Human adenovirus type 19: genomic and bioinformatics analysis of a keratoconjunctivitis isolate

Christopher M. Robinson[a,c], **Fatemeh Shariati**[a,b], **Jeremy Zaitshik**[e], **Allison F. Gillaspy**[c,e], **David W. Dyer**[c,e], and **James Chodosh**[a,b,c,d,*]

a*Molecular Pathogenesis of Eye Infection Research Center, Dean A. McGee Eye Institute, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma*

b*Department of Ophthalmology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma*

c*Department of Microbiology & Immunology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma*

d*Department of Cell Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma*

e*Laboratory for Genomics and Bioinformatics, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma*

Abstract

Human adenovirus type 19 (HAdV-19) is a major etiologic agent of epidemic keratoconjunctivitis (EKC), a common and severe eye infection associated with long-term visual morbidity due to persistent corneal inflammation. Ironically, while the prototype strain of HAdV-19 does not cause eye infections, other isolates of the serotype have caused major outbreaks of EKC. Here we have sequenced a clinical isolate of HAdV-19 (HAdV-19 strain C) from a human patient with EKC. Global pairwise alignment of HAdV-19 C to other HAdV species D serotypes identified areas of sequence divergence in the penton base (host cell internalization signal), hexon (principal viral capsid structural protein), E3 (site of immunomodulatory genes), and fiber (host cell binding ligand) regions. Comparison of HAdV-19 strain C to the recently sequenced HAdV-37, another EKC causing serotype, identified sequence diversity in the penton base and hexon, but sequence conservation in the E3 and fiber regions. Elucidation of the HAdV-19 C genome will facilitate future studies into the pathogenesis of EKC, and may shed light on the genetic determinants of corneal tropism.

Keywords

human adenovirus type 19; epidemic keratoconjunctivitis; genome

Human adenoviruses (HAdV) belong to the genera of *Mastadenovirus* within the family of *Adenoviridae* and cause a wide array of clinical diseases including acute respiratory infections, gastroenteritis, and ocular surface infections (Dingle and Langmuir, 1968; Harding et al., 1988; Wood, 1988). AdVs have a linear, double stranded DNA genome that ranges from 26

to 46 kb in size. HAdVs were first identified in human adenoids and characterized by two research teams (Hilleman and Werner, 1954; Rowe et al., 1953). There are 51 known human serotypes, and these have been classified into 6 species (A-F), based on restriction enzyme analysis and hemaglutination assays, later confirmed by genomic and phylogenetic analyses. A fifty-second serotype has been recently proposed and classified into a new species G (Jones et al., 2007).

HAdV-19 was originally isolated in 1955 from a child with trachoma in Saudi Arabia (Bell et al., 1959; Bell et al., 1960). Ironically, while the prototype strain of HAdV-19 does not cause eye infections, other isolates of HAdV-19 serotype later became known as major etiological agents of epidemic keratoconjunctivitis (EKC) (Desmyter et al., 1974; Hierholzer et al., 1974), a severe and highly contagious infection associated with long term visual morbidity (Butt and Chodosh, 2006). HAdV-19 isolates were first reported in outbreaks of keratoconjunctivitis in 1973.

HAdV species D contains the most serotypes of any human adenovirus species, yet relatively few have been completely sequenced. Our lab recently sequenced and annotated the HAdV-37 genome, another EKC-associated serotype (Robinson et al., 2008). In this study, we have sequenced and annotated HAdV-19 strain C (HAdV-19 C), an isolate collected directly from a cornea of a patient with EKC (Chodosh et al., 1995), and described its overall genomic organization. Global pairwise genome alignment with other HAdVs species D revealed areas of non-conserved sequence in the penton, hexon, E3, and fiber regions. Global pairwise comparison of HAdV-19 C and HAdV-37 revealed differences in the penton and hexon regions, but conserved sequence in the E3 and fiber regions. Phylogenetic analysis of whole genome sequence further confirmed the close similarity of HAdV-19 C and HAdV-37.

Standard PCR methodology was used to amplify regions of the HAdV-19 C genome to be sequenced. The sequencing reactions and genome assembly were performed as previously described (Robinson et al., 2008) and in the supplementary methods.

The genome of HAdV-19 C was determined to be 35 231 base pairs in length. The finished genome assembly contained 722 high quality reads with an average length of 949 base pairs. The fold coverage for both strands of the genome was 19. The Phrap average quality score was 89.4. The nucleotide composition of HAdV-19 C is 22.9% A, 20.6% T, 28.3% C, and 28.2% G. The GC content of 56.5% is on the lower end of the 57-59% previously reported for HAdV species D (Shenk, 1996).

Genome annotation was performed using JCVI's automated annotation system (http://www.jcvi.org/cms/research/projects/annotation-service), GeneMark Heuristic Model gene prediction (http://exon.gatech.edu/GeneMark), and NCBI's ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Splice sites were predicted using a splice site finder program (http://www.genet.sickkids.on.ca/~ali/splicesitefinder.html). Open reading frames were searched against available databases in GenBank, PIR, SWISS-PROT, and JCVI's CMR database. We identified 4 early, 2 intermediate, and 5 late transcription units, similar to those previously identified in other sequenced HAdVs, including 35 open reading frames (ORFs) and 8 hypothetical ORFs. The 5′ and 3′ ends of the HAdV genome are composed of inverted terminal repeat (ITR) sequence which for HAdV-19 C were determined to be 160 base pairs in length each.

E1A is the first transcript unit expressed during adenovirus infection (Shenk, 1996). Several alternatively spliced E1A transcripts are produced from a common transcript from this region (Berk and Sharp, 1978). Based on splice donor and acceptor sites, two ORFs corresponding to putative proteins of 191 and 253 amino acids were identified in the HAdV-19 C genome (Table

1). The putative TATA box for this transcription unit was identified at nucleotide 480 and the polyadenylation signal predicted to be at position 1452.

Proteins expressed by the E1B transcription unit block apoptosis and catalyze viral replication (Shenk, 1996). ORFs encoding putative proteins of 182 and 495 amino acids in length were identified in HAdV-19 C, corresponding to the 19- and 55-kDa proteins previously reported for other sequenced HAdVs. The polyadenylation signal for these transcripts was predicted to begin at nucleotide 3863.

The E2 region of the genome consists of three transcripts broken down into two groups, E2A and E2B. These transcripts produce three proteins that are essential for viral replication known as the DNA binding protein (DBP, E2A), terminal protein precursor (pTP, E2B), and DNA polymerase (E2B) (Shenk, 1996). The E2 genes were identified on the complementary strand and predicted to be 490, 635, and 1091 amino acids in length, respectively (Table 1).

The E3 region within HAdV genomes encodes proteins that modulate the host immune response to infection and are not required for viral growth *in vitro* (Horwitz, 2004; Windheim et al., 2004). We identified seven classical ORFs within this region of the HAdV-19 C genome. The TATA box for this transcription unit was predicted to begin at nucleotide 25896 and a polyadenylation site was identified at nucleotide 30853.

The E4 transcription unit ORFs produce proteins with a wide variety of functions during HAdV infection (Leppard, 1997). We found 6 predicted E4 ORFs in the HAdV-19 C genome located on the complementary strand. The E4 ORF 1 from HAdV-19 C was predicted to produce a protein of 65 amino acids in length similar to our previous report for HAdV-37 (Robinson et al., 2008).

The intermediate genes of HAdV consist of IVa2 and IX. We identified the spliced product of IVa2 in the HAdV-19 C genome using the splice site finder (http://www.genet.sickkids.on.ca/~ali/splicesitefinder.html). The IVa2 gene was located on the complementary strand. The predicted protein of 448 amino acids in length from the IVa2 ORF was 99% identical to the HAdV-9 homologue. The ORF corresponding to the IX protein was also identified on the complementary strand, and was identified at nucleotides 3455-3859.

Transcribed from the major late promoter (MLP), the late transcription units express proteins that are involved in capsid production for mature virions (Shenk, 1996). The L1 transcription unit encodes two proteins, 52/55K and IIIa. These proteins were predicted in HAdV-19 C to be 374 and 565 amino acids in length, respectively. The polyadenylation site for this transcription unit was predicted to start at nucleotide 13490.

The L2 transcription unit encodes four proteins involved in capsid formation: penton base, V, VII, and X. The gene for HAdV-19 C penton base is located at nucleotides 13537-15081, corresponding to a predicted protein of 514 amino acids. Arg-Gly-Asp (RGD) sequence in the penton base interacts with host cell integrins to induce internalization of the virus (Wickham et al., 1993), and sequence for this peptide was identified in the HAdV-19 C homologue at residues 304-306. The predicted penton base protein was 99% identical to that of a previously predicted penton base for another clinical isolate of HAdV-19, HAdV-19a (GenBank: **AAG00908**), differing in only two amino acids (Arnberg et al., 2000; Wadell and de Jong, 1980). A valine and threonine were predicted for HAdV-19 C at positions 373 and 374, respectively, in contrast to glycine and histidine in HAdV-19a. However, the sequence we determined for HAdV-19 C in this region was identical to that of HAdV-2, 3, 4, 8, 12, 37, and 40, representing all 6 HAdV species, as well as simian AdV-25 (Madisch et al., 2007; Robinson et al., 2008). It is not clear why the previously published HAdV-19a sequence differs from our current findings. The V, VII, and X proteins were predicted to have molecular weights of 37.9,

21.7, and 8.2-kDa, respectively. The L2 transcripts share a putative polyadenylation signal beginning at nucleotide 16967.

Within the L3 transcription unit, we identified three classical open reading frames corresponding to the VI, hexon, and protease proteins. The HAdV-19 C VI ORF was located at nucleotides 17016-17720, corresponding to a protein with predicted molecular weight of 25.5-kDa, and 99% identical to its HAdV-48 homologue. The HAdV-19 C hexon ORF was predicted to give rise to a protein 961 amino acids in length with a predicted molecular weight of 108-kDa, 99% identical to the previously published HAdV-19 hexon protein sequence (GenBank: **ABA00002**). A predicted HAdV-19 C protease ORF was also identified at nucleotides 20649-21278 corresponding to 209 amino acids.

Within the L4 transcription unit of HAdV-19 C, we identified 3 ORFs corresponding to the 100-kDa, 22K, and VIII proteins. A 732 amino acid 100-kDa protein was predicted from the ORF at nucleotides 22811-25009. The putative HAdV-19 C protein was 98% identical to the HAdV-46 homologue. The 22K protein was predicted to be 137 amino acids in length. The VIII protein ORF was predicted at nucleotides 25531-26214. The 227 amino acid protein has a predicted molecular weight of 24.6-kDa. The polyadenylation site for the transcription site for the VIII protein was identified at nucleotide 26524.

The L5 transcription unit consists of only one ORF, encoding the fiber protein. This protein is the primary ligand for host cell binding. The HAdV-19 C putative fiber ORF was identified at nucleotides 31055-32152 and corresponds to a protein of 365 amino acids in length. The HAdV-19 C fiber was found to be 100% identical to the previously published HAdV-19 fiber (GenBank: **AAB71733**) as well as 100% identical to the HAdV-37 fiber (GenBank: **AAB71734**).

During our annotation of the HAdV-19 C genome, we identified 8 hypothetical ORFs similar to ORFs previously archived in GenBank for other HAdV genomes (Table 1). Blast scores for these putative proteins were all determined to be less than e^{-5}. Five of the hypothetical ORFs were located on the sense strand, and 3 on the complementary strand.

An online sequence alignment program, mVISTA Limited Area Global Alignment of Nucleotides (LAGAN) (http://genome.lbl.gov/vista/index.shtml) was used for global pairwise sequence alignment to compare paired viral sequences (Brudno et al., 2003). We compared genomic sequence correspondence across the whole genome of HAdV-19 C to the 7 completely sequenced HAdV species D serotypes (Fig. 1). Comparison of HAdV-19 C to HAdV-9, 17, 26, 46, 48, and 49 (Table 2) revealed areas of sequence divergence in the penton, hexon, E3, and fiber regions. Direct comparison between HAdV-19 C and HAdV-37, another EKC serotype, revealed divergence in the penton and hexon regions, but conservation in the E3 and fiber regions.

Phylogenetic analysis of the whole HAdV-19 C genome compared with other completely sequenced AdVs located within GenBank allowed us to view HAdV-19 C in the context of adenoviral evolution. A bootstrap confirmed neighbor joining tree further confirmed the close sequence relationship with HAdV-37 and other HAdV species D viruses (Fig. 2).

In summary, we determined the 35 231 base pair genome of a clinical HAdV-19 isolate, taken directly from a patient with EKC. Through annotation, we identified 35 putative HAdV-19 C genes along with 8 hypothetical ORFs. Comparison of the HAdV-19 C genome to other HAdV species D identified areas of sequence divergence in the penton, hexon, E3, and fiber regions. Comparison of HAdV-19 C to another cornea tropic adenovirus, HAdV-37, revealed divergence in only the penton and hexon regions. The similarity between HAdV-19 C and HAdV-37 in the E3 and fiber regions was expected and similar to previous reports (Arnberg

et al., 1997; Burgert and Blusch, 2000). These data suggests the possibility of a prior recombination event or genetic drift to account for the very similar genomes of these two viruses. Identification of identical E3 and fiber regions in HAdV-19 and 37 compared to other non-cornea tropic viruses distinguishes these as possible determinants for corneal tropism. HAdV-37 fiber protein interacts with CD46, its primary receptor, and this interaction has been thought to restrict tissue tropism (Wu et al., 2001; Wu et al., 2004). However, the use of CD46 for viral binding cannot alone account for corneal tropism, because HAdV species B have been shown to also use CD46 for viral binding and do not typically cause corneal infections. Therefore, it seems likely that other regions in the HAdV genome work alone or in conjunction with the fiber protein to confer corneal tropism. Our analysis revealed the E3 region of HAdV-19 C to be identical to that of HAdV-37, but highly divergent from other HAdV species D viruses. The coding sequence within this area is also not well conserved across other HAdV species and could represent an important effector of tissue tropism across the entire spectrum of adenoviral infections.

Whole genome sequencing of human adenoviruses is vital to understanding adenoviral evolution. Past techniques for identification of serotypes through serological methods provided a less distinct picture of the true diversity of HAdVs, and phylogenetic trees built solely on single genes reveal only part of the evolutionary story. Sequence variance between related serotypes likely represents selection events during viral evolution. The recent identification of a natural recombinant of HAdV-22, HAdV-8, and HAdV-37 (Aoki et al., 2008; Engelmann et al., 2006) may allow determination of those genes important to the pathogenesis of EKC, specifically the genetic foundations of corneal tissue tropism. Further analysis of divergent regions in the adenoviral genome may help to identify new therapeutic targets for adenoviral infections, and allow for more targeted gene therapy using adenoviral vectors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Aoki K, Ishiko H, Konno T, Shimada Y, Hayashi A, Kaneko H, Ohguchi T, Tagawa Y, Ohno S, Yamazaki S. Epidemic keratoconjunctivitis due to a novel hexon-chimeric intermediate human adenovirus 22,37/H8. J Clin Microbiol. 2008

Arnberg N, Kidd AH, Edlund K, Olfat F, Wadell G. Initial interactions of subgenus D adenoviruses with A549 cellular receptors: sialic acid versus alpha(v) integrins. J Virol 2000;74(16):7691–3. [PubMed: 10906228]

Arnberg N, Mei Y, Wadell G. Fiber genes of adenoviruses with tropism for the eye and the genital tract. Virology 1997;227(1):239–44. [PubMed: 9007080]

Bell SD Jr, Mc CD, Murray ES, Chang RS, Snyder JC. Adenoviruses isolated from Saudi Arabia. I. Epidemiologic features. Am J Trop Med Hyg 1959;8(4):492–500. [PubMed: 13670377]

Bell SDJ, TF R, McComb DE. Adenoviruses isolated from Saudi Arabia. II. Six new serotypes. Am J Trop Med Hyg 1960;9:523–526.

Berk AJ, Sharp PA. Structure of the adenovirus 2 early mRNAs. Cell 1978;14(3):695–711. [PubMed: 688389]

Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Green ED, Sidow A, Batzoglou S. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. Genome Res 2003;13(4):721–31. [PubMed: 12654723]

Burgert HG, Blusch JH. Immunomodulatory functions encoded by the E3 transcription unit of adenoviruses. Virus Genes 2000;21(12):13–25. [PubMed: 11022786]

Butt AL, Chodosh J. Adenoviral keratoconjunctivitis in a tertiary care eye clinic. Cornea 2006;25(2): 199–202. [PubMed: 16371782]

Chodosh J, Miller D, Stroop WG, Pflugfelder SC. Adenovirus epithelial keratitis. Cornea 1995;14(2): 167–74. [PubMed: 7743800]

Desmyter J, De Jong JC, Slaterus KW, Verlaeckt H. Letter: Keratoconjunctivitis caused by Adenovirus Type 19. Br Med J 1974;4(5941):406. [PubMed: 4425899]

Dingle JH, Langmuir AD. Epidemiology of acute, respiratory disease in military recruits. Am Rev Respir Dis 1968;97(6):1–65. [PubMed: 4869845]

Engelmann I, Madisch I, Pommer H, Heim A. An outbreak of epidemic keratoconjunctivitis caused by a new intermediate adenovirus 22/H8 identified by molecular typing. Clin Infect Dis 2006;43(7):e64–6. [PubMed: 16941356]

Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 1985;39:783–791.

Harding SP, Mutton KJ, van der Avoort H, Wermenbol AG. An epidemic of keratoconjunctivitis due to adenovirus type 37. Eye 1988;2(Pt 3):314–7. [PubMed: 2841171]

Hierholzer JC, Guyer B, O'Day D, Schaffner W. Letter: Adenovirus type 19 keratoconjunctivitis. N Engl J Med 1974;290(25):1436. [PubMed: 4364287]

Hilleman MR, Werner JH. Recovery of new agent from patients with acute respiratory illness. Proc Soc Exp Biol Med 1954;85(1):183–8. [PubMed: 13134329]

Horwitz MS. Function of adenovirus E3 proteins and their interactions with immunoregulatory cell proteins. J Gene Med 2004;6:S172–83. [PubMed: 14978760]

Jones MS 2nd, Harrach B, Ganac RD, Gozum MM, Dela Cruz WP, Riedel B, Pan C, Delwart EL, Schnurr DP. New adenovirus species found in a patient presenting with gastroenteritis. J Virol 2007;81(11): 5978–84. [PubMed: 17360747]

Leppard KN. E4 gene function in adenovirus, adenovirus vector and adeno-associated virus infections. J Gen Virol 1997;78(Pt 9):2131–8. [PubMed: 9291998]

Madisch I, Hofmayer S, Moritz C, Grintzalis A, Hainmueller J, Pring-Akerblom P, Heim A. Phylogenetic analysis and structural predictions of human adenovirus penton proteins as a basis for tissue-specific adenovirus vector design. J Virol 2007;81(15):8270–81. [PubMed: 17522221]

Robinson CM, Shariati F, Gillaspy AF, Dyer DW, Chodosh J. Genomic and bioinformatics analysis of human adenovirus type 37: new insights into corneal tropism. BMC Genomics 2008;9:213. [PubMed: 18471294]

Rowe WP, Huebner RJ, Gilmore LK, Parrott RH, Ward TG. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. Proc Soc Exp Biol Med 1953;84(3):570–3. [PubMed: 13134217]

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4(4):406–25. [PubMed: 3447015]

Shenk, T. Adenoviridae: The Viruses and Their Replication. In: K, DM.; Fields, BN.; Howle, PM., editors. Fields Virology. Vol. Thrid Edition. Lippincott - Raven Publishers; Philadelphia: 1996. p. 2111-2148.

Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007;24(8):1596–9. [PubMed: 17488738]

Wadell G, de Jong JC. Restriction endonucleases in identification of a genome type of adenovirus 19 associated with keratoconjunctivitis. Infect Immun 1980;27(2):292–6. [PubMed: 6247272]

Wickham TJ, Mathias P, Cheresh DA, Nemerow GR. Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. Cell 1993;73(2):309–19. [PubMed: 8477447]

Windheim M, Hilgendorf A, Burgert HG. Immune evasion by adenovirus E3 proteins: exploitation of intracellular trafficking pathways. Curr Top Microbiol Immunol 2004;273:29–85. [PubMed: 14674598]

Wood DJ. Adenovirus gastroenteritis. Br Med J (Clin Res Ed) 1988;296(6617):229–30.

Wu E, Fernandez J, Fleck SK, Von Seggern DJ, Huang S, Nemerow GR. A 50-kDa membrane protein mediates sialic acid-independent binding and infection of conjunctival cells by adenovirus type 37. Virology 2001;279(1):78–89. [PubMed: 11145891]

Wu E, Trauger SA, Pache L, Mullen TM, von Seggern DJ, Siuzdak G, Nemerow GR. Membrane cofactor protein is a receptor for adenoviruses associated with epidemic keratoconjunctivitis. J Virol 2004;78 (8):3897–905. [PubMed: 15047806]
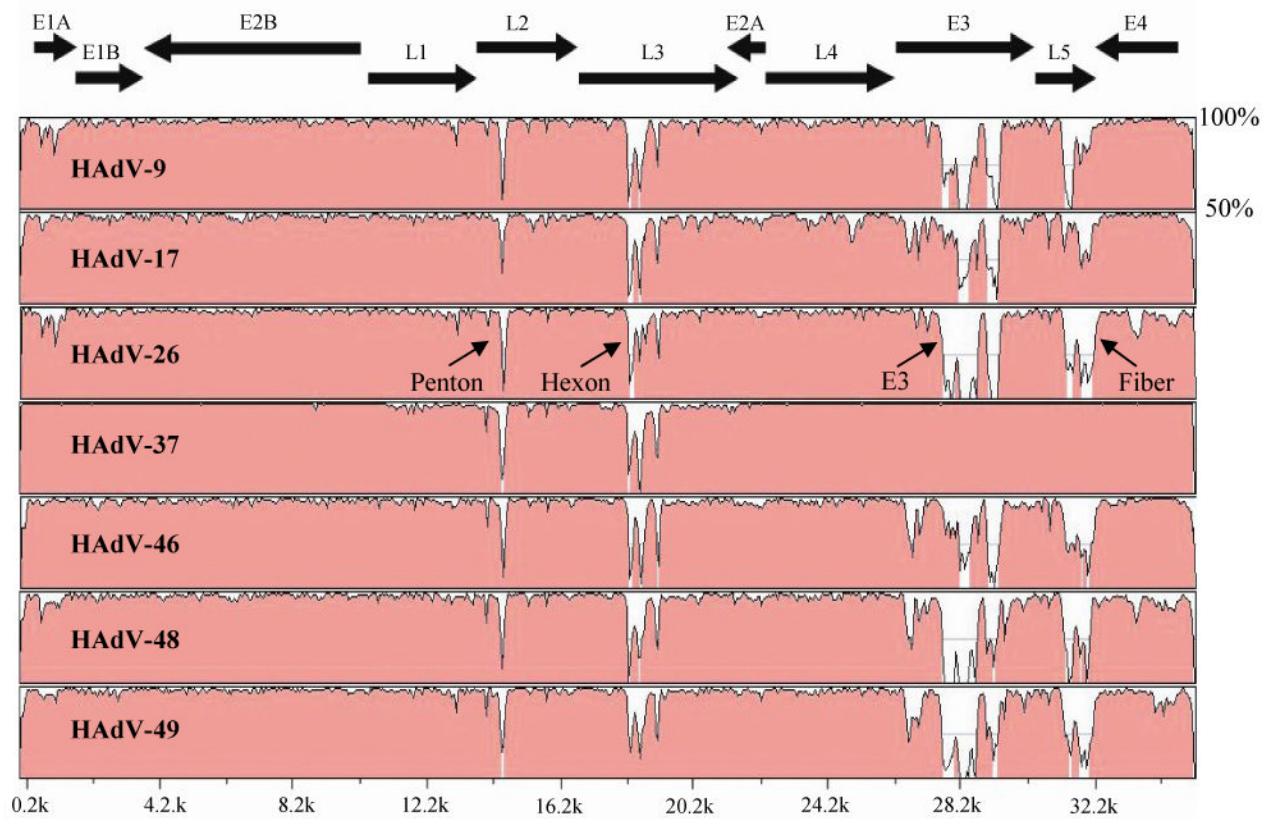
Fig. 1.
Global pairwise sequence comparison of HAdV-19 C with 7 other HAdV species D serotypes using the online sequence alignment program, mVISTA LAGAN. Percent sequence conservation is reflected in the height of each data point along the y axis. The penton, hexon, E3, and fiber regions of HAdV-19 C diverged from those of all other serotypes within species D, except for HAdV-37. The E3 and fiber regions were 100% identical between HAdV-19 and HAdV-37.
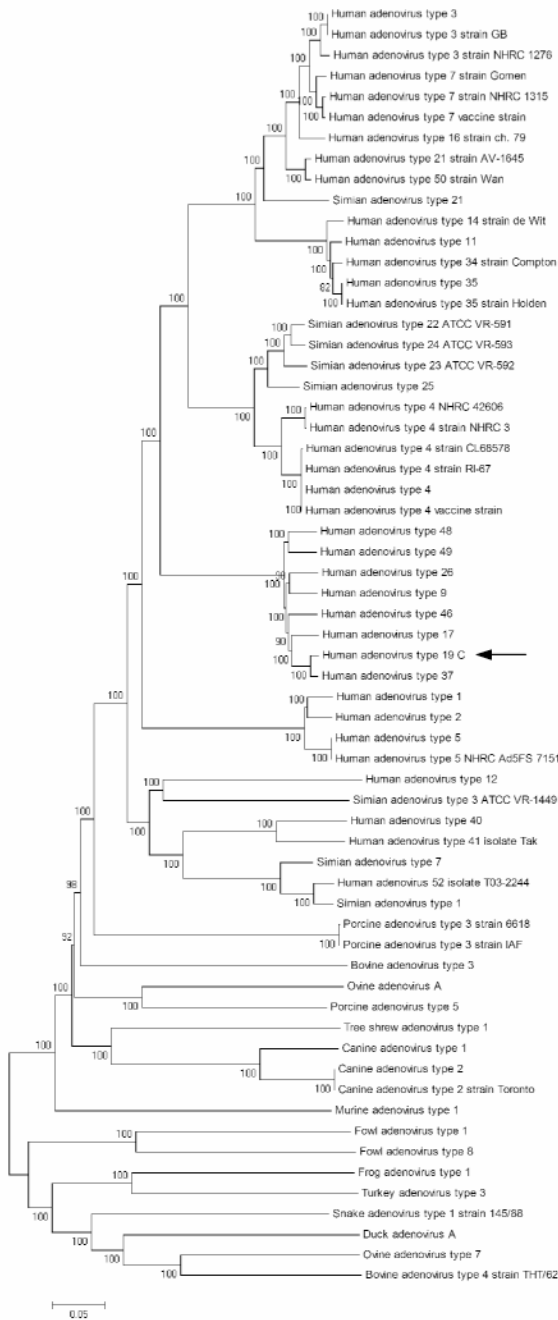
Fig. 2.
Bootstrap confirmed neighbor joining tree designed from MEGA 4.0.2 demonstrate phylogenetic relationships between HAdV-19 C (arrow) and all other completely sequenced adenovirus genomes (Saitou and Nei, 1987; Tamura et al., 2007). The nucleotide p-distance model was used, with pairwise deletion options. Bootstrap confidence levels (1000 replicates) are shown as percentages on each branch (Felsenstein, 1985).

Table 1

Genome Organization of HAdV-19 C

Region	Gene Product	TATA Box	Location	Protein Length (aa)	Poly(A) signal
E1A	21.2 kDa	480	570-933, 1215-1426 *	191	1452
	28.2 kDa	480	570-1119, 1215-1426 *	253	1452
E1B	21.1 kDa	1519	1578-2126	182	3863
	55.3 kDa	1519	1883-3370	495	3863
Intermediate	IX	ND	3455-3859	134	3863
	IVa2	ND	3903-5236, 5515-5527 c *	448	ND
E2B	DNA polymerase	ND	5006-8281 c	1091	ND
	Hypothetical 20.0 kDa	ND	5200-5766 c	188	ND
	Hypothetical 9.6 kDa	ND	6150-6419	89	ND
	Hypothetical 34.7 kDa	ND	6515-7504	329	ND
	Hypothetical 16.6 kDa	ND	7787-8254	155	ND
	Hypothetical 19.2 kDa	ND	8031-8564 c	177	ND
E2B	pTP	ND	8326-10224, 13502-13510 c *	635	ND
	Hypothetical 11.4 kDa	ND	9649-9984	111	ND
L1	52/55K	5827	10639-11763	374	13490
	pIIIa	5827	11786-13483	565	13490
L2	penton base	5827	13537-15081	514	16967
	pVII	5827	15085-15675	196	16967
	pV	5827	15708-16706	332	16967
	X	5827	16736-16960	74	16967
L3	pVI	5827	17016-17720	234	ND
	hexon	5827	17761-20646	961	ND
	protease	5827	20649-21278	209	ND
E2A	DNA binding protein	ND	21322-22794 c	490	ND
L4	100K	5827	22811-25009	732	26524
	Hypothetical 9.2 kDa	ND	23953-24198 c	81	ND
	22K	5827	24792-25205	137	26524
	pVIII	5827	25531-26214	227	26524
E3	12.1 kDa	25896	26215-26535	106	30853
	21.8 kDa	25896	26489-27079	196	30853

Region	Gene Product	TATA Box	Location	Protein Length (aa)	Poly(A) signal
	18.6 kDa	25896	27046-27546	166	30853
	48.9 kDa	25896	27571-28857	428	30853
	Hypothetical 31.6 kDa	25896	28884-29726	280	30853
	10.4 kDa	25896	29733-30008	91	30853
	14.7 kDa	25896	30011-30403	130	30853
	14.8 kDa	25896	30396-30788	130	30853
L5	fiber	5827	31055-32152	365	32159
E4	34.0 kDa	34685	32428-33306 c	292	32200
	17.1 kDa	34685	32684-33139 c	151	32200
	14.0 kDa	34685	33236-33598 c	120	32200
	13.5 kDa	34685	33601-33954 c	117	32200
	14.5 kDa	34685	33951-34343 c	130	32200
	7.4 kDa	34685	34384-34581 c	65	32200

Note: c, complementary strand

*, predicted splicing

ND, Not determined

Table 2

Nucleotide accession numbers for fully sequenced HAdV species D viruses

Human Adenovirus (HAdV) serotype	GenBank Accession Number
HAdV-9	AJ854486
HAdV-17	AC_000006
HAdV-19 C	EF121005
HAdV-26	EF153474
HAdV-37	DQ900900
HAdV-46	AY875648
HAdV-48	EF153473
HAdV-49	DQ393829