 10% FBS. LMH cells (chicken hepatoma cells) were maintained in Waymouth's medium supplemented with 10% FBS. DT-90 (chicken lymphoblastoma) cells were maintained in DMEM supplemented with 15% FBS, 5% chicken serum, and 0.015%  $\beta$ -mercaptoethanol. Human primary fibroblasts were obtained from Clonetics and maintained in serum-free proprietary medium supplied by the manufacturer. AAV (ATCC VR-865) was propagated in 10-day-old Spafas pathogen-free embryonated chicken eggs coinfecting with the Phelps strain of fowl adenovirus type 1 (FAV1; ATCC VR-486). AAV at 10<sup>4</sup> to 10<sup>7</sup> and FAV1 at 10<sup>5</sup> infectious particles in saline were simultaneously injected into the chorioallantoic cavity of eggs and incubated for 96 h at 37°C. At the end of the incubation, allantoamniotic fluids were harvested and clarified by centrifugation at 6,000  $\times$  g for 10 min.

**Viral DNA isolation, cloning, and sequencing.** Virus from infected clarified allantoamniotic fluids was precipitated by centrifugation at 100,000  $\times$  g for 2 h. The supernatant was discharged, and the virus-containing pellet was resuspended in proteinase K digestion buffer (50 mM Tris [pH 8], 20 mM EDTA, 0.5% sodium dodecyl sulfate [SDS], 200  $\mu$ g of proteinase K per ml) and incubated at 45°C for 2 h. Following a phenol-chloroform extraction and ethanol precipitation, the viral DNA was resuspended in Tris-EDTA buffer containing 0.1 M NaCl. The single-stranded viral DNA was annealed by heating to 95°C for 5 min followed by slow cooling to 65°C for 6 h. The annealed viral DNA was separated electrophoretically in 1% agarose gel, and the double-stranded AAV DNA of approximately 4.7 kb was excised and purified with a gel extraction kit (Qiagen). The viral DNA was further processed to fill in the ends by treating with DNA polymerase (Klenow fragment) at 37°C for 15 min in the presence of deoxynucleoside triphosphates. The whole genome was then blunt end cloned in the pPCR-script cloning vector containing the *lacZ* gene, allowing blue-white screening of ampicillin-resistant colonies (Stratagene). Colonies that contained large inserts (4.7 kb) were initially screened by restriction digestion, and three clones were selected for sequencing. No sequence differences were found in these three clones. The sequence of the entire genome (except ITRs) was determined by using an ABI 373A automated sequencer and FS dye terminator chemistry (ABI). Due to the high degree of secondary structure, ITRs were sequenced by isothermal noncycling sequencing chemistry by using radiolabeled dCTP (Epicentre). One of the clones (pAAAV) that contained the entire consensus sequence of AAV was further used to generate packaging and vector plasmids for construction of recombinant AAV (rAAAV) virus. The complete DNA sequence of AAV has been submitted to GenBank (accession number AY186198).

**Sequence analysis.** DNA and protein sequence alignments were performed by using the Clustal W multiple sequence alignment tool of the Biology Workbench web-based software (SDSC). Promoter, transcription initiation, and splice sites were predicted by using the Neural Network Promoter Prediction web-based software (BDGP). The presence of potential transcription binding sites was analyzed with the MatInspector computer program (54). Putative motifs in the Rep proteins were identified with the BLIMPS program that searches for motifs in the Blocks protein database (28).

**Southern blot hybridization.** The ability of pAAAV to support self-excision, packaging, and generation of nuclease-resistant wild-type (wt) AAV particles was examined. 293T cells seeded in 6-well plates were transfected by using calcium phosphate coprecipitation with pAAAV alone, pAAAV plus pAd12 (a helper plasmid containing the E2 and E4 ORFs and VA RNAs of Ad5), and pAAAV plus infection with Ad5. In addition, LMH cells seeded in gelatin-coated 6-well plates were similarly transfected with pAAAV alone or with pAAAV plus infection with FAV1. After 48 h, clarified lysates were prepared by using three freeze-thaw cycles and centrifugation at 3,800  $\times$  g for 20 min. The lysate (~100  $\mu$ l) was treated with 5 U of DNase for 2 h at 37°C to remove vector and un packaged progeny. Subsequently, the solutions were adjusted to contain 20 mM EDTA (pH 8), 0.5% SDS, and 200  $\mu$ g proteinase K per ml and incubated at 45°C for 2 h. After one phenol-chloroform extraction, nucleic acids were precipitated with the addition of an equal volume of isopropanol, and the pellet was resuspended in 30  $\mu$ l of Tris-EDTA buffer containing 0.1 M NaCl. The

samples were heated to 95°C for 5 min, slowly cooled down to 65°C, and incubated for 5 h. After electrophoresis and blotting, the membrane was probed with a <sup>32</sup>P-labeled 1.2-kb *Bam*HI fragment of pAAAV.

**Generation of recombinant AAV particles.** For production of recombinant particles, we generated and examined the efficiency of three different helper plasmids, pMA<sub>3</sub>VRC, pCA<sub>3</sub>VRC, and pA<sub>3</sub>VRC, containing the AAV *rep* and *cap* genes under control of a mouse mammary tumor virus (MMTV), cytomegalovirus (CMV), or native p5 promoter, respectively. For generation of pMA<sub>3</sub>VRC, the *rep* and *cap* ORFs (nucleotides 243 to 4482) was produced by PCR with PFU polymerase (Stratagene) as specified by the manufacturer by using primers containing *Bst*Z107 and *Not*I sites. The PCR products were digested with *Bst*Z107 and *Not*I and ligated in a *Bst*Z107/*Not*I fragment of pMMTV2.1 (18) containing an MMTV promoter and SV40 poly(A). For generation of pCA<sub>3</sub>VRC, the *rep* and *cap* ORFs (nucleotides 243 to 4482) were produced by PCR with PFU polymerase and blunt-end ligated in the pCMV-script (Stratagene) vector, which contains a CMV promoter and SV40 poly(A). For generation of pA<sub>3</sub>VRC, the *rep* and *cap* genes of AAV including the p5 promoter and poly(A) signal (nucleotides 142 to 4516) was produced by PCR using PFU polymerase and blunt end ligated in pPCR-script. Orientation of inserts was verified by restriction digestion analysis, and final clones were confirmed by sequencing. For generation of the vector carrying the β-galactosidase gene flanked by AAV ITRs, the plasmid pAAAV was digested with *Bsm*BI (NEB). *Bsm*BI does not cut in the plasmid backbone but cuts at positions 838, 1111, 2590, 4419, and 4530 of the AAV genome. The resulting fragment that contained the plasmid backbone and 700 bp of AAV genome flanked by ITRs was used to ligate a *Bsm*BI-*Bsm*I linker. The resulting plasmid was digested with *Pml*I (cuts at nucleotide 146 of AAV genome) and *Bsm*I and used to ligate a *Bst*Z107-*Bsm*I fragment of pAAV<sub>2</sub>RnLacZ (18) that contains the β-galactosidase gene under the control of an RSV promoter and SV40 poly(A) tail. The resulting plasmid (pA<sub>3</sub>VRSVβ-Gal) was cotransfected with one of the helper plasmids described above and pAd12 in 293T cells plated in 150-cm dishes. Forty-eight hours posttransfection, cells were harvested and quantitated with a hemacytometer, and rAAAV was prepared by using standard CsCl gradient purification. The number of rAAAV genomes was estimated by using real-time quantitative PCR (QPCR) and expressed as nuclease-resistant particles per cell recovered after transfections. Titration of rAAAV was performed in exponentially growing CEF, DF-1, LMH, QNR, QT6, DT-90, 293T, COS, and primary embryonic chicken kidney cells and nondividing primary pituitary cells plated in 96-well plates and transduced with serial dilutions of recombinant virus for 48 h as previously described (20).

**RESULTS**

To obtain AAV genomic DNA for cloning, a stock of AAV was obtained from ATCC (VR-865) and coinfecting with FAV1 in day 10 embryonated chicken eggs. Virus was concentrated after subjecting infected allantoamniotic fluids to high-speed centrifugation. Viral DNA was released by SDS-proteinase K digestion and purified by gel electrophoresis after annealing the complementary single strands by heating the purified DNA to 95°C and slowly cooling to 65°C. Preliminary experiments indicated that 10<sup>5</sup> infectious particles of FAV1 resulted in productive infection without killing the embryo prematurely. Coinfection with at least 10<sup>5</sup> infectious particles of AAV was required to detect viral DNA (~4.7 kb) by ethidium bromide staining (data not shown). After recovery and end filling, the double-stranded AAV genome was blunt end ligated and cloned into pPCR-script. Several clones that contained an insert of approximately 4.7 kb were initially screened by restriction digestion (data not shown) and all gave bands similar in size to those previously reported (30). We subsequently sequenced three of these clones and all gave identical sequences. One of the clones was randomly selected and used in subsequent analysis (pAAAV).

To verify that pAAAV can support self-excision, viral DNA replication, and packaging in mammalian and avian cells, we prepared viral lysates from 293T and LMH cells transfected

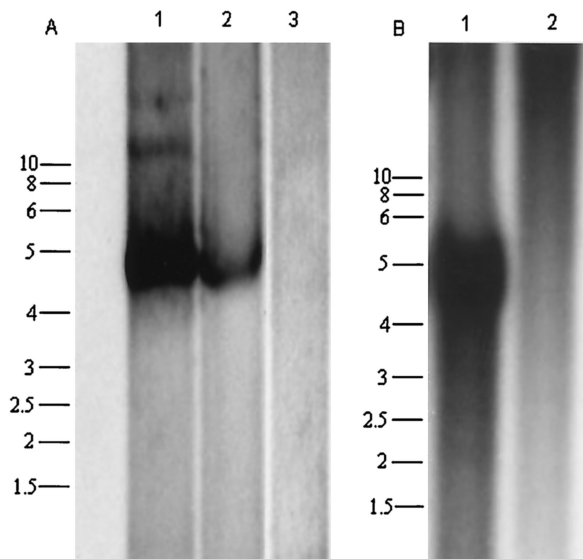


FIG. 1. Southern blot analysis of AAV nuclease-resistant particles in 293T and LMH cells. (A) 293T cells were transfected with pAAAV alone (lane 3), pAAAV plus pAd12 (lane 2), and pAAAV plus infection with wt Ad (lane 1). (B) LMH cells were transfected with pAAAV alone (lane 2) or pAAAV plus infection with FAV1 (lane 1). Viral DNA was isolated as described in Materials and Methods and fractionated on agarose gel before Southern blot analysis with a <sup>32</sup>P-labeled pAAAV DNA.

with pAAAV and infected with wild-type (wt) Ad5 or FAV1, respectively. In addition, the ability of an Ad5 plasmid to provide helper functions was examined in 293T cells. Southern blot analysis showed encapsidated (nuclease-resistant particles) AAV progeny in the presence of wt Ad5 or Ad helper plasmid in 293T cells and FAV1 in LMH cells but not in the absence of these (Fig. 1A and B). This result suggests that pAAAV can support rescue of AAV in mammalian and avian cells in the presence of mammalian or avian adenoviral genes.

The AAV ITR is composed of 142 nucleotides with the first 122 forming the characteristic T-shaped palindromic structure (Fig. 2), and it is 60 to 62% homologous with the ITRs of serotypes 2, 3, 4, and 6 and 48% homologous with AAV5. A tandem repeat of GAGY in the ITR, which serves as the binding element of Rep78 and Rep68 (RBE), is conserved between AAV and the other AAVs (Fig. 3, 4). The *trs* rec-

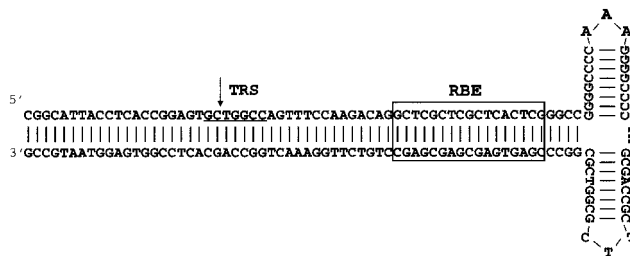


FIG. 2. AAV ITR. The sequence of the ITR is shown in the hairpin conformation. The putative Rep binding site is boxed, while the putative *trs* is underlined and the cleavage site is indicated by an arrow. RBE, Rep binding element.



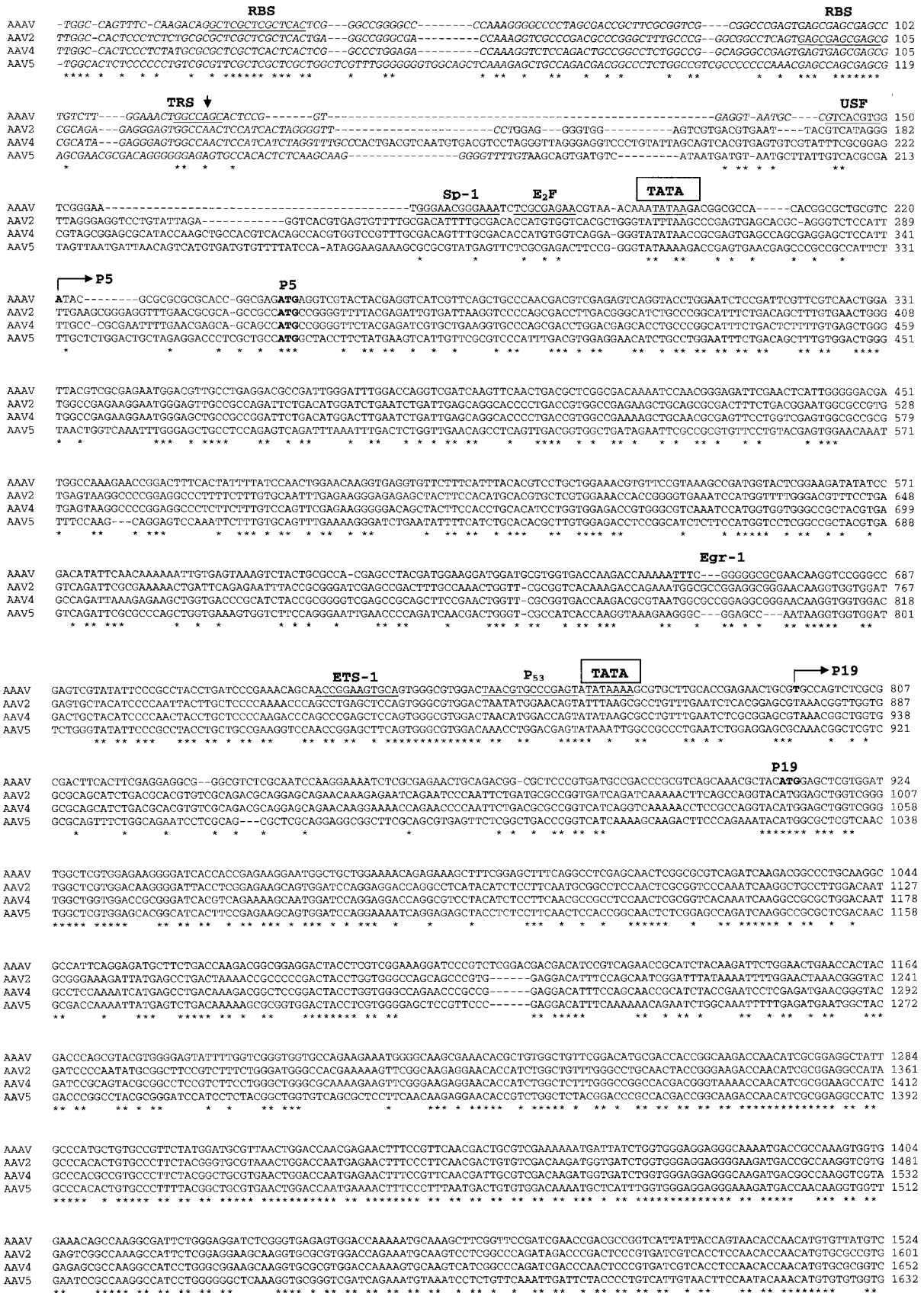


FIG. 3. Sequence of the AAV genome. The genomes of AAV1, AAV2, AAV4, and AAV5 were aligned by using Clustal W. The sequences of the ITRs are presented in italics. The putative *tr*s is indicated by a vertical arrow, and the putative Rep binding site is underlined. Proposed transcription factor binding sites and the polyadenylation signal are also underlined. Proposed transcription initiation sites of the p5, p19, and p40 promoters and splice donor and acceptor sites are indicated by horizontal arrows. Initiation and termination codons are presented in bold letters. USF, upstream factor.



AAAV CATCCGTCCTCCAAAGCGATGCGCTTTAGACTCTTTAACAATCCAGGTTAAAGAGGTCACGGTCCAAAGACTTCAACACCACCATCGGCAACAACCTCACCAGTACGGTCCAGGTCTTTGCGGA 3301  
AAV2 ATTCGGACCCAAAGAGACTCAACTTCAAGCTCTTTAACAATCAAGTCAAGGTCACGCAGAAATGACGGTACGACGACGATTCGCAATAACCTTACCAGCAGGTTACAGGTGTTTACTGA 3237  
AAV4 CATCGGACCCAAAGCGATGCGGGTCAAAATCTTCAACAATCCAGGTTAAAGAGGTCACGCAGTCAAGCGGACGACAAACGGTGGCTAATAACCTTACCAGCAGGTTACAGTCTTTGCGGA 3266  
AAV5 CTTCAACCCCGTCCCTCAGATCAAAATCTTCAACAATCAAGTCAAGGTCACGGTCCAGGACTCCACCACCATCGGCAACAACCTCACCCTCCAGCTGTTTACGGA 3218  
\* \* \* \* \*

AAAV CAAGGACTACCAACTGCGTACCTCTCGATCGGCTACCGAAGGACCTTCCCGCGCTTCCAGCGGATATCTACACGATCCCGCAGTACGGGTACTGACGCTAAACTACAAACAAG- 3420  
AAV2 CTCGGAGTACCAGCTCCGCTACGCTCTCGGCTCGGGCATCAAGGATGCTTCCCGCGCTTCCAGCAGACGCTTCTCATGGTGGCAGTATGGATACCTACCCCTGAACACCGGAGT- 3355  
AAV4 CTCGTCGTACCAACTGCGCTACGCTGATGGATCGGGTCAAGAGGGCAGCTGCTCTCTTTCCCAACGAGGCTTTTATGGTGCCTCAGTACGGCTACTGTGACTGTGTGACCGGCAAC 3386  
AAV5 CGACGACTACCAGCTGCGCTACGCTCGGCAACGGGACCGGGATGCTGCGGCGCTTCCCTCCGAGGCTTTACGTTCCGCGAGTACGGTTACGCGAGCTGAACCGGCAAC- 3336  
\* \* \* \* \*

AAAV -AGCGGTGGATCGTTCG- ----GCCTTCTACTGCTGGACTTTCCTCAGACATGCTGCGGACAGGAAATAACTTTGAGTTTACTTACACCTTCGAGGACGCTTCTTTCCATAG 3532  
AAV2 CAGGCAGTAGGACGCTC- ----TTCAATTTACTGCTGGAGTACTTTCCTTCTCAGATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGGAGCGTCTCTTTCCACAG 3468  
AAV4 TCTGTCAGCGCAGCAGCTACGAAATGCTTCTACTGCTGGAGTACTTTCCTTCCGAGATGCTGCGGACTGGCAACAACCTTTGAAATACGTACAGTTTGGAGAACTGCTTCCACTC 3506  
AAV5 -ACAGAAATCCACCAGAGGAGCAGCTCTTCTGCTAGAGTACTTTCACGCAAGTGTCTGAGAACGGGCAACAACCTTTGAGTTTACTTACACTTTTGGAGGAGTGCCTTCCACTC 3455  
\* \* \* \* \*

AAAV CATGTTTGGCCCAACAGCAGCTAGACCGGCTGATGAATCCCTCGTGGATCAGTACCTTGGGCTTTCAGCTCCGTCAGCCAA- ----GCAGGCTCATCTGGACGAGCTCTT 3641  
AAV2 CAGCTACGCTCAGCAGCAGCTCGGACCGTCTCATGAATCCCTCATCGACAGTACTGTATTTACTTGGACGAAACAACCT- ----CAAAGTGAACAACACGAGCTCAAGGCTTC 3584  
AAV4 GATGTACCGCAGCAGCAGGCTCGGACCGCTGATGAACCTCTCATCGACCACTGCTGCGGACTGCAATCGACCAACCGGAAACCCCTGAAATGCGGGGACTGCCA-CCACC 3624  
AAV5 CAGCTTCGCTCCGACTCAGAACCTGTCAAGCTGGCCAAACCGCTGGTGGACAGTACTTGTACCCTCTGTCGAGCAACAATA- ----ACACTGGCGGAGCT- ----C 3553  
\* \* \* \* \*

AAAV CATTTACTCGCGGGCAGTAAACCAACATGGCGGCTCAATATAGGAATGTTTCTTCCGCTTCTTCCGTTGATCAGCAATCTTTACGGCGCTAGCAACATCACTAAAAAATACGTC 3761  
AAV2 AGTTTTCTCAGGCGGAGCG-AGTGACATTCGGGACAGCTTGGAGACTGCTTGGAGCTTTCAGGCTGCTTACCGCAGCAGGATCAAAAGACTTCAACCAACAACAAGTGAATAC 3703  
AAV4 AACTTTTCAACAGCTCGCGCTTCCAACTTTTCCAACTTTAAAAAGACTGGCTGCGCGGCTTCAATCAAGCAGCAGGCTTCTCAAGACTGCAATCAAAACTACAA-GATCCCTGC 3743  
AAV5 AGTTCAACAAAGACTGCGCGGAGATACGCC-AACACCTCAAAAATGTTTCCCGGGCCATGGGCGCAACCCAGGCTGGAACCTGGGCTCCGGGTCAACCGGCGAGTGTACG 3672  
\* \* \* \* \*

AAAV TTTAGCGTTTGGGAAAAGGCAAGCAATGGGAATCGCAATCGGCAACCTAAATGACGCGCGCTTCCGCGCAGCAGCACCCTTTAGCGGAGAACCTGACCGTCAAG- ----CCAT 3874  
AAV2 TCGTGGACT--GGAG- ----CTACCAAGTACCACCTCAATGGCA- ----GAGACTCTCTGGTGAATCCGGCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTC- ----CTCAG 3805  
AAV4 CACCGGCT--AGACAGTCTCAATAACAGAGCAGCAGCAGCTC-TGGACGGAAGTGGAGTCCCTGACCCCGGACCTTCAATGGCCACGGCTGGACCTGGCGACAGCAAGTTTCAG 3860  
AAV5 GCCTTGCGC--ACGA- ----CAAATGAGTGGAGCTCGAGGCGCGAGTACCAGGTGCCCCCGACGGCAAGGCAATGACCAACAACCTTCCAGGCGACCAACCTATG- ----CCCTGGA 3782  
\* \* \* \* \*

AAAV GCAAAACAGCTGGCTTTAGCAGGACCGTCTACGATCAAAACGACCGCCACGACCGATCTGTAACAGATACTCATCCAACGAAAGCAAAATCAGACCCCAACCTCGTTCGGTATCGA 3994  
AAV2 -AGCGGGTCTCATCTTTGGGAAGCA-AGGCTCAGAGAAAACAATGTGAACATGAAAAGGTCATGATT--ACAGACGAAGAGGAAATCGGAACAACCAATCCGCTGGCTACGGA 3918  
AAV4 CAACAGCCAGCTCATCTTTGGGGGCT- ----AAACAGAAACGGCAACAGC-GCCACCTTCCCGGACTCTGATTTTCACTTGGAGGAGGCTGGCAGCCACCAACCCAGATACGGA 3977  
AAV5 GAACACTATGATCTTCAACAGCCAGCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAAC 3902  
\* \* \* \* \*

AAAV CGCGTGGGAGCAGTTCACCAACAACAGCTGATCTGACCCCGGCTTCCGCGCGGCTTCAACAATCAAGGGGCGCTTCCCGGATGGTGTGGCAAAACAGAGACATTTACCTTAC 4114  
AAV2 GCAGTATGGTCTGTATATCAAACTCCAGAGAGGCAACAGACAGTACCAGAGATGTAACACACAAGGCGTCTTCCAGCGATGGTCTGGCAGGACAGAGATGTTACTT-C 4037  
AAV4 CATGTGGGCAACCTACTTGGCGGTGACCCAGAGCAACAGCAACTCGGACCGTGGACAGACTGACAGCTTGGGAGCCGTCGCTGGAATGGTCTGGCAAAACAGAGCAATTTCACTAC-C 4096  
AAV5 CGTGGCGGCGAGTGGCCACCAACAACAGAGCTCCACCCTGCCCCGCGACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAAC 4021  
\* \* \* \* \*

AAAV AGGGACCATTTGG-CCAAAATTCGCGACACTGACAACTCACITCCATCCGCTCCCGCTTATTTGGGCGTTTGGCTGCAAGCATCCCCPCCCGAGATTTTCAATAAAAACACACCGGTC 4233  
AAV2 AGGGGCCATTCGGGCAAGATTCACACACAGCAGCGACATTTTCAACCTTCCCGCTTATGGGTTGAGTTCGAGACTTAAACACCCCTTCCACAGATTTCTCAAGAAACACCCCGGTAC 4157  
AAV4 AGGTCCATTTGGGCAAGATTCCTCATACCGATGACACTTTTCAACCTTCCCGCTTATGGGTTGAGTTCGAGACTTAAACACCCCTTCCACAGATTTTCAATAAAAACACACCGGTC 4216  
AAV5 AAGGACCATTCGGGCAAGATTCAGAGACGGGGCGCACTTTCAACCTTCCCGCTTATGGGCGGATTCGAGCTCAACACCCCGGCTTATGCTCATCAAGAAACAGCGCTGTGC 4141  
\* \* \* \* \*

AAAV CTGCCAACCTTCGGAACCTTCCAGCGGCAAGTGGCTTCTTCAACACAGTACTCGACCGGACA-GTGCACCGTGGAAATCTTTGGGAATCAAGAAGGAAACCTTCAAGCGC 4352  
AAV2 CTGGCAATCTTCGACCACTTCAAGTGGGCAAGTTTGGCTTCTTCAACACAGTACTCCAGCGGACCGGTCAGCTGGAGATCGAGTGGGAGTGCAGAAAGAAAACAGCAACCGC 4277  
AAV4 CTGGCAATCTTCGCAACGACTTCACTTCCGTTAACTTCTTCACTACTCAGTACAGCACTGGCCA-GGTTCGTTGAGATGACTGAGGAGTCCAGAAAGGAGCGTCCAAACGC 4335  
AAV5 CCGGAAATATC- ----ACCAGCTTCTCGGAGTCCCGTTCAGCAGCTTTCATCACCGCACTACAGCAACCGGCA-GGTCAACCGTGGAGTGGAGTGGGAGTCAAGAAGGAAACCTCAAGAGC 4257  
\* \* \* \* \*

AAAV TGGAAACCCGAAATCCAGTTTACCCTCCAATTT- ----GGCAACCGGCGCA-CATCCAGTTTGGCGTTCGACACGGGATCTTATCCGAACTCGTCCCACTCGGTACCGGTTACCTTA 4467  
AAV2 TGGAAATCCGAAATTCAGTACACTTCAACTAC- ----AACAAGTCTGTTAATGTTGAGCTTA-CGGTGGATACTAATGGCGTGTATTCAGAGCTCCGCCATTTGGCAACAGATACCTGA 4392  
AAV4 TGGAAACCCGAGGTCAGTTTACCCTCCAACCTACGACAGCAAACTCTGTTTGGGCTCC- ----GATCGGCTGGGAAATCACTGAGCTTAGGGCTATCGGTACCGGCTACCTGA 4450  
AAV5 TGGAAACCCGAGATCCAGTACACAAACTAC- ----AAGCA-CCCCAGTTTGTGACTTTGCCCGGACGACCGGGAAATACAGAAACCCAGACTATCGGAAACCCGATCCCTTA 4372  
\* \* \* \* \*

AAAV **VP-stop** CCAAACCTCTGTAA- ----ATTAA- ----ACCCTCAA-TAAACCG- ----TTTATGCGTAACTGTATTTCCGCTC- ----CTGTCGTTATTCAGTACATGA- ---- 4550  
AAV2 CTCGTAATCTGTAAATGCTGTGTAATCAATAAACCGTTTAAATCGTTTCACTGAACTTTGGTCTCGGTATTTCTT-TCATATAGTTTCCATGGCTAC- ----GT-AGATAAGTAGC- 4506  
AAV4 CCCACACCTGTAAATAACCGTTTAAATCAATAAACCGTTTAAATCGTTTCACTGAACTTTGGTCTCGGTATTTCTT-TCATATAGTTTCCATGGCTAC- ----GT-AGATAAGTAGC- 4569  
AAV5 CCCGACCTT**TAA**- ----CCATTCAT- ----GTCCATACCTCAA- ----TAAACCGT- ----TATCGTGTCAATAAATACTGCTC- ----TTGTGTCATTCAT- ----GAATAA-CAGC- 4470  
\* \* \* \* \*

AAAV -----TGCG- ----CATTACTC- ----ACCGAGTCTGG- ----CCAGTTTCCAAAGCAGGCTCGCTCGCTCACTCGGCGCG-G 4617  
AAV2 -----ATGGCGGTTAATCATTAACATAAGAAACCCCTAGTATGGAATGGCCACTCCCTCTGCGGCTCGCTCGCTCACTGAGCCGG- 4594  
AAV4 GCCTGGCGCTTGGCTTCCGCTTTACAACCTGCGGTTAATCAGTAACTTGGCAAACTTGGTCTCGGTGAGTGGCCACATAGCTATGAGTGGCCACATAGCTATGAGCTACCTCGGCTCACTCGGCTCGCT 4686  
AAV5 -----TTAGAATCTTCAAAAACCCCTTGTGAGAGTGGCACTTCCCCCTGTCGCTCGCTCGCTCGCTTGGGGGGCG 4559  
\* \* \* \* \*

AAAV GCCCAAAGGGGCGCTAGCAGCG- ----CTTGGCGTTCGGCGCGG- ----AGTGACGAGCAGGCTCTCTT-CAAAGT-CCCA- 4694  
AAV2 GCGACAAAGGTCGCGCAGCGCGCGG-CTTTGGCGGCGCGCTC- ----AGTGACGAGCAGGCGCGCAGAGGGAGTG-CCCAA 4675  
AAV4 GAGACAAAGGTCCTCAGACTGCGCG-CTTGGCGGCGAGGCGC- ----AGTGAGTGAGCAGGCGCAGAGGGAGTG-CCCAA 4767  
AAV5 ACGCGCAGAGGCGCTGCTGCGAGCTCTTTGAGCTGCCACCCCAACAGGACGAGCGAGCGCAACCGGCAAGGGGGAGAGTGCCA- 4652  
\* \* \* \* \*

FIG. 3—Continued.



ognition motif of serotypes 2, 3, 4, and 6 (CCGGT'TG) is highly homologous with that of the putative AAV *trs* (CCG GT'CG) and weakly homologous with the AAV5 *trs* site (ACG GT'GT). In addition, the spacing between the RBE and the putative *trs* is similar to that found in other serotypes, a characteristic that has been shown to be essential for Rep activity (12).

It has been proposed that a potential inverted repeat flanking the core *trs* sequence of AAV serotypes might be required for Rep *trs* nicking (11). Such an inverted repeat is not found around the AAV *trs* sequence. This observation may indicate that avian Rep nicking does not require any secondary structure around the core *trs* element. Methylation interference experiments have indicated the importance of the CTTTG motif found at the tip of one palindrome in AAV2 Rep binding (9, 57). Most of this motif is conserved in AAV ITR (CTTCG) and only one T residue is changed to C. Interestingly, the AAV4 ITR has a similar substitution in this motif (CTCTG). Thus, regardless of the overall nucleotide sequence homology, the secondary structure and the elements required for viral replication are conserved in the AAV ITR.

The entire AAV genome (Fig. 3) is 4,694 nucleotides in length and has a similar organization to that of other AAVs. It has two inverted terminal repeats and two distinct ORFs. The entire genome of AAV displays 56 to 65% identity at the nucleotide level with the other known AAVs. The p5 promoter region of AAV is much shorter and shows some divergence from homologous regions of other AAV serotypes. Core regulatory elements such as the TATAA box and Ebox/USF are conserved; however, YY1 and Rep binding sites are not present. This suggests that AAV gene expression might be regulated differently from that of other AAVs. The p19 promoter, the p40 promoter, and poly(A) can also be identified in the AAV genome by homology to those in primate AAV serotypes. Based on the general organization and sequence, these elements are highly conserved.

Clustal W protein sequence alignment indicates that the left ORF of AAV is 46 to 54% identical and equally divergent from that of the primate AAVs and the goose parvovirus (GPV) Rep ORF (Fig. 4A) and only 18 to 22% identical with the Rep ORF of other mammalian autonomous parvovirus (data not shown). In comparison, the Rep ORFs of isolates 1, 2, 3, 4, 6, 7, and 8 show greater than 90% similarity and are approximately 67 to 70% identical with that of the AAV5 Rep ORF. The central region of the AAV Rep ORF (amino acids [aa] 322 to 470), which is present in all Rep proteins, displays the greatest identity (82%) with the same region of the other AAVs and the GPV. This region of the Rep proteins is necessary for ATPase and helicase activity and contains an ATP-binding site (aa 334 to 349) and a divalent cation binding site at amino acid residue 421 (44, 61, 65). The amino terminus (aa 1 to 251) shows 42 to 45% similarity between AAV and the other AAVs. This region of the Rep78 and Rep68 proteins is required for DNA binding and *trs* endonuclease activities (22, 50). A tyrosine residue at 155 is homologous to the Tyr156 in AAV2 that functions as the active nucleophile in the *trs* endonuclease site (22, 62). The active site is assembled by the spatial convergence of a divalent metal ion that is tetrahedrally coordinated by Asp24, Glu83, His90, and His92. In addition, Glu6 is required for the correct orientation of the two active-

site imidazoles from His90 and His92 (31). All of these amino acid residues are strictly conserved among AAV serotypes, including AAV. Furthermore, a helix region important for Rep multimerization (aa 159 to 179) is also conserved in AAV. The carboxyl terminal portion (aa 490 to 662) of the unspliced AAV Rep proteins appears highly divergent, displaying less than 15% homology with the primate serotypes. However, a characteristic zinc finger motif was identified by using the BLIMPS algorithm. This feature is conserved in all AAV serotypes.

The right ORF of AAV, which encodes the three viral capsid proteins, is approximately 54 to 57% identical to the capsid ORF of the other AAVs and the GPV (Fig. 4B). It has been previously reported (6) that the AAV capsid proteins VP1, VP2, and VP3 have apparent molecular masses of 92, 69, and 61 kDa, respectively, as determined by SDS-polyacrylamide gel electrophoresis. The calculated molecular masses based on amino acid composition for VP1, VP2, and VP3 are 83, 67, and 60 kDa. We also subjected purified AAV virions to SDS-polyacrylamide gel electrophoresis and found that they have molecular masses of 91, 68, and 60 kDa (data not shown). As with the primate AAVs and the goose and duck autonomous parvovirus, the AAV cap gene contains two ATG initiator codons, one for VP1 and the other for VP3. The unusual ACG initiator codon for VP2 is also conserved in AAV.

Clustal W alignment of the VP ORFs indicated the presence of conserved and divergent regions. The N terminus of VP1 (aa 1 to 143), which is required for particle formation, is relatively conserved among AAV, AAV2, AAV4, AAV5, and GPV. However, the start sites for VP2 and VP3 are found in a divergent region. Based on the published three-dimensional structure of the canine parvovirus and comparisons of parvovirus capsid sequences (15, 67), most of the divergent regions among AAV, AAV2, AAV4, AAV5, and GPV are located on the exterior of the virus, thus suggesting different uptake mechanisms and altered tissue tropism.

In the present study, we constructed recombinant AAV particles containing the gene for nucleus-localized  $\beta$ -galactosidase. Virus was produced as previously described (19, 20) by constructing a vector plasmid containing the  $\beta$ -galactosidase gene under the control of a Rous sarcoma virus promoter flanked by AAV ITRs (pA3V $\beta$ -Gal, Fig. 5A) and a helper plasmid containing the AAV *rep* and *cap* genes. Virus was isolated from 293T cell lysates by CsCl banding, and the distribution of recombinant virus across the gradient was determined by QPCR analysis of gradient fractions. The majority of packaged genomes were found in fractions with a density of 1.42 g/cm<sup>3</sup>, which is similar to that of wt AAV. We also examined the yield of rAAV when using helper plasmids with the *rep* gene under the control of three different promoters, CMV, MMTV, or the native P5 promoter (Fig. 5A). The different helper plasmids (pCA3VRC, pMA3VRC, and pA3VRC) were cotransfected with pA3V $\beta$ -Gal and an adenovirus helper plasmid in 293T cells, and rAAV was purified from the three different clarified viral lysates by using CsCl gradients. The number of rAAV genomes was determined by QPCR. In three independent trials, the yield of rAAV was 5-fold and 15-fold greater when using the stronger CMV promoter than the yield with the MMTV and the native P5 promoter, respectively (Fig. 5A). This finding with rAAV is in

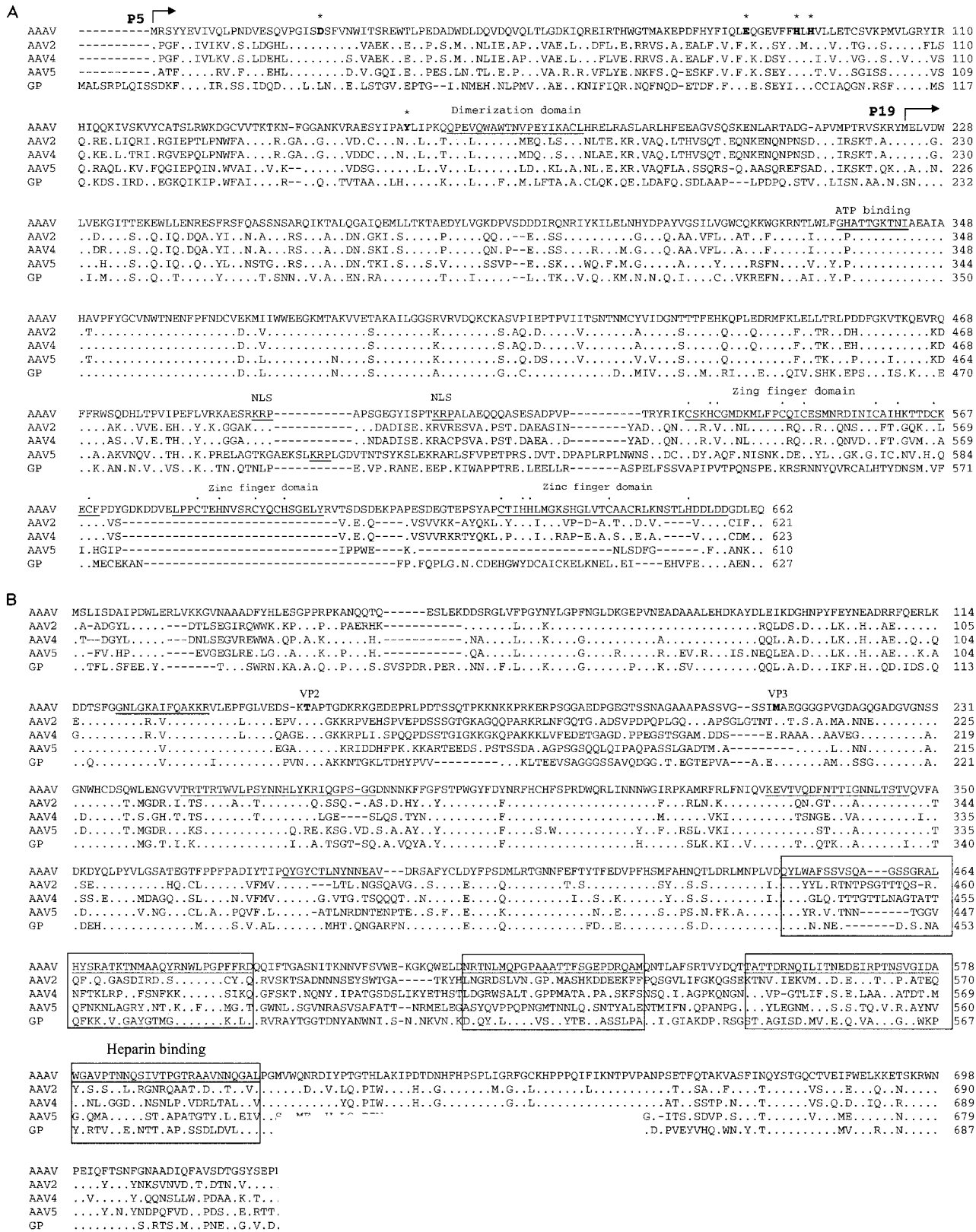


FIG. 4. Comparisons of *rep* and *cap* ORFs. The *rep* and *cap* ORFs of AAV, AAV2, AAV4, AAV5, and GPV (GP) were aligned by using Clustal W. Identical amino acids are indicated by a dot. Dashes indicate gaps in the sequence added by the alignment program. (A) Horizontal arrows indicate the initiator codon of the p5 and p19 Rep proteins. The Rep endonuclease site established by Tyr155 and the tetrahedrally coordinated Asp24, Glu83, His90, and His92 are presented in bold letters and are overlined by an asterisk. The region important for Rep multimerization, the ATP binding site, and the basic amino acids of the nuclear localization signal are underlined. The zinc finger motifs in the carboxy terminus are underlined and the coordinating cysteine and histidine residues are indicated by dots. (B) The theoretical initiator codons of VP2 and VP3 are indicated in bold letters. Regions that have been proposed to be on the surface of AAV2 are underlined and divergent regions are boxed. The heparin binding region in the capsid of AAV2 is also indicated.



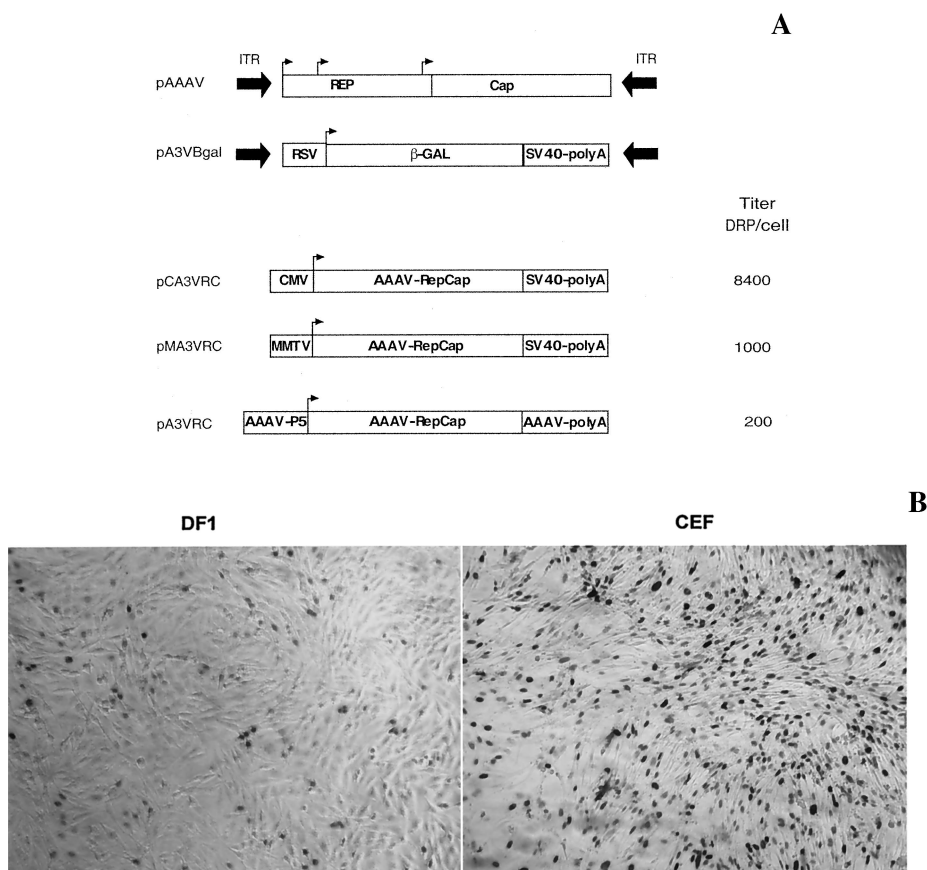


FIG. 5. Vector constructs for the generation of recombinant AAV and transduction of chicken fibroblasts. (A) Wild-type AAV, vector plasmid (pA3V $\beta$ -Gal), and production yields of rAAAV using helper plasmids providing the *rep* gene under the control of CMV, MMTV, or the native P5 promoter. The helper plasmids pCA3VRC, pMA3VRC, and pA3VRC were individually cotransfected with pA3V $\beta$ -Gal and an adenovirus helper plasmid in 293T cells, and rAAAV was produced as described in Materials and Methods. The number of rAAAV genomes produced in each group was determined by quantitative PCR and is expressed as DNase-resistant particles per cell (DRP/cell). (B) Relative transduction efficiency of primary chicken embryonic fibroblasts (CEF) and immortalized chicken embryonic fibroblasts (DF1) with equal particles of rAAAV expressing LacZ.

contrast to previous work with AAV2, which demonstrated that the use of a CMV promoter inhibited the production of rAAV2 (39).

In preliminary studies, we observed that the addition of detergents during virus purification affected infectivity. To better understand the effect of detergents, we prepared rAAAV in the presence of the following conditions: 0.5% deoxycholate, 0.5% 3-[(3-chloamidopropyl)-dimethylammonio]-1-propane-sulfonate (CHAPS), 0.5% octylglucoside, or no detergent, respectively. The virus from the four groups was purified by using CsCl gradients, and rAAAV genomes were quantitated by using quantitative PCR. No effect was observed on the yield of viral particles or density of rAAAV in the four preparations. After dialysis against phosphate-buffered saline, transduction efficiency was measured by titration on CEF cells. Addition of OCG or CHAPS had no significant effect on transduction efficiency. However, deoxycholate, which is a stronger ionic detergent, reduced transduction efficiency almost 10-fold (data not shown).

Tissue tropism of rAAAV was determined in CEF, DF1, LMH, DT-90, QNR, QT6, 293T, COS, primary chicken embryonic kidney cells, primary chicken pituitary cells, and pri-

mary human fibroblasts and compared with that of rAAV2, rAAV4, and rAAV5 (Table 1). Transduction efficiency of rAAAV was 10- to 300-fold higher in avian cells than in rAAV2, rAAV4, and rAAV5. In contrast, transduction of the mammalian cells in the panel by rAAAV was almost absent. This observation suggests that AAV is using a different uptake or transduction mechanism compared with the primate AAVs. Interestingly, rAAAV exhibited ~15-fold higher transduction efficiency in primary chicken embryonic fibroblasts than did the immortalized embryonic fibroblasts (Fig. 5B).

## DISCUSSION

Although the molecular and biological properties of AAV were largely unknown, serological and immunological data have indicated that AAV is distinct from the primate AAV (68, 69). That evidence prompted us to isolate, clone, and sequence an infectious clone of AAV and construct a recombinant vector.

Previous studies have indicated difficulties in directly cloning full-length infectious clones of AAVs. These difficulties have been attributed to the genetic instability of parvoviral inverted

TABLE 1. Titers for rAAAV, rAAV2, rAAV4, and rAAV5 expressing LacZ in avian and mammalian cell lines and primary cells

Cell type	Transducing units per 10 <sup>6</sup> genomes <sup>a</sup>			
	rAAAV	rAAV <sub>2</sub>	rAAV <sub>4</sub>	rAAV <sub>5</sub>
CEF	7,140 ± 380	25 ± 3.5	84 ± 6.3	58 ± 5.7
DF-1	530 ± 35	8 ± 0.9	45 ± 4.7	60 ± 6.1
LMH	2,380 ± 145	230 ± 25	34 ± 5.6	40 ± 4.9
DT-90	ND	ND	ND	ND
QNR	1,260 ± 90	176 ± 18	42 ± 5.2	185 ± 26
QT6	930 ± 62	112 ± 21	23 ± 3.8	33 ± 5
Chicken primary embryonic kidney cells	8,080 ± 560	422 ± 46	350 ± 40	235 ± 38
Chicken primary pituitary cells	4,640 ± 375	144 ± 17	70 ± 12	91 ± 8.4
293T	ND	4,500 ± 355	3,130 ± 270	684 ± 57
COS	5 ± 0.7	6,920 ± 420	3,550 ± 165	592 ± 53
A549	ND	2,190 ± 315	1,360 ± 140	26 ± 4.3
Human primary fibroblasts	ND	1,990 ± 170	1,130 ± 145	292 ± 31

<sup>a</sup> Transductions were performed as described in Materials and Methods, and efficiency is expressed as transducing units per 10<sup>6</sup> recombinant particles. Numbers represent mean ± standard deviation from four independent assays. ND, none detected.

terminal repeats. For that reason, investigators have used *recBC* bacterial strains (10, 70) or low-copy-number plasmids (56, 59) or constructed the full genome from cloned subfragments (20, 37, 48). Surprisingly, in the present study we did not encounter any difficulties in directly cloning the full AAV genome in a medium- to high-copy-number pUC18 derivative plasmid. This may indicate a higher genetic stability of the AAV ITRs than that of the primate isolates.

The nucleotide sequence of AAV is 56 to 65% identical with the other known AAVs and contains all the structural components and genetic elements that characterize the family of AAVs. These elements include the ITRs, promoters, ORFs, transcription start and stop sites, and intron splice junctions. Particularly, the ITRs of all serotypes (including AAV) are similar in length and symmetry, contain a conserved *rep*-binding site, and retain the ability to form the characteristic hairpin structure. Previous studies have demonstrated that the ability to form the terminal hairpin structure is important for AAV replication (38). This observation is further supported by the conservation of this structure between AAV and the primate AAVs.

The high degree of conservation of the *rep* ORF between the primate AAVs indicates the importance of this gene to the life cycle of the virus. The *rep* ORF of AAV is significantly more divergent than it is in other serotypes; however, the core region (aa 322 to 470) containing the ATPase and helicase activity is highly conserved (82% identity). This region is highly conserved in all vertebrate parvovirus, both autonomous and dependovirus, indicating the region's importance in the parvovirus life cycle. The N-terminal region of *rep* has been shown to be important for DNA binding; however, the exact amino acids involved are not known. The *rep*-binding site in the ITRs is highly conserved among AAVs including AAV. Therefore, it is anticipated that the motif in the N-terminal region of *rep* involved in DNA binding must also be conserved. The N-terminal region of AAV (aa 1 to 251) only shows 43% similarity with that of the other serotypes. Thus, the low degree of homology may help in identifying the conserved motif involved in DNA binding.

The carboxyl terminal of the unspliced Rep proteins encodes a zinc finger motif, and it is conserved among the AAV serotypes. The function of this region is largely unknown; however,

previous studies have indicated involvement in transactivation (23). In addition, this region has been shown to be important for interaction with the cellular kinases PrKX and PKA, causing inhibition of kinase activity (17, 23). The carboxyl terminus of AAV *rep* is highly divergent, displaying less than 15% homology. This fact may explain the increased titers of rAAAV obtained when using a helper plasmid that drives expression of the *rep* gene from a CMV promoter.

The predicted amino acid sequence of the capsid proteins indicates several regions with significant variation between serotypes (Fig. 4B). Several differences in the capsid proteins (aa 450 to 613) lie in regions that have been proposed to be on the exterior surface. These regions may play a role in serotype-specific properties such as antigenicity and/or binding to specific cellular receptors. However, not all of the changes are confined to the proposed exterior regions (aa 152 to 221), and they may also be important for other unique properties of AAV versus the primate AAVs.

Transduction efficiency in avian and mammalian cells was very distinct between rAAAV and rAAV2, rAAV4, and rAAV5. The only difference in our four recombinant constructs is the presence of serotype-specific ITRs flanking the expression cassette [RSV-β-Gal-SV40poly(A)] and the serotype-specific capsid. Although it is possible that each serotype-specific ITR could interact differently with host-specific intracellular factors, it is more likely that transduction efficiencies are affected by the presence of distinct cell surface receptors. These data, combined with the extensive divergence of the *cap* ORF, suggest that AAVs utilize a different uptake mechanism from those utilized by other serotypes.

The original hypothesis of a host-dependent evolution of parvovirus (3) is in contrast to the high sequence homology reported between the goose and Muscovy duck autonomous parvovirus and the primate AAVs (13, 40, 70). This observation raised the possibility of horizontal transfer of parvovirus between different species during evolution. The AAV might be the missing link between the avian autonomous parvovirus and the AAVs or may constitute a distinct branch in the evolution of dependovirus. The AAV genome (both *rep* and *cap* genes) is equally divergent between the avian autonomous parvovirus and AAVs. However, the structure and function of the AAV ITR are very similar to those of AAVs. Thus, the

proposed classification of parvoviruses based on the properties of the ITRs (1) gains further merit.

#### REFERENCES

- Astell, C. R. 1990. Terminal hairpins of parvovirus genomes and their role in DNA replication, p. 59–79. *In* P. Tijssen (ed.), *Handbook of parvoviruses*. CRC Press, Boca Raton, Fla.
- Atchison, R. W., B. C. Casto, and W. Hammon. 1965. Adenovirus-associated defective virus particles. *Science* **149**:754–756.
- Bando, H., J. Kusuda, T. Gojobori, T. Maruyama, and S. Kawase. 1987. Organization and nucleotide sequence of a densovirus genome imply a host-dependent evolution of the parvoviruses. *J. Virol.* **61**:553–560.
- Bantel-Schall, U., and H. zur Hausen. 1984. Characterization of the DNA of a defective human parvovirus isolated from a genital site. *Virology* **134**:52–63.
- Bauer, H. J., and G. Monreal. 1986. Herpesviruses provide helper functions for avian adeno-associated parvovirus. *J. Gen. Virol.* **67**:181–185.
- Bauer, H. J., R. Schneider, H. R. Gelderblom, R. Lurz, V. Friehmelt, and G. Monreal. 1991. Biological and physicochemical characterization of the major (1.40) and minor (1.45) component of infectious avian adeno-associated virus. *Arch. Virol.* **120**:123–133.
- Berns, K. I., and S. Adler. 1972. Separation of two types of adeno-associated virus particles containing complementary polynucleotide chains. *J. Virol.* **9**:394–396.
- Berns, K. I., M. Bergoin, M. Bloom, M. Lederman, N. Muzyczka, G. Siegl, J. Tal, and P. Tattersall. 1994. *Parvoviridae*. Vth report of International Committee on Taxonomy of Viruses, p. 166. *In* F. A. Murphy, C. M. Faquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers (ed.), *Virus taxonomy*. Springer-Verlag, Vienna, Austria.
- Bohenzky, R. A., R. B. LeFebvre, and K. I. Berns. 1988. Sequence and symmetry requirements within the internal palindromic sequences of the adeno-associated virus terminal repeat. *Virology* **166**:316–327.
- Boissy, R., and C. R. Astell. 1985. An *Escherichia coli* recBCSbcBrecF host permits the deletion-resistant propagation of plasmid clones containing the 5'-terminal palindrome of minute virus of mice. *Gene* **35**:179–185.
- Brister, J. R., and N. Muzyczka. 1999. Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification. *J. Virol.* **73**:9325–9336.
- Brister, J. R., and N. Muzyczka. 2000. Mechanism of Rep-mediated adeno-associated virus origin nicking. *J. Virol.* **74**:7762–7771.
- Brown, K. E., S. W. Green, and N. S. Young. 1995. Goose parvovirus—an autonomous member of the dependovirus genus? *Virology* **210**:283–291.
- Buller, R. M. L., and J. A. Rose. 1978. Characterization of adenovirus-associated virus-induced polypeptides in KB cells. *J. Virol.* **25**:331–338.
- Chapman, M. S., and M. G. Rossmann. 1993. Structure, sequence and function correlations among parvoviruses. *Virology* **194**:491–508.
- Chejanovsky, N., and B. J. Carter. 1989. Mutagenesis of an AUG codon in the adeno-associated virus rep gene: effects on viral DNA replication. *Virology* **173**:120–128.
- Chiorini, J. A., B. Zimmermann, L. Yang, R. H. Smith, A. Ahearn, F. Herberg, and R. M. Kotin. 1998. Inhibition of PrKX, a novel protein kinase, and the cAMP-dependent protein kinase, PKA, by the regulatory proteins of adeno-associated virus type 2. *Mol. Cell. Biol.* **18**:5921–5929.
- Chiorini, J. A., C. M. Wendtner, E. Urcelay, B. Safer, M. Hallek, and R. M. Kotin. 1995. High-efficiency transfer of the T cell co-stimulatory molecule B7-2 to lymphoid cells using high-titer recombinant adeno-associated virus vectors. *Hum. Gene Ther.* **6**:1531–1541.
- Chiorini, J. A., F. Kim, L. Yang, and R. M. Kotin. 1999. Cloning and characterization of adeno-associated virus type 5. *J. Virol.* **73**:1309–1319.
- Chiorini, J. A., L. Yang, Y. Liu, B. Safer, and R. M. Kotin. 1997. Cloning of adeno-associated virus type 4 (AAV4) and generation of recombinant AAV4 particles. *J. Virol.* **71**:6823–6833.
- Davidson, B. L., C. S. Stein, J. A. Heth, I. Martins, R. M. Kotin, T. A. Derksen, J. Zabner, A. Ghodsi, and J. A. Chiorini. 2000. Recombinant adeno-associated virus type 2, 4, and 5 vectors: transduction of variant cell types and regions in the mammalian central nervous system. *Proc. Natl. Acad. Sci. USA* **97**:3428–3432.
- Davis, M. D., J. Wu, and R. A. Owens. 2000. Mutational analysis of adeno-associated virus type 2 Rep68 protein endonuclease activity on partially single-stranded substrates. *J. Virol.* **74**:2936–2942.
- Di Pasquale, G., and S. N. Stacey. 1998. Adeno-associated virus Rep78 protein interacts with protein kinase A and its homolog PRKX and inhibits CREB-dependent transcriptional activation. *J. Virol.* **72**:7916–7925.
- El Mishad, A. M., K. J. McCormick, V. J. Yates, and J. J. Trentin. 1975. Detection and serological identification of adeno-associated virus in avian adenovirus stocks. *Infect. Immun.* **11**:287–299.
- 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.
- Green, M. R., and R. G. Roeder. 1980. Definition of a novel promoter for the major adenovirus-associated virus mRNA. *Cell* **22**:231–242.
- Green, M. R., and R. G. Roeder. 1980. Transcripts of the adeno-associated virus genome: mapping of the major RNAs. *J. Virol.* **36**:79–92.
- Henikoff, S., J. G. Henikoff, and S. Pietrokovski. 1999. Blocks+: a non-redundant database of protein alignment blocks derived from multiple compilations. *Bioinformatics* **15**:471–479.
- Hermonat, P. L., M. A. Labow, R. Wright, K. I. Berns, and N. Muzyczka. 1984. Genetics of adeno-associated virus: isolation and preliminary characterization of adeno-associated virus type 2 mutants. *J. Virol.* **51**:329–339.
- Hess, M., G. Paul, S. Kling, and G. Monreal. 1995. Molecular characterization of two strains of the avian adeno-associated virus (AAV). *Arch. Virol.* **140**:591–598.
- Hickman, A. B., D. R. Ronning, R. M. Kotin, and F. Dyda. 2002. Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep. *Mol. Cell* **10**:327–337.
- Hoggan, M. D., N. R. Blacklow, and W. P. Rowe. 1966. Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. *Proc. Natl. Acad. Sci. USA* **55**:1467–1474.
- Im, D. S., and N. Muzyczka. 1992. Partial purification of adeno-associated virus Rep78, Rep52, and Rep40 and their biochemical characterization. *J. Virol.* **66**:1119–1128.
- Kleinschmidt, J. A., M. Mohler, F. W. Weindler, and R. Heilbronn. 1995. Sequence elements of the adeno-associated virus rep gene required for suppression of herpes-simplex-virus-induced DNA amplification. *Virology* **206**:254–262.
- Labow, M. A., P. L. Hermonat, and K. I. Berns. 1986. Positive and negative autoregulation of the adeno-associated virus type 2 genome. *J. Virol.* **60**:251–258.
- Laughlin, C. A., H. Westphal, and B. J. Carter. 1979. Spliced adenovirus-associated virus RNA. *Proc. Natl. Acad. Sci. USA* **76**:5567–5571.
- Laughlin, C. A., J. D. Tratschin, H. Coon, and B. J. Carter. 1983. Cloning of infectious adeno-associated virus genomes in bacterial plasmids. *Gene* **23**:65–73.
- LeFebvre, R. B., S. Riva, and K. I. Berns. 1984. Conformation takes precedence over sequence in adeno-associated virus DNA replication. *Mol. Cell. Biol.* **4**:1416–1419.
- Li, J., R. J. Samulski, and X. Xiao. 1997. Role for highly regulated *rep* gene expression in adeno-associated virus vector production. *J. Virol.* **71**:5236–5243.
- 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.
- Lusby, E., K. H. Fife, and K. I. Berns. 1980. Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. *J. Virol.* **34**:402–409.
- Lusby, E., R. Bohenzky, and K. I. Berns. 1981. Inverted terminal repetition in adeno-associated virus DNA: independence of the orientation at either end of the genome. *J. Virol.* **37**:1083–1086.
- Mayor, H. D., K. Torikai, J. L. Melnick, and M. Mandel. 1969. Plus and minus single-stranded DNA separately encapsidated in adeno-associated satellite virions. *Science* **166**:1280–1282.
- McCarty, D. M., T. H. Ni, and N. Muzyczka. 1992. Analysis of mutations in adeno-associated virus Rep protein in vivo and in vitro. *J. Virol.* **66**:4050–4057.
- McLaughlin, S. K., P. Collis, P. L. Hermonat, and N. Muzyczka. 1988. Adeno-associated virus general transduction vectors: analysis of proviral structures. *J. Virol.* **62**:1963–1973.
- Mendelson, E., J. P. Trempe, and B. J. Carter. 1986. Identification of the *trans*-acting Rep proteins of adeno-associated virus by antibodies to a synthetic oligopeptide. *J. Virol.* **60**:823–832.
- Muralidhar, S., S. P. Becerra, and J. A. Rose. 1994. Site-directed mutagenesis of adeno-associated virus type 2 structural protein initiation codons: effects on regulation of synthesis and biological activity. *J. Virol.* **68**:170–176.
- Muramatsu, S. I., H. Mizukami, N. S. Young, and K. E. Brown. 1996. Nucleotide sequence and generation of an infectious clone of adeno-associated virus 3. *Virology* **221**:208–217.
- Nahreini, P., S. H. Larsen, and A. Srivastava. 1992. Cloning and integration of DNA fragments in human cells via the inverted terminal repeats of the adeno-associated virus 2 genome. *Gene* **119**:265–272.
- Owens, R. A., M. D. Weitzman, S. R. Kyostio, and B. J. Carter. 1993. Identification of a DNA-binding domain in the amino terminus of adeno-associated virus Rep proteins. *J. Virol.* **67**:997–1005.
- Parks, W. P., J. L. Melnick, R. Rongey, and H. D. Mayor. 1967. Physical assay and growth cycle studies of a defective adeno-satellite virus. *J. Virol.* **1**:171–180.
- Pronovost, A. D., V. J. Yates, and D. E. Fry. 1978. Effect of avian adeno-associated virus on pathogenicity of Tipton virus in chicks. *Avian Dis.* **22**:354–357.
- Pronovost, A. D., V. J. Yates, and D. E. Fry. 1979. Inhibition and enhance-



- ment of avian adenovirus plaque production by heavy and light avian adenovirus-associated viral particles. *Am. J. Vet. Res.* **40**:549–551.
54. **Quandt, K., K. Frech, H. Karas, E. Wingender, and T. Werner.** 1995. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23**:4878–4884.
55. **Rose, J. A., J. V. Maizel, Jr., J. K. Inman, and A. J. Shatkin.** 1971. Structural proteins of adenovirus-associated viruses. *J. Virol.* **8**:766–770.
56. **Rutledge, E. A., C. L. Halbert, and D. W. Russell.** 1998. Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2. *J. Virol.* **72**:309–319.
57. **Ryan, J. H., S. Zolotukhin, and N. Muzyczka.** 1996. Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats. *J. Virol.* **70**:1542–1553.
58. **Samulski, R. J., A. Srivastava, K. I. Berns, and N. Muzyczka.** 1983. Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV. *Cell* **33**:135–143.
59. **Samulski, R. J., K. I. Berns, M. Tan, and N. Muzyczka.** 1982. Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. *Proc. Natl. Acad. Sci. USA* **79**:2077–2081.
60. **Senapathy, P., J. D. Tratschin, and B. J. Carter.** 1984. Replication of adeno-associated virus DNA. Complementation of naturally occurring rep 2 mutants by a wild-type genome or an ori 2 mutant and correction of terminal palindrome deletions. *J. Mol. Biol.* **179**:1–20.
61. **Smith, R. H., and R. M. Kotin.** 1998. The Rep52 gene product of adeno-associated virus is a DNA helicase with 3'-to-5' polarity. *J. Virol.* **72**:4874–4881.
62. **Smith, R. H., and R. M. Kotin.** 2000. An adeno-associated virus (AAV) initiator protein, Rep78, catalyzes the cleavage and ligation of single-stranded AAV ori DNA. *J. Virol.* **74**:3122–3129.
63. **Smuda, J. W., and B. J. Carter.** 1991. Adeno-associated viruses having nonsense mutations in the capsid genes: growth in mammalian cells containing an inducible amber suppressor. *Virology* **184**:310–318.
64. **Tratschin, J. D., I. L. Miller, and B. J. Carter.** 1984. Genetic analysis of adeno-associated virus: properties of deletion mutants constructed in vitro and evidence for an adeno-associated virus replication function. *J. Virol.* **51**:611–619.
65. **Walker, S. L., R. S. Wonderling, and R. A. Owens.** 1997. Mutational analysis of the adeno-associated virus Rep68 protein: identification of critical residues necessary for site-specific endonuclease activity. *J. Virol.* **71**:2722–2730.
66. **Wonderling, R. S., S. R. M. Kyostio, and R. A. Owens.** 1995. A maltose-binding protein/adeno-associated virus rep68 fusion protein has DNA-RNA helicase and ATPase activities. *J. Virol.* **69**:3542–3548.
67. **Xie, Q., W. Bu, S. Bhatai, 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.
68. **Yates, V. J., A. M. el-Mishad, K. J. McCormick, and J. J. Trentin.** 1973. Isolation and characterization of an avian adenovirus-associated virus. *Infect. Immun.* **7**:973–980.
69. **Yates, V. J., G. J. Dawson, and A. D. Pronovost.** 1981. Serologic evidence of avian adeno-associated virus infection in an unselected human population and among poultry workers. *Am. J. Epidemiol.* **113**:542–545.
70. **Zadori, Z., R. Stefancsik, T. Rauch, and J. Kisary.** 1995. Analysis of the complete nucleotide sequences of goose and muscovy duck parvoviruses indicates common ancestral origin with adeno-associated virus 2. *Virology* **212**:562–573.