JOURNAL OF VIROLOGY, Nov. 1988, p. 3947–3957
0022-538X/88/113947-11$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 62, No. 11

Identification of Mouse Adenovirus Type 1 Early Region 1: DNA Sequence and a Conserved Transactivating Function

AMY OBERHAUSER BALL, MARK E. WILLIAMS, AND KATHERINE R. SPINDLER*
Department of Genetics, University of Georgia, Athens, Georgia 30602

Received 25 May 1988/Accepted 11 July 1988

The left end of the genome of mouse adenovirus type 1 (also known as strain FL) was characterized by determination of the DNA sequence, amino acid similarities with early region proteins of primate adenoviruses, and a functional assay. Several specific DNA sequence features were similar to those found in human adenoviruses, and open reading frames from this region could encode proteins similar to human adenovirus early region 1A and early region 1B proteins. DNAs from this region were tested in transient-expression assays in human and mouse cells and were found to transactivate the human adenovirus type 5 early region 3 promoter fused to the chloramphenicol acetyltransferase gene. The data indicate structural and functional homologies between mouse adenovirus type 1 early region 1 and early region 1 of primate adenoviruses.

A large amount of information is known about the molecular biology of human adenoviruses, including the processes of DNA replication, mRNA transcription and processing, and oncogenic transformation. Relatively little is known, however, about the molecular basis of adenovirus pathogenesis. There are 41 distinct types of human adenoviruses, which cause various acute respiratory, ocular, gastrointestinal, and urinary diseases (reviewed in reference 53). Because adenoviruses are species specific, an animal model system must be developed that will allow the use of a molecular genetic approach to study the mechanisms by which adenoviruses cause cell injury and disease. To this end, we have begun to characterize the molecular genetics of mouse adenovirus type 1, MAV-1 (also known as strain FL; see reference 27 for nomenclature), originally isolated by Hartley and Rowe (23). The pathogenesis of MAV-1 is not directly comparable with that of any one type of human adenovirus. MAV-1 has an adrenal and cardiac tropism (reviewed in reference 27). Infection of suckling mice by MAV-1 results in a wasting disease, followed by death; adult mice recover from infection but may shed the virus for up to 24 months after infection (27).

MAV-1 has morphological and biophysical characteristics similar to those of human adenoviruses, and it was originally characterized as an adenovirus on serological grounds (23, 65). Studies have shown that MAV-1 also has some of the molecular characteristics of other adenoviruses. For example, the MAV-1 DNA termini have inverted terminal repeats (56) and are associated with a protein (33). Also, the MAV-1 DNA-protein complex can be replicated in vitro in nuclear extracts from adenovirus type 2 (Ad2)-infected HeLa cells (56).

The genome organization of MAV-1 must be determined so that regions may be identified for future mutagenic studies. Both orientations have been proposed for the MAV-1 restriction map with respect to the physical and biological features of the virus (32, 33, 56). One standard criterion for the orientation of adenoviruses is to position the transforming region, early region 1 (E1), at the left end of the genome (2). There is no evidence that MAV-1 is oncogenic in vivo (12), thus the criterion of a transforming region does not orient MAV-1. However, proteins encoded in E1 of adenoviruses have other functional and sequence similarities. The human Ad2 and Ad5 289-amino-acid early region 1A (E1A) proteins possess a transactivating activity (reviewed in reference 3). Three polypeptide regions of the Ad2 and Ad5 289-amino-acid E1A proteins have sequence similarity to other primate adenovirus E1A proteins (29, 57; reviewed in reference 41). One of these regions has been correlated with the transactivating function of E1A (18, 19, 28, 34, 35, 42, 48). There is also significant sequence similarity among adenovirus E1B proteins (59).

Previously, the left end of MAV-1 DNA was defined by one group as that end corresponding to the *Eco*RI-C restriction fragment (32), on the basis of several criteria. These criteria include molecular hybridization studies that identified a region of MAV-1 DNA with sequence similarity to the DNA that encodes the major viral structural protein of Ad2, hexon protein (1, 32), and preferential packaging of one end of the virus in defective particles (32). Temple et al. (56) oriented the virus in the opposite way, with the right end of MAV-1 corresponding to the *Eco*RI-C fragment, on the basis of the AT-rich half of the viral DNA being positioned to the right (2).

We have chosen to orient the MAV-1 genome as Larsen et al. (32) did, by designating the left end as that corresponding to the *Eco*RI-C fragment. However, to conclusively orient the left end with respect to biological and physical characteristics, it was necessary to demonstrate a transactivating activity from MAV-1 DNA or to identify polypeptides (by prediction from DNA sequence) with similarity to the conserved regions of primate adenovirus E1A and E1B proteins, or both. In this report, we present the DNA sequence of the left end of MAV-1 DNA, including and extending beyond the *Eco*RI-C fragment. There is significant polypeptide sequence similarity between the open reading frames (orfs) predicted by the MAV-1 DNA sequence and the E1A and E1B proteins of primate adenoviruses. We also show that there is an activity encoded in this part of the genome which has functional similarity to the 289-amino-acid E1A protein. A plasmid containing the leftmost 6% of the MAV-1 genome was able to transactivate the human Ad5 E3 promoter in cotransfection transient-expression assays. From the sequence data and the functional assay, we have identified MAV-1 E1A and E1B at the left end of the genome.

* Corresponding author.

MATERIALS AND METHODS

Cells. Mouse L cells (from J. Holland, University of California, San Diego) and mouse 3T6 cells (from A. Winters, University of Alabama, Tuscaloosa) were maintained in Dulbecco modified Eagle medium (DMEM) plus 5% heat-inactivated calf serum (KC Biologicals). HeLa cells (from A. Berk, University of California, Los Angeles) were maintained in DMEM plus 7% newborn-calf serum (KC Biologicals or GIBCO Laboratories). All media were supplemented with 100 U of penicillin, 100 μg of streptomycin, and 20 U of nystatin (GIBCO) per ml.

Virus. Our pool of virus, derived from a single-plaque isolate of MAV-1 (strain FL), was a gift from S. H. Larsen (Indiana University Medical Center, Indianapolis). MAV-1 infections were carried out essentially as described previously (33). Briefly, the medium was removed, and virus was added to cell monolayers at a multiplicity of infection of 1 to 5 PFU per cell. After 1 h of adsorption at 37°C, DMEM containing 1% heat-inactivated calf serum was added to the cells. Unlike human adenoviruses, MAV-1 does not remain associated with the cells but is shed into the medium (65). Therefore, to prepare virus stocks, the medium from infected cell monolayers was collected at 3 days postinfection and frozen at −70°C (33, 56). Titers were determined by standard plaque assays used for human adenoviruses (66), except that mouse L or 3T6 cells were used.

Viral DNA isolation. Viral DNA was isolated by using a modification of a procedure described previously (33). Medium from infected cells was collected, and virions were precipitated by the addition of NaCl and polyethylene glycol to final concentrations of 0.5 M and 8%, respectively, and incubation overnight at 4°C. The pellet containing virions was obtained by centrifugation at 3,000 × *g* for 20 min and then suspended in 10 mM Tris hydrochloride (pH 8.0) for cesium gradient centrifugation (15). This suspension was loaded onto a step gradient of CsCl of ρ = 1.2 and 1.4 g/ml in 50 mM Tris hydrochloride (pH 8.0) and centrifuged for 90 min at 220,000 × *g* in an SW41 rotor. Virus material at the interface of the two densities of CsCl was either removed with a Pasteur pipette or collected from the bottom of the tube. The virus suspension was diluted with an equal volume of 10 mM Tris hydrochloride (pH 8.0), loaded on a linear gradient of CsCl in 50 mM Tris hydrochloride (pH 8.0) with ρ from 1.2 to 1.4 g/ml, and centrifuged for 18 h at 220,000 × *g* in an SW41 rotor. The lower band containing the virus was collected from the bottom of the tube and dialyzed for at least 2 h against 50 mM Tris hydrochloride (pH 8.0)–1 mM EDTA. Virus was digested with 250 μg of proteinase K per ml in the presence of 0.5% sodium dodecyl sulfate for 15 min at 55°C and then extracted twice with phenol and twice with chloroform (45). Viral DNA was then extensively dialyzed against 10 mM Tris hydrochloride (pH 8.0)–1 mM EDTA and stored at 4°C.

Plasmids and cloning of MAV-1 DNA. Plasmids were kindly supplied as follows: pAd2.17, containing the left end of Ad2 from 0 to 17 map units (m.u.), from D. Barker (D. D. Barker, Ph.D. thesis, University of California, Los Angeles, 1987); p3CAT, containing the Ad5 E3 promoter fused to the bacterial chloramphenicol acetyltransferase (CAT) gene, from R. Gaynor, University of California, Los Angeles (64); p*Eco*RIBAd5, containing the right end of Ad5 (from 84.0 to 100 m.u.), from G. Ketner, The Johns Hopkins University, Baltimore, Md. (4); pBE5X, containing Ad2 from 0 to 9.5 m.u. from D. Rosser (D. S. E. Rosser, Ph.D. thesis, University of California, Los Angeles, 1985); pSVA13, a pBR322 derivative containing the simian virus 40 (SV40) enhancer, origin of replication, and T-antigen-coding sequences, from B. Semler, University of California, Irvine (30); and pKJB55, a shortened form of pBR322 with a polylinker, and pUC18CM, a chloramphenicol-resistant derivative of pUC18, from K. Buckley (K. J. Buckley, Ph.D. thesis, University of California, San Diego, 1985).

Enzymes were obtained from Bethesda Research Laboratories, Inc., New England BioLabs, Inc., or Promega. The instructions of the manufacturers and standard techniques were followed for all cloning procedures (39). To clone all but the terminal fragments of the genome, viral DNA (passaged once in NIH 3T3 cells and three times in L cells after receipt of the pool from S. Larsen) was cleaved with restriction endonuclease *Pst*I, *Bam*HI, or *Sma*I and ligated to pUC18CM. Clones with MAV-1 DNA inserts were identified by Southern analysis (51). To clone the terminal fragments, viral DNA was digested with *Hpa*I and dC tailed (52), followed by ligation to *Pst*I-digested, dG-tailed pUC18CM or pKJB55. From the published restriction map (32) and our own restriction map (unpublished), we determined that clones representing the entire MAV-1 genome were isolated.

For transfection experiments, fragments of the MAV-1 genome were subcloned (Table 1). pMXD, containing the *Xba*I-D fragment (from 0 to 16.5 m.u.) as a tailed *Pst*I-*Xba*I insert, was constructed by a three-piece ligation of a fragment from pMHF2 (providing the vector pKJB55 and MAV-1 DNA from 0 to 5.0 m.u.), an *Sst*I-*Bam*HI fragment (from 5.0 to 5.8 m.u.) from pMBE, and a *Bam*HI-*Xba*I fragment (from 5.8 to 16.5 m.u.) from pMBC.

pMHC55, containing the tailed *Hpa*I-C fragment (from 87.5 to 100 m.u.) as a *Pst*I insert, was constructed by a three-piece ligation, because it has internal *Pst*I sites. The small *Pst*I-*Sst*II fragment (from 87.5 to 87.8 m.u.) of pMHC, the *Sst*II-*Eco*RI fragment (from 87.8 to 100 m.u.) of pMHC, and *Eco*RI-*Pst*I digested pKJB55 were ligated together.

pAB1 contained the *Sst*II-E fragment (from 87.8 to 100 m.u.) and was cloned from pMHC. pMHC was digested with *Sst*II at 87.8 m.u. to linearize it. The *Sst*II site was filled in with T4 DNA polymerase, and *Eco*RI linkers were added. After cleavage with *Eco*RI to remove excess linkers and to cut at the *Eco*RI site at the opposite end of the insert, the fragment was ligated into the *Eco*RI site of pSVA13. pAB2 and pAB3 contained the *Xba*I-D fragment (from 0 to 16.5 m.u.) from pMXD. pMXD was digested with *Xba*I and *Hin*dIII to cut out the insert. The *Eco*RI site at 7.4 m.u. was methylated, and the *Xba*I and *Hin*dIII sites were filled in with T4 DNA polymerase. *Eco*RI linkers were added, and the DNA was cleaved with *Eco*RI and ligated into *Eco*RI-digested pSVA13. In pAB2, the 16.5-m.u. end of the MAV-1 insert was located next to the ampicillin resistance gene, and the 0-m.u. end was next to the SV40 origin. In pAB3 the MAV-1 DNA insert was oriented in the opposite direction (Fig. 1). Restriction digests confirmed all constructions.

DNA sequencing. A set of nested deletions for sequencing were obtained by the BAL 31 exonuclease method (46) by using the BAL 31 enzyme and conditions as described by the manufacturer (Bethesda Research Laboratories, Inc.). pMBE was linearized with *Eco*RI and digested with BAL 31 exonuclease for various lengths of time. Digests of appropriate size ranges were pooled, digested with *Hin*dIII to cut out the insert, and cloned into *Sma*I-*Hin*dIII-digested pBS+ (Stratagene). A set of nested deletions from the opposite orientation was obtained by initially linearizing pMBE with *Hin*dIII, digesting with BAL 31, digesting with *Eco*RI, and

TABLE 1. MAV-1 DNA-containing plasmids

Plasmid	MAV-1 DNA insert[a]	Span (m.u.)[b]	Vector[c]	Source of insert
pAB1	*Sst*II-E[d]	87.8–100	pSVA13 (*Eco*RI)	Subcloned DNA[e]
pAB2	*Xba*I-D[d]	0–16.5	pSVA13 (*Eco*RI)	Subcloned DNA[e]
pAB3	*Xba*I-D[d]	0–16.5	pSVA13 (*Eco*RI)	Subcloned DNA[e]
pMBC	*Bam*HI-C	5.8–31.7	pUC18CM (*Bam*HI)	MAV-1 DNA
pMBE	*Bam*HI-E	1.6–5.8	pUC18CM (*Bam*HI)	MAV-1 DNA
pMHC	*Hpa*I-C	87.5–100	pUC18CM (*Pst*I)	MAV-1 DNA
pMHC55	*Hpa*I-C	87.5–100	pKJB55 (*Pst*I)	Subcloned DNA[e]
pMHF2	*Hpa*I-F	0–6.0	pKJB55 (*Pst*I)	MAV-1 DNA
pMPI	*Pst*I-I	8.2–13.7	pUC18CM (*Pst*I)	MAV-1 DNA
pMSCL	Left portion of *Sma*I-C	0.3–8.2	pKJB55 (*Sma*I-*Pst*I)	MAV-1 DNA
pMXD	*Xba*I-D	0–16.5	pKJB55 (*Pst*I-*Xba*I)	Subcloned DNA[e]

[a] The MAV-1 restriction map is found in reference 32.
[b] One map unit on MAV-1 DNA corresponds to about 315 bp. All plasmids that have an insert starting at 0 or 87.5 m.u. or ending at 6.0 or 100 m.u. also contain GC tails at those positions (see Materials and Methods).
[c] The restriction enzymes in parentheses indicate the vector sites into which the MAV-1 insert DNA was cloned.
[d] *Eco*RI linkers were added to blunt-ended insert fragments.
[e] The construction of these plasmids is described in Materials and Methods.

then cloning into *Sma*I-*Eco*RI-digested pBS− (Stratagene). Clones differing by 200 to 300 base pairs (bp) were chosen for sequencing. Fragments from 0 m.u. across the *Bam*HI site at 1.2 m.u., across the *Bam*HI site at 5.8 m.u. to the *Pst*I site at 8.2 m.u., and across the *Pst*I site were subcloned from pMHF2, pMSCL, and pMXD, respectively, into either bacteriophage M13 (40) or pBS+ or pBS−. A set of BAL 31 nested deletion clones in both directions from pMPI were obtained to provide information from the *Pst*I site at 8.2 m.u. to nucleotide (nt) 2950 at approximately 9.4 m.u.

Single-stranded DNAs were prepared from pBS+ or pBS− plasmids and M13 phage by using the protocol of the manufacturer. The DNAs were sequenced by the dideoxy-chain termination method (47), by using either the large fragment of DNA polymerase I or Sequenase enzyme (United States Biochemical Corp.).

Sequence analysis. Sequence analysis was done with Intelligenetics and International Biotechnologies, Inc.-Pustell sequencing programs. DNA similarities were analyzed by using the forward homology matrix comparison application. Amino acid sequences within each reading frame were compared with those of other adenovirus polypeptides or with those of proteins in the National Biomedical Research Foundation data bank by using the Lipman-Pearson FASTP algorithm (36) implemented by the International Biotechnologies, Inc.-Pustell Cyborg software. The default parameters from the PAM250 matrix (49) for amino acid matches were used in all cases. *Z* values (similarity scores) were calculated as (initial score − average score)/standard deviation. Random permutations of amino acid sequences were obtained by a random-shuffling method (14).

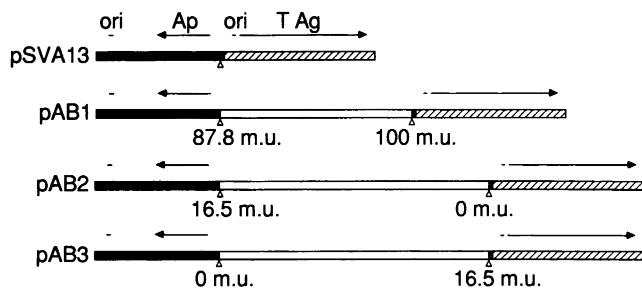


FIG. 1. Structure of plasmids used in DEAE transfections. The terminal regions of MAV-1 DNA were inserted into pSVA13 (30) as described in Materials and Methods. Symbols: ▬, sequences of pBR322; ▨, SV40 DNA sequences; ▭, MAV-1 DNA; △, *Eco*RI restriction sites. Abbreviations: Ap, ampicillin resistance gene; ori, origin (of pBR322 or SV40, as indicated by shading); T Ag, SV40 large T antigen. Arrows indicate the direction of transcription.

Transactivation assays. The reporter plasmid, p3CAT, and various test plasmids were introduced into cell lines by either calcium phosphate precipitation or DEAE-dextran, and chloramphenicol acetyltransferase (CAT) activity was assayed after 48 h. Human HeLa cells and mouse L cells were transfected by the calcium phosphate method as described elsewhere (20). The cells were plated 1 day prior to transfection at 0.5×10^6 to 2.5×10^6 cells per 60-mm-diameter plate. At 3 h before transfection, the medium was replaced with DMEM plus 7% heat-inactivated calf serum (L cells) or DMEM plus 10% newborn-calf serum (HeLa cells). The precipitates were formed by bubbling a mixture containing 3 μg each of p3CAT and test plasmid, 4 μg of salmon sperm DNA, 50 mM $CaCl_2$, and H_2O in a total volume of 500 μl into 500 μl of 2× *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES)-buffered saline (275 mM NaCl, 42 mM HEPES, 12 mM D-glucose, 10 mM KCl, 1.76 mM Na_2HPO_4; pH 7.06 ± 0.05). The DNA precipitates were added to each plate and incubated for 20 min, and medium was added. After 4 h, the medium was removed, and the cells were shocked with glycerol (1 ml of 10% glycerol in phosphate-buffered saline [2.7 mM KCl, 1.5 mM KH_2PO_4, 137 mM NaCl, 8 mM Na_2HPO_4]) for 1 min. The cells were washed with phosphate-buffered saline, and normal growth medium was added. The cells were harvested after 48 h, and cellular extracts were prepared to assay CAT activity.

Mouse 3T6 cells were transfected by the DEAE-dextran method as described by Lopata et al. (37) and modified by Selden et al. (50). Briefly, 3T6 cells were plated at 5×10^5 cells per 60-mm-diameter plate and transfected 2 days later. The medium was removed from the cells and replaced with 2 ml of DMEM plus 10% NuSerum (Collaborative Research, Inc.). Test plasmid and p3CAT (2 μg each) and DEAE-dextran to a final concentration of 200 μg/ml were added to each plate. After 2 h the medium was removed, and the cells were shocked for 1 min with 2 ml of 10% dimethyl sulfoxide in phosphate-buffered saline. The dimethyl sulfoxide–phosphate-buffered saline was removed, and DMEM containing 10% heat-inactivated calf serum was added. After 72 h the cells were harvested, and cell extracts were assayed for CAT

activity. The volume of extracts assayed for CAT activity was standardized to the amount of total protein in the cell extracts as determined by the method of Bradford (6), with reagents purchased from Bio-Rad Laboratories.

CAT assays were performed as previously described (20). Cell extract (30 μl) was added to a reaction mixture containing 70 μl of 0.25 mM Tris hydrochloride (pH 8.0), 25 μl of H_2O, 20 μl of 4 mM acetyl-CoA, and 0.1 to 0.2 μCi of [^{14}C]chloramphenicol (Amersham Corp.) and incubated for 30 to 60 min. The [^{14}C]chloramphenicol was extracted with ethyl acetate and subjected to ascending chromatography on thin-layer plates (J. T. Baker Chemical Co.), which were then exposed to XAR-5 X-ray film (Eastman Kodak Co.). The percentage of total chloramphenicol that was converted to acetylated forms was determined by excising the spots and counting them in toluene-based scintillation fluid.

RESULTS

Cloning of MAV-1 DNA. A complete library of MAV-1 DNA was obtained by cloning *Bam*HI, *Sma*I, and *Pst*I restriction endonuclease fragments of MAV-1 DNA into plasmid vectors. The clones obtained were compared with restriction maps (32, 33; K. Spindler, unpublished results) and included all but the terminal restriction fragments. Clones of the terminal fragments were obtained by GC tailing, as described by Stow for Ad2 (52). The terminal clones were sequenced. A comparison of the DNA sequence of these terminal clones to the published sequences of MAV-1 DNA termini (56) indicated that the complete termini (both left and right ends of MAV-1) had been cloned.

Sequence of MAV-1 DNA from 0 to 9.4 m.u. MAV-1 DNA is divided into 100 m.u. and is estimated to be about 31,500 bp, on the basis of restriction fragment data (data not shown). This estimate is in good agreement with the published molecular weight of MAV-1 DNA, 19.5×10^6 to 20.7×10^6 (33, 56). Because E1 of Ad2 extends from 0 to 11.2 m.u., sequencing of MAV-1 was begun at the left end of the viral DNA to determine whether the left end corresponds to E1 in this virus. Both strands were sequenced in their entirety by the dideoxy-chain termination method of sequencing (Fig. 2). The sequence of 2,950 bp, from 0 to 9.4 m.u., is shown in Fig. 3. The first 190 nt were previously sequenced by using viral DNA and direct chemical sequencing (56); the inverted terminal repeats common to all adenoviruses are 93 bp in MAV-1. We, like Larsen et al. (32), oriented the MAV-1 genome in the direction opposite that used by Temple et al. (56). Thus the sequence published by Temple et al. (56), which is designated the right end, corresponds to the left end of MAV-1 reported here. The sequence we obtained is identical to that of Temple et al. (56) until nt 177; thereafter the two sequences diverge (Fig. 3). Similarly, the sequence we obtained from the right end of the MAV-1 genome (left end in reference 56) is identical to that reported by Temple et al. (56) until nt 118 (numbering starts with nt 1 at 100 m.u.) (Spindler, unpublished). We obtained the sequence shown from both strands of several independent clones of the viral DNA, so the sequence difference does not appear to be a result of a cloning artifact. The viral DNA we used for cloning was obtained from a virus pool that was a gift from S. Larsen (see Materials and Methods); Temple et al. (56) also obtained a virus pool from the same laboratory (from D. Nathans). Temple et al. (56) isolated a single plaque and prepared their viral DNA from a pool made from this plaque. This difference in passage histories for the two virus pools sequenced may explain the differences in the sequences obtained.

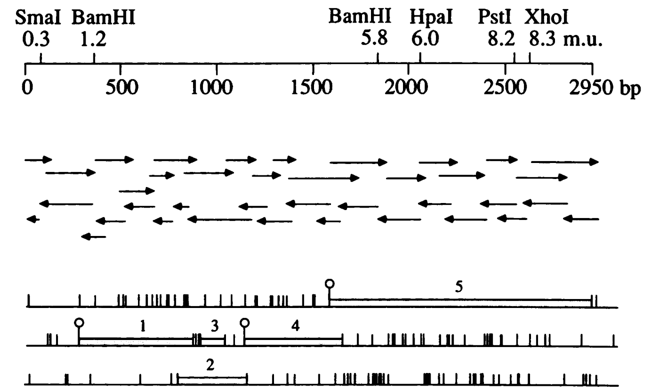


FIG. 2. Restriction map, sequencing strategy, and orfs of left end of MAV-1 DNA. The restriction map at the top indicates sites referred to in this work and their position on the genome in map units; nucleotide numbers are given below the line. Arrows indicate the clones that were used in determining the DNA sequence. Horizontal lines at the bottom indicate the three translational reading frames. Vertical lines indicate stop codons; vertical lines marked with an open circle indicate methionine codons. Numbered shaded regions indicate the orfs referred to in the text.

DNA sequence features. The DNA sequence of MAV-1 was compared to other adenovirus DNA sequences, and significant similarity was detected only in short, localized regions. As previously reported, an inverted terminal repeat (93 bp in MAV-1; 56) and the conserved sequence from nt 1 to 18 involved in DNA replication (8, 31, 55), common to adenoviruses, are found in MAV-1 DNA. We considered three major orfs, orfs 1, 4, and 5 (Fig. 2), as potential E1A- and E1B-coding regions (see below for discussion of protein similarities). By using a forward homology matrix comparison program, the DNA sequences of MAV-1 corresponding to these orfs and two minor orfs (orfs 2 and 3) were compared with those of Ad5 (60, 62) or Ad7 (10, 11). Within the regions corresponding to the conserved regions of primate E1A proteins (41), the DNA similarity between Ad7 and MAV-1 was 50 to 80%. The highest values corresponded to the short conserved region 2, as expected from a comparison of the protein sequence (see below for a discussion of protein similarities). The DNA similarity was not significant between the conserved regions of Ad5 and MAV-1, although the DNA sequence similarity between Ad5 and Ad7 for the conserved regions ranges from 60 to 90%. Outside the conserved regions and in the orfs corresponding to the E1B proteins, there was no significant similarity between the MAV-1 and Ad5 or Ad7 DNA sequences.

DNA sequences which have been identified as encapsidation signals for human adenoviruses include specific sequences identified either by in vitro mutational analysis (24) or by analysis of naturally occurring mutants and identification of conserved sequences within specific regions (7, 22). Hearing et al. (24) identified a repeated sequence in Ad5, designated A/A′, AGTAA(G)ATTTGG(G/C)C, at nt 239 (A) and 259 (A′), which functions as a *cis*-acting virus-packaging sequence. The G nucleotide at position 6 is not found in the A copy. An inverted match of MAV-1 DNA sequence to the consensus Ad5 DNA sequence (11 of 13 nt), GGaCtAATTT ACT, was found beginning at MAV-1 nt 194 (upper-case letters designate match). A second match (9 of 13 nt), ctTAAtTTTtGGC, was found beginning at MAV-1 nt 205. Another sequence identified by its presence in a region required for Ad16 packaging (22) and found in the first 500 bp

```
CATCATCAAT AATATACAGT TAGCAAAAAA TGGCGCCTTT GTTTGGCTTT GTTCCAACTG TTTTTGGCCC GAGTTGGGTT TCGTTTTCCC   90

GGGTTTCACG TGGTGCGTCA GGTGATGTTG AGTTTTTATA CATGTTTGTG TGTGTTTTGT GCTTGTGTAA AGGAGTTTTG GTTTTTTACT  180
ITR end                                                                                    ...CTT

TTGCGGAGCT TACGGACTAA TTTACTTAAT TTTTGGCAAA TTATATCTGT TTTTGGGTAT TTTGTTTGTT ATTTGGTGCG ATTTTTCGAC  270
GCGACACGAC A *

TTTTGACTCA ATGTCGCGGC TCCTACGTCT TTCTCTTTCT TCGCGTGTGT GGTTAGCAGC TCAGGAAGCT ACTCGCAATG TTTCTGAGGA  360
           met E1A (orf 1)

TCCTGTTGTA TGTCGTACGC CATGGGATGG TTCGCCTACT TGCACAGCTG TTCGTGTTGT ACGAGCGGAG GTTCTGGCGG ACGGTACGAT  450

GGACCTTGAC ATCGTGTTTC CTGAAGCTGC TGTACAGGCT GTTTTTAGTC GCACTCCATG GCAGGATTCT ACTACTGCCA CTTCTGCCGA  540

AGAACCGTCT GCTTCGACCG ATTCGATCTC CTCTGATCCC TTGCCTATAA GCTGTGTGGA GAGTTTTGAG GACATGGATC TGCGGTGTTA  630

TGAACAGTTG TCTCCCAGCC CCGAGTCTAT TGAAACTATT GAGGTGTTCC CGCCATGTTC TACGTGTGGA GGACATGAAG TTAATGGGTT  720

TTGCAGTCTG TGTTACTTGC GCGGTCTGAC TGGTAAGGTT TTTCTTTTAC GGATGATGTC TTTTTATGGT TCGTTGTTTG TTTCTAATAG  810
                                                            orf 2 -->

TCCTGATCTT TTTTTCAGAT CTGCTGCCTC AGGCGGATGA CGCAGGGGAA GCTGAAGTTC CTGATGAGTC TGCCAAGGAT TTATGTTTTA  900
                                     term orf 1                       orf 3 -->

TGGACCTGCT GACCTGGGCC ATGGAGGACA AGACGGAATG TTCCCGTCAC GGTCAGTGCT TGCTTCTTTG GCGGGAGAAT GTTGGCTTAG  990

AGTTGTGGAA AATTTTGTTT AGGTTGTATT GTTGTTTTTT GGCAGCTGTT ATGTTAGTAC TGTATAAAAA GGACAGGACT TGGTCAGGGG 1080
                  term orf 3

TGCATTATTT GGTTTCTGCG GCGGAGTGGT TAGGGTAGGC TGCAATATGT TACCTGTGTA TCCTTTCTTG GCTCCTGTCT TTGATACTTT 1170
                                    term orf 2  met E1B 21K (orf 4)

TGAGTCGACT CGTGCTTTAT TTTTGGAGGC GACTCTGTCT TTACAGGAGC CTTGGTGGTT TCGACTGCTT AGTTTTTTTA GATCCCCTCG 1260

TTACACGGGG CTAGCAGATG TGGTTAGAGC CATTGCTTGT GAGCGCGAGG CTGATTTTAC CGCTATTACT CCCTCTTCTT TAATCGAGGA 1350

TATTGCAGAG GGACGTTTTG TGATGCTTGT TGGGCTTATG GAGTATTTGC GTTTTGATAC TGCGGGCCAA GGCATTGTTT CATTTGCTTT 1440

TTTGTCTTAT CTTATTGATA GGGTTACGAA GCAAAGCCCT CTTGCGCATT TTACTGTTGT AGAGGTTGTG TCTTTGGTGG TATGGAGGAC 1530
                                                                                     met E1B 55K
                                                                                        (orf 5)

CGTCAAGCTC TCGAGGAGGC GGGAGAGGAG ATGGGTTTCC CAGTTGTCAA CATTAGCGGA GGAGGACGAG GACGAAGAGG GAACTACTCT 1620

AACGACGGAA GCGGAGCAAG AGAGCTCAGC TTGAGAGTTA ATGAGGAAGT TACTGAACCT CCTTACAATT CTATTACTTA TGCGGAGGTG 1710
                                 term E1B 21K (orf 4)

TTGCGTCTTT ATCGTCAGAA TGGGTTTTCT TTTTTTCTTT CTAACTTGCA TTTTGAGCAG ATAAAGTACT TTGAGCATAG AGAGGGTGAA 1800

GGTGATTTGG GCGCTATTAT TCGTGAGCAT GGAAAAAATAG TGTTGGATCC TAATGTTGTT TACCATTTAA GGCGTCCACT GAAGATTCAG 1890

AGTCTTTGTT ACATTGTAGG CAGGGGCGCT GTTATAAAGG TTCATGCAAG CGTGGGACAG CAAGCTGTTC ATGTTTATCG ACAGCTCGAT 1980

ATAAGTCCCC GCATTATGGG CATGTTAAAT GTGACATTTT TGGAATGTAA GTTTCACTGG GCGGATGATG AAGGGGTTAA CTTTGAAAGA 2070

CTTGCACGTC ACTTTACTTT GTTTTTTCTA ACATATGAAA GTGTTTTTTT TGCATGTGAT TTTGTTGGCT TTCCAGGCTT AACGTTAAGG 2160

AGCACCTGTC TTTTAAGGGT CGAGGGCTGT ACCTTCACTA GCTGTGCTGT GGGCATTCAC CACGCGGATG TTGCAGACCT GAAGGTCAAA 2250

GCCTGTTACT TTAATCATTG TTCTGTGTGT ATATTTGCTG ATGGACCTGC TGACGTATTG CGCAACTGTG CAACTAATTG TGATTGTTTT 2340

GTAATAATGG AGCAGGAAGG GTCTGTGGTA GGGAATTCTG TAGTGTATAA GCTACCTCCC TCTCGTGCAG ATGATCTAAT GATTTGTCAA 2430

CAAGGTTACA TGTATCCATT ATCAAGCATT CATATTGTTC GCAATCTGTC TTGTGAGTAC CCGAAGCTGA AAAACAATGT ATTTAGTCAT 2520

GTGGATATGC ATGTGGGTTT GCGACAAGGA CTGCAGCATT TTAACCAGTG TAATTTGAGC TTTGTGTATT GTATGTTAGA AACTGAAAGT 2610

GTTCCGAAAG TGAGCTTTTA CAGTTCGTAC ATACAAACGC TGACTGTGGC CCAGATTGTG CAATTCTCTC GAGAACATTC ACGAGAGTGT 2700

CAGTGTTTTT GCGGCGGAAG ACATAGATTG CTCTTTCCGT CCGTGGTTCA TATCACTCCT ACTGTGGTGC CGGATCGCAC GCGGTTTACG 2790

GTGGATGTGG AGGAATTGAG TGATGGTATG TACGAGACCA ATATGGGTGA GGGAAAATGT TTACTGTGTA ATTAAAAATT CTTTGTGTTT 2880
                                                                              term E1B 55K (orf 5)

GCAGATGAAT AGCCGTATGC GACGGTGGGC CGCTACTTCT CGGCTGTTGC ACGAAGATCC GCCTGCGACA                      2950
```

FIG. 3. DNA sequence of MAV-1 from 0 to 9.4 m.u. Several sequence features are indicated, including the end of the inverted terminal repeat (ITR end) and initiator (met) and termination (term) codons. The asterisk indicates the end of a previously published sequence (56) that is identical to the sequence determined here from nt 1 to 177 and differs as indicated from nt 178 to 191.

of Ad5, Ad7, Ad12, and tupaia adenovirus (7) was also present in MAV-1. In MAV-1, the sequence TATTTG GTGcGATTTtTTCGACTTTTGA, a match of 14 of the 16 conserved nucleotides (underlined), was found beginning at nt 250. The significance of either of these packaging sequences in MAV-1 remains unknown. Sequences that may cause DNA bending and that are proposed to be important for encapsidation (A tracts) have been identified at the left end of several adenoviruses (13). We have not analyzed the DNA bending of MAV-1 DNA, but we note that there are 11 A or T tracts that are at least 4 bases long 5' to the E1A-coding region (Fig. 3), an observation suggesting that DNA bending may occur. We have observed that a MAV-1 DNA fragment from nt 1 to 360 has anomalous mobility in acrylamide gel electrophoresis, a result consistent with DNA bending (data not shown).

E1A enhancer sequences have been located in Ad5 and identified in other transcriptional control regions (25, 26). Two elements were defined. Element I, (A/C)GGAAG TGA(A/C), which is found near nt 200 and near nt 300, regulates E1A transcription specifically and is also present in other viral transcriptional control regions. Element II, consensus (C/G)GCG(T/A)AA, which is composed of two sets of direct repeats inverted relative to each other and found between nt 250 and 280, enhances the expression of all early messages (26). A match of 7 of 10 nt to element I was found in the MAV-1 DNA sequence at nt 184, CGGAgcTtAC; inverted matches of 7 of 10 nt were found at nt 201 and 297; and a match of 8 of 10 nt was found at nt 333, AGGAAGct AC. A perfect copy and an inverted match of element II were found in MAV-1 DNA at nt 194 and 63, respectively. Ten inverted matches of 6 of 7 nt were found in the first 300 nt of MAV-1. One match began at nt 220, just downstream of the forward match at nt 184. There was a cluster of elements and packaging sequence homologies between nt 180 and 230 in the MAV-1 DNA sequence that was similar to the clustered packaging and enhancer elements in Ad5, a result suggesting that this region may be important for transcriptional regulation or packaging.

Protein similarities. The amino acid sequence was deduced from the DNA sequence, and orfs were identified (Fig. 2). Amino acid sequences within each reading frame were compared with those of other adenovirus polypeptides or with those of other proteins in the National Biomedical Research Foundation data bank, as described in Materials and Methods. Orf 1 had sequence similarity to primate adenovirus E1As; orfs 4 and 5 had sequence similarities to the human adenovirus E1B 21,000-molecular-weight (21K) and 55K proteins, respectively. Orfs 2 and 3 were also examined because they, like the primate adenovirus E1A second exon, lie upstream of the orf for the 21K E1B protein.

The first exons of the larger E1A proteins of primate adenoviruses are characterized by three conserved regions interspersed with nonconserved sequences (29, 41, 57). Orf 1 was similar to primate adenovirus E1As only in these three short regions of exon 1. Alignments similar to those reported by van Ormondt et al. (61) were obtained. The alignment of MAV-1 E1A with Ad4, Ad5, and Ad7 E1A sequences is shown in Fig. 4A. The similarity within any one of the conserved regions is about 70% among Ad4, Ad5, and Ad7 E1A proteins; when MAV-1 E1A sequences were included, the similarity for all four polypeptides was about 40%. Significant similarities with MAV-1 were not found outside these conserved regions. The similarity was best in conserved region 2, between amino acids 112 and 128 of MAV-1 E1A, in which more than 50% of the amino acids in all four polypeptides were identical or similar. Overall, exon 1 of MAV-1 E1A is slightly more similar to E1A of Ad4 or Ad7 than to that of Ad5 (28 versus 24%). In exon 2 of primate adenovirus E1As, the sequence similarity is only about 15%. We were unable to positively identify a polypeptide region in the sequenced MAV-1 DNA corresponding to the portion of the primate adenovirus E1As encoded in exon 2 in orfs 1, 2, or 3 (Fig. 2 and 4A). However, there was a stretch of amino acids in orf 2 (amino acids 18 to 60) which showed some similarity to the portion of E1A encoded in exon 2 of human adenovirus E1A.

Orfs 4 and 5 were compared to the E1B proteins from primate adenoviruses. These orfs were similar to the E1B 21K and 55K proteins, respectively (Fig. 2, 4B, and 4C). Both E1B alignments were similar to those published previously for Ad5, Ad7, and Ad12 (59). An alignment for the proposed MAV-1 E1B 21K protein with the Ad5 and Ad7 E1B 21K proteins is shown in Fig. 4B. The similarity between Ad5 and Ad7 is about 60% overall; the similarity between MAV-1 and either Ad5 or Ad7 was about 37%. Other E1B 21K proteins from Ad4 (57), Ad12 (54), and the proposed tupaia adenovirus E1B (16) showed less similarity with the MAV-1 protein and were omitted from Fig. 4B for clarity.

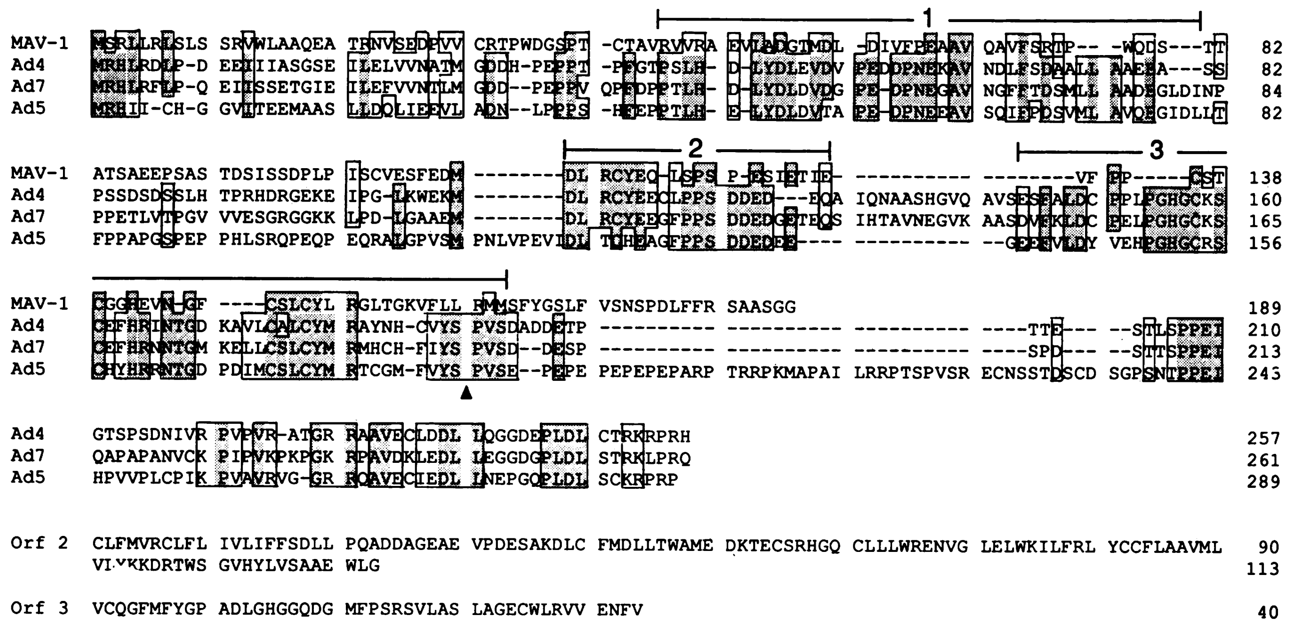
The alignment of the putative MAV-1 E1B 55K protein with proteins from Ad5 and Ad7 is shown in Fig. 4C. Ad5 and Ad7 55K proteins are about 70% similar, whereas the similarity between MAV-1 and either Ad5 or Ad7 was about 42%. The Ad12 (54) and proposed tupaia E1B 55K proteins (16) were compared, but they were not as similar and were not included in the figure. For both the 21K and 55K proteins, MAV-1 matched Ad5 slightly better than it did Ad7 (44 versus 41%).

To test the significance of the matches of the E1B proteins, 100 random permutations of the MAV-1 E1B 21K and 55K polypeptide sequences were generated and matched against Ad5 E1B 21K or 55K proteins. For both the 21K and 55K E1B proteins, the similarity scores for the authentic MAV-1 peptides were greater than 6, a score indicating significant matches, whereas the scores for the scrambled sequences were less than 2, a score indicating matches that were not significant (36, 49). This analysis was not done for the E1A peptide because the overall similarity was very low.

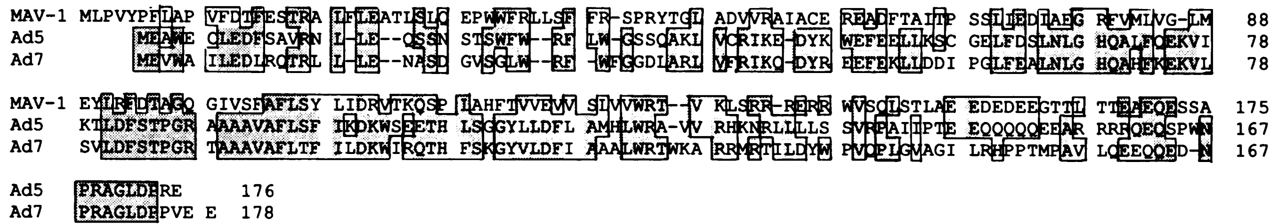
The longest orf in the opposite direction was 139 amino acids between nt 1866 and 1449. No significant similarity was found between this orf and any adenovirus peptides or orfs or with other proteins in the National Biomedical Research Foundation protein data bank.

Transactivation by the left end of the MAV-1 genome. The human adenovirus E1A protein can transactivate a human Ad5 E3 promoter when plasmids encoding the E1A protein are cotransfected into cells with reporter plasmids containing the E3 promoter (63, 64). We tested both the left and right ends of MAV-1 DNA to determine whether either end encodes a transactivating protein analogous to that of human adenovirus E1A. For the experiments described here, the reporter plasmid p3CAT contained the E3 promoter upstream of the CAT gene (64). The level of CAT activity produced after transfection of the p3CAT plasmid has been shown to correlate with the level of steady-state mRNA transcribed from the E3 promoter in this plasmid (17). p3CAT and test plasmids were cotransfected into HeLa or L cells by calcium phosphate precipitation (Fig. 5A). The cells were harvested and assayed for CAT activity 2 days after transfection. Extracts from HeLa cells cotransfected with

A



B



C

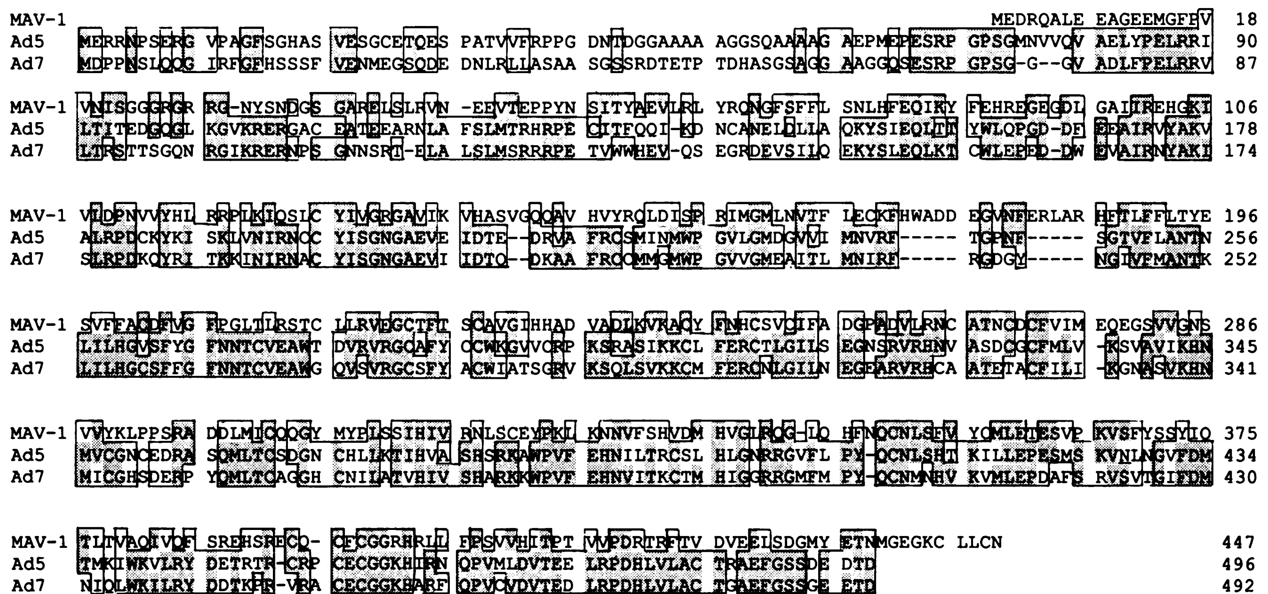


FIG. 4. Alignment of predicted MAV-1 polypeptides with adenovirus E1 proteins. Amino acid sequences predicted from the MAV-1 DNA sequence were aligned for maximum similarity to human Ad4, Ad5, or Ad7 E1 proteins, as described in Materials and Methods. Single-letter abbreviations for the amino acids are used. Boxes are drawn around amino acids that are similar or identical in three of four viruses (A) or in two of three viruses (B and C); identical amino acids are shaded. (A) Comparison of MAV-1 orf 1 to the larger E1A proteins of Ad4, Ad7, and Ad5 (289-amino-acid protein). Regions indicated by horizontal lines 1, 2, and 3 indicate the three conserved regions identified for primate adenovirus E1A proteins (41). The solid triangle indicates the position of the splice junction between exons 1 and 2 of Ad4, Ad7, and Ad5. There is no alignment of MAV-1 orf 1 to the E1As after this point, nor is there any good alignment of orfs 2 or 3 (shown at the bottom of the figure) to amino acids of the E1A exon 2 (see text). (B) Comparison of MAV-1 orf 4 to the E1B 21K proteins of Ad5 and Ad7. (C) Comparison of MAV-1 orf 5 to the E1B 55K proteins of Ad5 and Ad7.

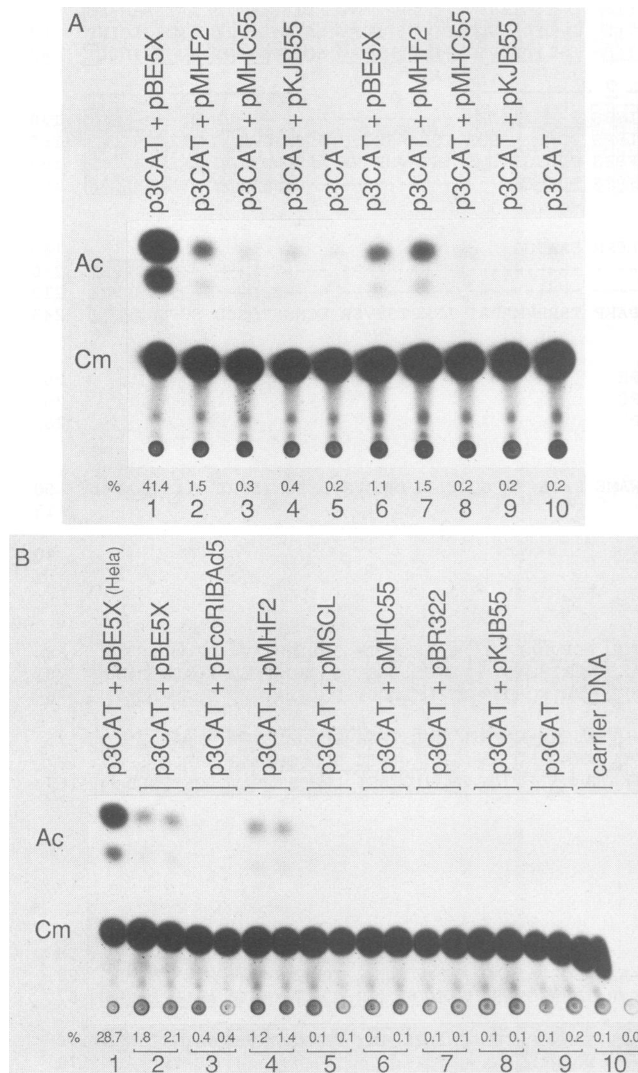


FIG. 5. Transactivation by the left end of MAV-1 in HeLa and L cell transient-expression assays. Cells were cotransfected with various plasmids by the calcium phosphate precipitation method, and at 48 h posttransfection, the cell extracts were assayed for CAT activity. Chloramphenicol and its major acetylated forms are indicated by Cm and Ac, respectively. The percent conversion of chloramphenicol to acetylated forms is given above the lane numbers. (A) HeLa cells (lanes 1 to 5) or L cells (lanes 6 to 10) were cotransfected with p3CAT and the indicated plasmid (lanes 1 to 4 and 6 to 9) or p3CAT alone (lanes 5 and 10). (B) Lane 1, HeLa cells were cotransfected with p3CAT and pBE5X; lanes 2 to 9, L cells were cotransfected in duplicate with p3CAT and the indicated plasmid; lane 10, only carrier (salmon sperm) DNA was used in the duplicate transfections.

p3CAT and pBE5X, which contained human Ad2 E1A (from 0 to 9.5 m.u.), served as a positive control. Conversion of the substrate to acetylated products in extracts from pBE5X-transfected cells demonstrated the previously documented transactivation by Ad2 E1A (Fig. 5A, lanes 1 and 6). The left end of MAV-1 (pMHF2; Fig. 5A lanes 2 and 7) also showed transactivation. In HeLa cells, the amount of CAT activity induced by pMHF2 was lower than that induced by pBE5X, as evidenced by the 1.5% conversion of substrate to acetylated product for pMHF2 versus 41.4% for pBE5X. However, in L cells, induction of CAT activity by pMHF2 was

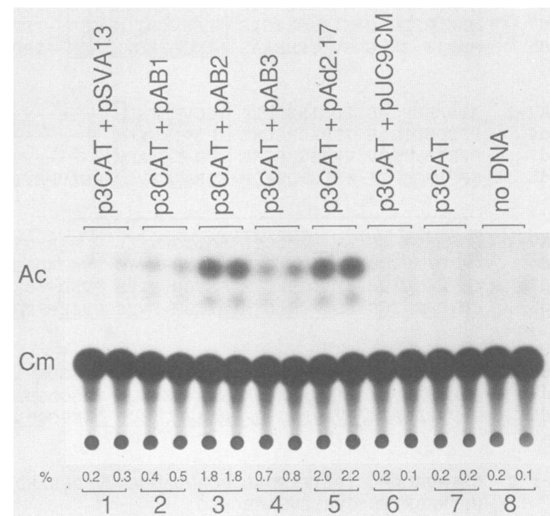


FIG. 6. Transactivation by the left end of MAV-1 introduced into mouse 3T6 cells by DEAE-dextran. p3CAT and the indicated test plasmids were introduced into mouse 3T6 cells in duplicate by the DEAE-dextran method. At 72 h after transfection, the cell extracts were assayed for CAT activity. Cm, Ac, and percent conversion are as defined in the legend to Fig. 5. In lanes 7 and 8, the transfections contained only p3CAT or no DNA, respectively.

approximately the same as that by pBE5X (1.5 versus 1.1%). Transfection by the right end of MAV-1 (pMHC55; Fig. 5A, lanes 3 and 8) did not result in appreciable CAT activity, and no CAT activity was seen in negative controls: pKJB55 (vector without insert; Fig. 5A, lanes 4 and 9) and p3CAT alone (Fig. 5A, lanes 5 and 10). In a second experiment (Fig. 5B), duplicate cotransfections were done in L cells. Again, transactivation was observed in extracts from cells transfected by pBE5X and pMHF2 and not from extracts of cells transfected by the right end of Ad5 (Fig. 5B, lanes 3), the right end of MAV-1 (Fig. 5B, lanes 6), or by negative controls (Fig. 5B, lanes 7 to 10). Although pMSCL (nt 90 to 2550) lacked only the first 90 bp, compared with MHF2 (nt 1 to 2055), it was unable to transactivate in three separate experiments (Fig. 5B; data not shown). In the experiments shown here and several others (data not shown), the transactivation activity induced by MAV-1 left-end plasmids was low but ranged from 3- to 12-fold over the background in all cases.

To try to increase the signal of transactivating activity over background levels, transfections were performed by using the DEAE-dextran method and mouse 3T6 cells (Fig. 6). Neither the vector pSVA13 (Fig. 6, lanes 1) alone nor pAB1 (Fig. 6, lanes 2), which contained the rightmost 12.2 m.u. of the genome, were able to transactivate the E3 promoter under these conditions. pAd2.17 (Fig. 6, lanes 5), which contained the left end of Ad2, served as a positive control. Plasmids pAB2 and pAB3 (Fig. 6, lanes 3 and 4), which contained the leftmost 16.5 m.u. of MAV-1 in opposite orientations, were able to transactivate p3CAT and induced CAT activity levels by two- to fourfold above the background. Transfection of pAB2 resulted in about a twofold-higher level of conversion to acetylated CAT than was seen with pAB3 (1.8 versus 0.8%). In a second experiment (data not shown), pAB2 transfection again resulted in more conversion to acetylated CAT than when pAB3 was used (5.9 versus 1.9%). In both pAB1 and pAB2 (right end and left end, respectively), the bidirectional SV40 enhancer was

adjacent to the proposed MAV-1 transcription start sites (A. Ball and K. Spindler, unpublished results), whereas in pAB3 it was not. Thus, pAB1 and pAB2 should be compared to each other in order to establish that the left end of MAV-1 encodes a transactivating activity and the right end does not.

DISCUSSION

To study the molecular basis of the pathogenesis of adenoviruses, an outline of the genome organization of MAV-1 is needed so that mutants can be made in the appropriate genes and then studied in vitro and in vivo. Larsen et al. (32) oriented the MAV-1 DNA with the *Eco*RI-C fragment at the left end of the genome by using two criteria. We have used the same orientation as that used by Larsen et al., which is based on the preferential packaging of incomplete genomes and the probable coding region for the major structural protein, hexon protein. The sequence analysis of the left end of MAV-1 was undertaken to search for protein or DNA sequence similarities, or both, with the E1 regions of other adenoviruses. The sequence obtained included orfs with significant similarity to exon 1 of the coding region for E1A and the entire coding regions for the two major E1B proteins. We identified a transactivating activity encoded in this region similar to the transactivating activity found in Ad5 E1A. Remarkably, the MAV-1 E1 transactivated the heterologous human adenovirus E3 promoter in both mouse and human cells. The sequence and functional homologies of this MAV-1 region to sequences and transactivating activities of other adenoviruses identify it as E1. This finding is also consistent with the orientation of other adenoviruses, which have E1 at the left end of their genomes. Our results conclusively identify an orientation for the genome of MAV-1.

The DNA similarity of human adenoviruses within a subgroup is approximately 80%, whereas between subgroups, the DNA sequence similarity in human adenoviruses drops to about 20% (21). Given these numbers, it is not surprising that there is little overall DNA sequence similarity between MAV-1 and the other adenoviruses. Local areas of DNA sequence similarity common to other adenoviruses can be found in MAV-1, including DNA sequence similarities within the regions corresponding to the conserved regions of the E1A proteins. MAV-1 also has the terminal sequences involved in DNA replication and sequences with similarity to enhancer and packaging sequences characterized in Ad5 or Ad16 which are conserved in other adenoviruses. The functional significance of these enhancer or packaging sequences has not been demonstrated in MAV-1.

The human adenoviruses have been grouped as highly oncogenic (subgroup A; for example, Ad12), weakly oncogenic (subgroup B; Ad7), or nononcogenic (subgroups C [Ad2 and Ad5], D, and E [Ad4]) (21, 58). The ability of adenoviruses to cause tumors is encoded in E1. In the regions corresponding to E1, MAV-1 DNA and amino acid sequences are more similar to those of the nononcogenic or weakly oncogenic adenoviruses Ad4, 5, and 7 than to that of the highly oncogenic adenovirus Ad12. This conclusion is consistent with observations that MAV-1 is nononcogenic (12).

Orf 1 was identified as exon 1 of E1A on the basis of sequence similarity in the three conserved regions, although the similarity between the MAV-1 E1A polypeptide and the primate adenovirus E1A proteins was much less than that among primate adenovirus E1A proteins. We were unable to conclusively identify a MAV-1 peptide sequence corresponding to the polypeptide encoded by exon 2 of primate E1As, although there was a short sequence in orf 2 which showed some similarity to the polypeptide encoded in exon 2 of Ad5 E1A. However, this orf and orfs 1 and 3 did not include a nuclear localization sequence just upstream of a termination codon (38). In addition, the conserved protein sequence Y<u>S</u>PVS, which includes the serine that is crucial for transactivation in Ad5 (19), was not present in the amino acid sequence in orf 1, 2, or 3. The absence of these conserved sequences suggests a much greater divergence of sequence or mRNA structure than noted previously among the adenoviruses.

Potential transcription signals for the E1A message can be identified by analysis of DNA sequence similarities. For example, a cluster of elements with similarity to enhancer elements is located between nt 180 and 230, 50 to 100 bp upstream of the ATG start for the E1A protein. Sequences similar to the TATA homology are found upstream of the E1A-coding region. However, without knowledge of the message structure, the significance of these sequences is unknown. We found no consensus polyadenylation signal (AAUAAA, AUUAAA, or ACUAAA) 5′ of the coding region for the E1B 21K protein, where one would be expected for the E1A message (44). In addition, preliminary cDNA and transcription analyses suggest an mRNA organization different from that found in other adenoviruses (A. Ball, M. Stringer, and K. Spindler, unpublished results).

In addition to the determination of the DNA sequence, our studies also revealed information about transactivation by MAV-1-encoded polypeptides. We have taken advantage of the ability of human adenovirus E1A proteins to transactivate other promoters and have used transactivation to test for an E1A-like activity encoded by MAV-1. The data indicate that the left end of MAV-1 encodes a transactivating activity whereas the right end does not. In addition, this transactivating activity functioned in a heterologous system. MAV-1 transactivated the E3 promoter from human Ad5, and the transactivation was observed in human cells, that is, in a nonpermissive host for the virus. Two MAV-1 left-end plasmids differed in their ability to transactivate; both plasmids contained the entire identified E1A-coding region. A plasmid that extended from nt 1 to 2055 (pMHF2; the leftmost 6% of the genome) transactivated the Ad5 E3 promoter. pMSCL (nt 90 to 2550), however, did not transactivate the Ad5 E3 promoter, a result indicating that sequences between nt 1 to 90, which comprise most of the MAV-1 inverted terminal repeat, may be important for the expression of transactivating activity in mouse cells. These findings are consistent with preliminary evidence that there are mRNAs from this region that initiate at several sites between nt 120 and 160 (Ball and Spindler, unpublished). Transactivation by the left end of MAV-1 was also demonstrated by using a different plasmid vector which included the SV40 origin of replication, early and late bidirectional transcription enhancer (reviewed in reference 9), and T-antigen gene coding sequences. In three experiments, pAB2 resulted in more transactivation than did pAB3. The difference between the two constructs, which both had the MAV-1 sequence from 0 to 16.5 m.u., was the orientation of the MAV-1 insert relative to the vector. In pAB2, the bidirectional enhancer was adjacent to the proposed start(s) of transcription at the left end (Ball and Spindler, unpublished). Perhaps an enhancer effect is responsible for the higher level of transcription from the MAV-1 insert in pAB2 relative to pAB3.

The organization of the MAV-1 E1B region is consistent

with previously identified E1B proteins (5). The orf for the MAV-1 E1B 21K protein preceded that for the E1B 55K protein by 130 nt (compared with 300 nt for Ad5) in an overlapping reading frame. Transcription regulatory signals have been extensively analyzed for the human adenovirus E1B messages, including TATA boxes, Sp1-binding sites, and other protein-binding sites (43, 67). Potential regulatory sites upstream of MAV-1 E1B-coding sequences included a TATA box located at nt 1053 and a GC-rich sequence similar to the consensus binding site for the transcription factor Sp1 at nt 969.

MAV-1 has retained many typical adenovirus features, including morphological and biophysical characteristics and serological cross-reactivity. In this work, we have identified protein structural and functional homologies of MAV-1 to human adenoviruses. The MAV-1 early region 1 proteins have strong similarity to conserved protein sequences in primate adenovirus E1As and E1Bs. We also identified a transactivating function in MAV-1 that is analogous to transactivating functions found in other adenoviruses. There are also some differences between the genome organizations in MAV-1 and human adenoviruses (Ball and Spindler, unpublished). It will be important to investigate the impact of these similarities and differences with respect to the disease determinants of adenoviruses. It may be possible to create chimeric viruses by exploiting these similarities and differences for the study of adenovirus pathogenesis. Further work on genome organization, mRNA transcription, and the proteins encoded by MAV-1 is in progress.

ACKNOWLEDGMENTS

We thank the scientists who provided the cells, plasmids, and virus stocks used in this work. We also thank Arnie Berk, Kenn Buckley, Richard Gaynor, Rich Meagher, Lois Miller, David O'Reilly, and Bert Semler for helpful discussions. We thank Clayton Beard, Becky Mullally, and Sambra Redick for providing part of the MAV-1 DNA sequence analysis. We are grateful to Marty Ball for assistance with random sequence generation and statistical analysis and to Keena Lowe for editorial comments.

This work was supported by grants from the American Cancer Society (MV 302) and the University of Georgia Research Foundation and by a Public Health Service grant (AI-2376202) from the National Institutes of Health. K.R.S. is the recipient of an American Cancer Society Junior Faculty Award.

ADDENDUM

DNA sequence analysis of MAV-1 cDNA clones corresponding to mRNAs that could encode the E1B proteins indicates that there is a short intron corresponding to nt 2816 to 2884 (Ball and Spindler, unpublished data). This alters the predicted amino acid sequence of the E1B 55K protein. Instead of the final 16 amino acids shown in Fig. 4C, the predicted sequence should be S-D-D-E.

LITERATURE CITED

1. **Antoine, G., P. Aleström, R. Schilling, U. Petterson, and E.-L. Winnacker.** 1982. Defective expression of mouse adenovirus Fl in human cells. EMBO J. **1**:453–459.
2. **Bachenheimer, S., et al.** 1977. Adenovirus strand nomenclature: a proposal. J. Virol. **22**:830–831.
3. **Berk, A. J.** 1986. Adenovirus promoters and E1A transactivation. Annu. Rev. Genet. **20**:45–79.
4. **Berkner, K. L., and P. A. Sharp.** 1983. Generation of adenovirus by transfection of plasmids. Nucleic Acids Res. **11**:6003–6009.
5. **Bos, J. L., L. J. Polder, R. Bernards, P. I. Schrier, P. J. van den Elsen, A. J. van der Eb, and H. van Ormondt.** 1981. The 2.2 kb E1b mRNA of human Ad12 and Ad5 codes for two tumor antigens starting at different AUG triplets. Cell **27**:121–131.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
7. **Brinckmann, U., G. Darai, and R. M. Flügel.** 1983. Tupaia (tree shrew) adenovirus DNA: sequence of the left-hand fragment corresponding to the transforming early region of human adenoviruses. EMBO J. **12**:2185–2188.
8. **Challberg, M. D., and D. R. Rawlins.** 1984. Template requirements for the initiation of adenovirus DNA replication. Proc. Natl. Acad. Sci. USA **81**:100–104.
9. **Das, G. C., S. K. Niyogi, and N. P. Salzman.** 1985. SV40 promoters and their regulation. Prog. Nucleic Acid Res. Mol. Biol. **32**:217–236.
10. **Dijkema, R., B. M. M. Dekker, and H. van Ormondt.** 1980. The nucleotide sequence of the transforming BglII-H fragment of adenovirus type 7 DNA. Gene **9**:141–156.
11. **Dijkema, R., B. M. M. Dekker, and H. van Ormondt.** 1982. Gene organization of the transforming region of adenovirus type 7 DNA. Gene **18**:143–156.
12. **Dussaix, E., N. Faucon-Biguet, D. Samolyk, C. Boni, and P. Tournier.** 1982. Restricted replication of mouse adenovirus strain FL in mouse Balb/c cells transformed by simian adenovirus 7. Virology **116**:641–645.
13. **Eckdahl, T. T., and J. N. Anderson.** 1988. Bent DNA is a conserved structure in an adenovirus control region. Nucleic Acids Res. **16**:2346.
14. **Efron, B.** 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia.
15. **Everitt, E., S. A. Meador, and A. S. Levine.** 1979. Synthesis and processing of the precursor to the major core protein of adenovirus type 2. J. Virol. **21**:199–214.
16. **Flügel, R. M., H. Bannert, S. Suhai, and G. Darai.** 1985. The nucleotide sequence of the early region of the Tupaia adenovirus DNA corresponding to the oncogenic region E1b of human adenovirus 7. Gene **34**:73–80.
17. **Garcia, J., F. Wu, and R. Gaynor.** 1987. Upstream regulatory regions required to stabilize binding to the TATA sequence in an adenovirus early promoter. Nucleic Acids Res. **15**:8367–8385.
18. **Glenn, G. M., and R. P. Ricciardi.** 1985. Adenovirus 5 early region 1A host range mutants *hr*3, *hr*4, and *hr*5 contain point mutations which generate single amino acid substitutions. J. Virol. **56**:66–74.
19. **Glenn, G. M., and R. P. Ricciardi.** 1987. An adenovirus type 5 E1a protein with a single amino acid substitution blocks wild-type E1a transactivation. Mol. Cell. Biol. **7**:1004–1011.
20. **Gorman, C.** 1985. High efficiency gene transfer into mammalian cells, p. 143–190. *In* D. M. Glover (ed.), DNA cloning: a practical approach. IRL Press, Oxford.
21. **Green, M., J. Mackey, W. S. M. Wold, and P. Rigden.** 1979. Thirty-one human adenovirus serotypes (Ad1-Ad31) form five groups (A-E) based upon DNA genome homologies. Virology **93**:481–492.
22. **Hammarskjöld, M.-L., and G. Winberg.** 1980. Encapsidation of adenovirus 16 DNA is directed by a small DNA sequence at the left end of the genome. Cell **20**:787–795.
23. **Hartley, J. W., and W. P. Rowe.** 1960. A new mouse virus apparently related to the adenovirus group. Virology **11**:645–647.
24. **Hearing, P., R. J. Samulski, W. L. Wishart, and T. Shenk.** 1987. Identification of a repeated sequence element required for efficient encapsidation of the adenovirus type 5 chromosome. J. Virol. **61**:2555–2558.
25. **Hearing, P., and T. Shenk.** 1983. The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. Cell **33**:695–703.
26. **Hearing, P., and T. Shenk.** 1986. The adenovirus type 5 E1a enhancer contains two functionally distinct domains: one is specific for E1a and the other modulates all early units in cis. Cell **45**:229–236.
27. **Ishibashi, M., and H. Yasue.** 1984. Adenoviruses of animals, p. 497–562. *In* H. S. Ginsberg (ed.), The adenoviruses. Plenum

Publishing Corp., New York.
28. **Jelsma, T. N., J. A. Howe, C. M. Evelegh, N. F. Cunniff, M. H. Skiadopoulos, M. R. Floroff, J. E. Denman, and S. T. Bayley.** 1988. Use of deletion and point mutants spanning the coding region of the adenovirus 5 E1A gene to define a domain that is essential for transcriptional activation. Virology **163**:494–502.
29. **Kimelman, D., J. S. Miller, D. Porter, and B. E. Roberts.** 1985. Ela regions of the human adenoviruses and of the highly oncogenic simian adenovirus 7 are closely related. J. Virol. **53**: 399–409.
30. **Kuhn, R. J., E. Wimmer, and B. L. Semler.** 1987. Expression of the poliovirus genome from infectious cDNA is dependent upon arrangements of eukaryotic and prokaryotic sequences in recombinant plasmids. Virology **157**:560–564.
31. **Lally, C., T. Dörper, W. Gröger, G. Antoine, and E.-L. Winnacker.** 1984. A size analysis of the adenovirus replicon. EMBO J. **3**:333–337.
32. **Larsen, S. H., R. F. Margolskee, and D. Nathans.** 1979. Alignment of the restriction map of mouse adenovirus FL with that of human adenovirus 2. Virology **97**:406–413.
33. **Larsen, S. H., and D. Nathans.** 1977. Mouse adenovirus: growth of plaque-purified FL virus in cell lines and characterization of viral DNA. Virology **82**:182–195.
34. **Lillie, J. W., M. Green, and M. R. Green.** 1986. An adenovirus Ela protein region required for transformation and transcriptional repression. Cell **46**:1043–1051.
35. **Lillie, J. W., P. M. Loewenstein, M. R. Green, and M. Green.** 1987. Functional domains of adenovirus type 5 Ela proteins. Cell **50**:1091–1100.
36. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
37. **Lopata, M. A., D. W. Cleveland, and B. Sollner-Webb.** 1984. High level transient expression of a chloramphenicol acetyl transferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment. Nucleic Acids Res. **12**:5707–5717.
38. **Lyons, R. H., B. Q. Ferguson, and M. Rosenberg.** 1987. Pentapeptide nuclear localization signal in adenovirus E1A. Mol. Cell. Biol. **7**:2451–2456.
39. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
40. **Messing, J., B. Gronenborn, B. Müller-Hill, and P. H. Hofschneider.** 1977. Filamentous coliphage M13 as a cloning vehicle: insertion of a *Hin*dII fragment of the *lac* regulatory region in M13 replicative form *in vitro*. Proc. Natl. Acad. Sci. USA **74**: 3642–3646.
41. **Moran, E., and M. B. Mathews.** 1987. Multiple functional domains in the adenovirus E1A gene. Cell **48**:177–178.
42. **Moran, E., B. Zerler, T. M. Harrison, and M. B. Mathews.** 1986. Identification of separate domains in the adenovirus E1A gene for immortalization activity and the activation of virus early genes. Mol. Cell. Biol. **6**:3470–3480.
43. **Parks, C. L., S. Banarjee, and D. J. Spector.** 1988. Organization of the transcriptional control region of the E1b gene of adenovirus type 5. J. Virol. **62**:54–67.
44. **Perricaudet, M., G. Akusjärvi, A. Virtanen, and U. Pettersson.** 1979. Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses. Nature (London) **281**:694–696.
45. **Pettersson, U., and J. Sambrook.** 1973. Amount of viral DNA in the genome of cells transformed by adenovirus type 2. J. Mol. Biol. **73**:125–130.
46. **Poncz, M., D. Solowiejczyk, M. Ballantine, E. Schwartz, and S. Surrey.** 1982. "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. Proc. Natl. Acad. Sci. USA **79**:4298–4302.
47. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
48. **Schneider, J. F., F. Fisher, C. R. Gooding, and N. C. Jones.** 1987. Mutational analysis of the adenovirus Ela gene: the role of transcriptional regulation in transformation. EMBO J. **6**: 2053–2060.
49. **Schwartz, R. M., and M. O. Dayhoff.** 1978. Matrices for detecting distant relationships, p. 353–358. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure, vol. 5, suppl. 3. National Biomedical Research Foundation, Washington, D.C.
50. **Selden, R. F., K. B. Howie, M. E. Rowe, H. M. Goodman, and D. D. Moore.** 1986. Human growth hormone as a reporter gene in regulation studies employing transient gene expression. Mol. Cell. Biol. **6**:3173–3179.
51. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
52. **Stow, N. D.** 1981. Cloning of a DNA fragment from the left-hand terminus of the adenovirus type 2 genome and its use in site-directed mutagenesis. J. Virol. **37**:171–180.
53. **Straus, S. E.** 1984. Adenovirus infections in humans, p. 451–496. *In* H. S. Ginsberg (ed.), The adenoviruses. Plenum Publishing Corp., New York.
54. **Sugisaki, H., K. Sugimoto, M. Takanami, K. Shiroki, I. Saito, H. Shimojo, Y. Sawada, Y. Uemizu, S. Uesugi, and F. Fujinaga.** 1980. Structure and gene organization in the transforming Hind III-G fragment of adenovirus type 12. Cell **20**:777–786.
55. **Tamanoi, F., and B. W. Stillman.** 1983. Initiation of adenovirus DNA replication in vitro requires a specific DNA sequence. Proc. Natl. Acad. Sci. USA **80**:6446–6450.
56. **Temple, M., G. Antoine, H. Delius, S. Stahl, and E.-L. Winnacker.** 1981. Replication of mouse adenovirus strain FL DNA. Virology **109**:1–12.
57. **Tokunaga, O., T. Yaegashi, J. Lowe, L. Dobbs, and R. Padmanabhan.** 1986. Sequence analysis in the E1 region of adenovirus type 4 DNA. Virology **155**:418–433.
58. **Tooze, J.** 1981. DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
59. **van Ormondt, H., and B. Hesper.** 1983. Comparison of the nucleotide sequences of early region E1b DNA of human adenovirus types 12, 7 and 5 (subgroups A, B and C). Gene **12**: 217–226.
60. **van Ormondt, H., J. Maat, A. de Waard, and A. J. van der Eb.** 1978. The nucleotide sequence of the transforming HpaI-E fragment of adenovirus type 5 DNA. Gene **4**:309–328.
61. **van Ormondt, H., J. Maat, and R. Dijkema.** 1980. Comparison of nucleotide sequences of the early Ela regions for subgroups A, B and C of human adenoviruses. Gene **12**:63–76.
62. **van Ormondt, H., J. Maat, and C. P. van Beveren.** 1980. The nucleotide sequence of the transforming early region E1 of adenovirus type 5 DNA. Gene **11**:299–309.
63. **Weeks, D. L., and N. C. Jones.** 1983. E1A control of gene expression is mediated by sequences 5′ to the transcriptional starts of the early viral genes. Mol. Cell. Biol. **3**:1222–1234.
64. **Weeks, D. L., and N. C. Jones.** 1985. Adenovirus E3-early promoter: sequences required for activation by E1A. Nucleic Acids Res. **13**:5389–5402.
65. **Wigand, R., H. Gelderblom, and M. Özel.** 1977. Biological and biophysical characteristics of mouse adenovirus, strain FL. Arch. Virol. **54**:131–142.
66. **Williams, J. F.** 1970. Enhancement of adenovirus plaque formation on HeLa cells by magnesium chloride. J. Gen. Virol. **9**:251–255.
67. **Wu, L., D. S. E. Rosser, M. C. Schmidt, and A. Berk.** 1987. A TATA box implicated in E1A transcriptional activation of a simple adenovirus 2 promoter. Nature (London) **326**:512–515.