3-14-1994

Phylogenetic analysis of segment 10 from African horsesickness virus and cognate genes from other orbiviruses

Rafael O. de Sá
University of Richmond, rdesa@richmond.edu

Marla Zellner

Marvin J. Grubman

Recommended Citation

de Sá, R. O. M. Zellner, and M. Grubman. 1994. Phylogenetic analysis of segment 10 from African Horsesickness Virus and cognate genes from other orbiviruses. Virus Research 33:157-165.

ELSEVIER

Virus Research 33 (1994) 157–165

Virus Research

Phylogenetic analysis of segment 10 from African horsesickness virus and cognate genes from other orbiviruses

Rafael O. de Sá [a], Marla Zellner [b], Marvin J. Grubman [b,*]

[a] *Department of Biology, University of Richmond, Richmond, VA 23173, USA*
[b] *United States Department of Agriculture, ARS, NAA Plum Island Animal Disease Center, P.O. Box 848, Greenport, NY 11944, USA*

(Received 5 November 1993; Revision received 11 March 1994; Accepted 14 March 1994)

Abstract

Utilizing the reverse transcriptase-polymerase chain reaction (RT-PCR) procedure, we have synthesized full-length copies of segment 10 from African horsesickness virus (AHSV) serotypes 1, 4 and 8. The genes were cloned, sequenced and compared with the sequence of the cognate gene from AHSV serotypes 3 and 9. Sequences were analyzed to assess evolutionary relationships among serotypes using cladistics. Based on this analysis the data support a close relationship between serotypes 4 and 9 and between serotypes 1 and 8 and a closer relationship of serotype 3 to the 4 and 9 group.

Key words: African horsesickness virus; Segment 10; Polymerase chain reaction; Evolutionary relationships

African horsesickness virus (AHSV), a member of the genus *Orbivirus,* family Reoviridae, contains a genome consisting of ten segments of double-stranded RNA (dsRNA) (Oellermann et al., 1970). Nine serological types of the virus have been identified (Howell, 1962). Within the orbivirus genus there are 11 serological groups (Gorman and Taylor, 1985). Two of these groups, AHSV and bluetongue virus (BTV), are of economic interest because they are the causes of important infectious diseases of livestock.

The nucleotide sequence of individual segments of a few AHSV serotypes have been determined (Mizukoshi et al., 1992; Roy et al., 1991; Van Staden and

* Corresponding author. Fax: +1 (516) 323-2507.

SSDI 0168-1702(94)00033-9

Huismans, 1991; Van Staden et al., 1991). However, phylogenetic relationships within the AHSV serogroup have not been reported. Bremer et al. (1990) used Northern blot hybridization analysis to determine intra- and inter-serogroup nucleic acid similarities using AHSV-3 cDNA clones representing all 10 segments as probes. As expected, they found that segment 2 coding for the major outer capsid protein, VP2 (Grubman and Lewis, 1992), is serotype-specific. Somewhat surprisingly, they found that segment 10, coding for a non-structural protein (Grubman and Lewis, 1992), only hybridized to itself and its cognate gene from serotype 7.

In this study, we report nucleotide sequence information for segment 10 from AHSV serotypes 1, 4, and 8. Based on this information and previously available sequences for this segment from AHSV serotypes 3 and 9 as well as for other *Orbivirus* serogroups, we present a phylogenetic analysis of AHSV.

The AHSV serotypes used in this study were provided by Drs. Carol House and James House, Foreign Animal Disease Diagnostic Laboratory, U.S. Department of Agriculture, Greenport, New York. They were originally obtained in 1968 from the Veterinary Research Institute, Onderstepoort, South Africa as low passage mouse brain extracts and were passed a number of times in equine dermis cells and/or Vero cells prior to plaque purification. Virus stocks were prepared by growing plaque-purified virus in Vero cells as previously described (Grubman and Lewis, 1992). Partially purified virus was prepared by treatment of stock virus with Triton X-100, centrifugation through a 40% (w/w) sucrose cushion, extraction with freon and centrifugation through a sucrose cushion as above. Virion RNA was isolated from partially purified virus or from the cytoplasm of infected cells by standard procedures.

Utilizing the sequences available for segment 10 from AHSV-3 (Van Staden and Huismans, 1991), we designed a pair of specific primers from sequences at the 5′ and 3′ ends of this segment. To facilitate subsequent cloning these primers were constructed to contain restriction sites. The primers were the 5′ and 3′ oligonucleotides AGTCGACGAATTCGTTTAAATTATCCCTTGTC and AGTCGACGGATCCGTAAGTCGTTATCCCGGCT, respectively. The underlined sequences correspond to *Eco*RI and *Bam*HI sites, respectively. Genomic RNA, 1 μg, or 2 μg total RNA from infected cells, was denatured for 10 min at room temperature with 10 mM methyl mercury hydroxide and reverse transcribed in the presence of the above primers with Moloney murine leukemia virus reverse transcriptase (BRL) at 37°C for 60 min. DNA was amplified with Taq polymerase (Perkin Elmer Cetus) for 30 cycles at 95°C for 1 min, 60°C for 1 min, and 75°C for 2 min and 1 cycle at 75°C for 4 min. The DNA was purified by agarose gel electrophoresis, digested with *Eco*RI and *Bam*HI (New England Biolabs), ligated

→

Fig. 1. Nucleotide sequences of segment 10 from AHSV serotypes 1, 3, 4, 8, and 9, and Palyam virus are presented. The nucleotide sequence representing the protein coding region was aligned based on the amino acid alignment, while the 5′ and 3′ noncoding regions were aligned to maximum similarity. Numbers above the sequence correspond to: 1–19, 5′ non-coding region; 20–691, protein coding sequences; 692–782, 3′ noncoding region. Underlined sequences correspond to the first and second ATG codon positions. N indicates regions that were not determined (Van Staden and Huismans, 1991).

```
          1                   19
          |                   |
PALYAM    GTTAAAATAATCAC-CGAC
AHSV-3    GTTTAAATTATCCCTTGTC
AHSV-4    GTTTAAATTATCCCTTGTC
AHSV-9    GTTTAA-TTATCCCTTGTC
AHSV-1    GTTTAAATTATCCCTTGTC
AHSV-8    GTTTAAATTATCCCTTGTC

          20                                  49                                  79
          |                                   |                                   |
PALYAM    ATG---TTGGCGCGGTCGCTAAATGAGTAT AAAGCGATGAGATCTGAATCGGAAATGAGT
AHSV-3    ATGAGTCTAGCTACGATCGCCGAAAATTAT ---ATGATGCATAATGGAAATCAGAGAGCA
AHSV-4    ATGAATCTAGCTGCAATCGCCAAGAATTAT ---AGTATGCATAATGGAGAGTCGGGGGCG
AHSV-9    ATGAATCTAGCTGCAATCGCCGAAAATTAT ---AGTATGCATAATGGAGAGTCGGGGGCG
AHSV-1    ATGAATCTTGCTAGCATCTCCCAAAGCTAT ---ATGTCACATAATGAGAATGAAAGATCA
AHSV-8    ATGAATCTTGCTAGCATCTCCCAAAGCTAT ---ATGTCACATAATGAGAATGAAAGATCA

PALYAM    GTTGTTCCATACCAACCACCGGCGTAT--- CCGACGGCTCCAGTGGGATTGAGAAGTGAA
AHSV-3    ATTGTACCGTATGTTCCACCCCCTTATGCG TATGCAAATGCTCCGACGCTTGGTGGTCAG
AHSV-4    ATTGTCCCTTATGTGCCACCGCCATATAAT TTCGCGAGCGCTCCGACGTTTTCTCAGCGT
AHSV-9    ATCGTCCCTTATGTGCCACCACCATACAAT TTCGCAAGTGCTCCGACGTTTTCTCAGCGT
AHSV-1    ATTGTACCATACATTCCGCCACCGTAT--- CATCCGACGGCTCCGGCGCTTGCTGTATCC
AHSV-8    ATTGTACCATACATTCCGCCACCGTAT--- CATCCGACGGCTCCGGCGCTTGCTGTATCC

PALYAM    ---AAAGATCTAAGTGGCATATCCCTGGGG GTTCTAAATAATGCGATGAACGATACCACA
AHSV-3    GCGGGTGAAATGGAGTCCATGTCGCTTGGG ATACTTAATCAAGCCATGTCAAGTACAACT
AHSV-4    ACGAGTCAAATGGAGTCCGTGTCGCTTGGG ATACTTAACCAAGCCATGTCAAGTACAACT
AHSV-9    ACGAGTCAAATGGAGTCCGTGTCGCTTGGG ATACTTAACCAAGCCATGTCAAGTACAACT
AHSV-1    GCCAGTCAAATGGAGACCATGTCGCTTGGG ATACTTAACCAAGCAATGTCAAGTTCAGCT
AHSV-8    GCCAGTCAAATGGAGACCATGTCGCTTGGG ATACTTAACCAAGCAATGTCAAGTTCAGCT

PALYAM    GCGGCGACCCAAGCAGAGAGAGAAGAGAAA GTTGCGTATGCATCGTTCGCTGAAGCGTTG
AHSV-3    GGTGCAAGTCGGGCTCTTAAGGATGAAAAA GCAGCGTTTGGTGCGATGGCGGAAGCATTA
AHSV-4    GGTGCGAGTGGGGCGCTTAAAGATGAAAAA GCAGCGTTCGGTGCTATGGCGGAACGATTG
AHSV-9    GGTGCGAGTGGGGCGCTTAAAGATGAAAAA GCGGCATTTGGTGCCATGGCGGAAGCATTG
AHSV-1    GGTGCGAGCGGAGCACTTAAGGATGAAAAG GCAGCGTATGGAGCGGTGGCAGAGGCGTTG
AHSV-8    GGTGCGAGTGGCGCACTTAAGGATGAAAAG GCAGCGTTTGGAGCGGTGGCGGAGGCGTTG

PALYAM    AGAGATCCGATGTGCGTGCGAGAGATCAAG AAGCGGGTTTCATCACGAACGATAGTTGCG
AHSV-3    CGTGATCCAGAACCGATACGTCAAATAAAG AAACATGTTGGATTAAGAACGCTCAAGCAT
AHSV-4    CGTGATCCAGAACCCATACGTCAAATTAAA AAGCAGGTGGGTATCAGAACTTTAAAGAAC
AHSV-9    CGTGATCCAGAACCCATACGTCAAATTAAA AAGCAGGTGGGTATCAGAACTTTAAAGAAC
AHSV-1    AGAGATCCGGAGCCGATCAGAAAAATTAAG CGACAAGTAGGTATCCAAACTCTAAAAACA
AHSV-8    AGAGATCCGGAGCCGATCAGAAAAATTAAG CGACAAGTAGGTATCCAAACTCTAAAAACA

PALYAM    TTGGAAAAAGAGTATCGTCACCAAAAAAGG ATATATGATTTTGTCCGGCTGATTTTATTT
AHSV-3    TTAAAGATAGAGTTGGCGTCAATGAGACGT AGGTATGCGATACTACGTGTAGTGATCTTT
AHSV-4    TTGAAGATGGAGTTAGCAACAATGCGTCGA AAAAAGTCGGCATTAAAAATAACGATTCTT
AHSV-9    TTAAAGATGGAGTTAGCAACAATGCGTCGA AAGAAATCGGCATTAAAAATAATGATTTTT
AHSV-1    CTGAAAGTTGAATTGAGCGGGATGCGAAGG AAGAAATTGATTTTGAAAATAATTATGTTT
AHSV-8    CTGAAAGTTGAATTGAGCGGGATGCGAAGG AAGAAATTGATTTTGAAAATAATTATGTTT
```

```
PALYAM  ATAATGAGCGTAATCGCAGTTGTAACATCA AGTTTAAGTGCGGCGATTGTTGTTATCCCA
AHSV-3  ATGAGCGGGTGCGTAACGATGGCTACCTCG ATGGCGGGCGGGTTAACGATTATTGATAAT
AHSV-4  ATCAGCGGATGTGTGACGTTAGCAACATCG ATGGTCGGGGGGTTAAGTATTGTCGATAAC
AHSV-9  ATTAGTGGGTGCGTGACGTTAGCTACATCG ATGGTTGGGGGATTGAGTATTGTTGATGAC
AHSV-1  ATTTGCGCAAACGTAACTATGGCTACTTCT CTAGTTGGAGGTATGTCAATCGTTGATGAG
AHSV-8  ATTTGCGCAAACGTAACTATGGCTACTTCT CTAGTTGGAGGTATGTCGATCGTTGATGAG

PALYAM  GAAACAAAGAATCTAAGAGAT--------- ------GAT---TGGATTAATATATTGGTA
AHSV-3  GAAATATATGAAGACCTTAGT--------- ---GGAGATGGTTGGCTGTCGAAGACGATT
AHSV-4  GAAATATTTGAAGATTATAAG--------- AAGAACGAT---TGGTTAATGAAAGCGATA
AHSV-9  CAAATATTAGATGATTATAAG--------- AAAAACGAT---TGGTTGATGAAGACTATA
AHSV-1  GATATTGCTAAGCATTTGGCGTTTGACGGA AAAGGGGAT---TGGGTGTCAAAAACGGTC
AHSV-8  GATATTGCTAAGCATTTGGCGTTTGACGGA AAAGGGGAT---TGGGTGTCAAAAACGGTC

PALYAM  CATGTAGGGAATTTATTTTCAACCGGAGTT TATGTTGGGATGAGCAAGTATGCGGAAAAG
AHSV-3  CACGGTTTGAATTTGCTGTGTACCACTATG TTGTTAGCGGCTGGAAAAATATCAGATAAA
AHSV-4  CATGGGCTGAATTTGTTATGTACCACAGTT TTGTTGGCGGCGGGTAAGATTTCTGATAAA
AHSV-9  CATGGGCTGAATTTGTTATGTACTACAGTT TTGTTAGCTGCGGGTAAGATTTCTGATAAA
AHSV-1  CATGGTTTAAATTTATTATGTACCACGATG CTGCTGGCAGCGAATAAAATATCGGAAAAG
AHSV-8  CATGGTTTAAATTTGTTATGTACCACAATG CTGCTAGCAGCGAATAAAATATCGGAAAAA

PALYAM  TTAGATAGTTTCCTTAAGAGAACACGAAAA GAGATAGTGAAGAAGGCGATCATATATAGN
AHSV-3  ATACAGGAGGAGATCTCACGCACAAAGCGG GATATAGCGAAGAGAGAATCATATGTTTCC
AHSV-4  ATACAAGAGGAGATTTCACGAACAAAGCGT GATATTGCGAAAAGAGAGTCTTACGTATCT
AHSV-9  ATACAAGAAGAAATTTCACGGACTAAACGT GACATTGCAAAAAGAGAGTCTTACGTATCA
AHSV-1  GTGAGAGAAGAGATTGCGAGGACAAAAAGA GACATCGCGAAAAGACAATCGTACGTATCA
AHSV-8  GTGAGAGAAGAGATTGCGAGGACAAAAAGA GACATCGCGAAAAGACAATCGTACGTATCA

PALYAM  NNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNN
AHSV-3  GCGGCTAGTATGTCTTGGAGTGGGGATACG AGCGTT---CTATTAAAAGAGGTAAAATAT
AHSV-4  GCGGCGAGTATGTCATGGAATGGAGATACT GAAGTA---TTATTGCAAGGAATTAAGTAT
AHSV-9  GCGGCGAGTATGTCGTGGAATGGAGATACT GAGATG---TTATTACAGGGAACTAAGTAT
AHSV-1  GCTGCGACGATGTCTTGGGATGGCGAT--- AGCGTAACTCAATTACGAGATGTAAAATAT
AHSV-8  GCTGCGACGATGTCTTGGGATGGCGAT--- AGCGTAACTCTATTACGAGATGTAAAATGT

                    691
                     |
PALYAM  NNNNNNNNNNNN  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
AHSV-3  GGCGACAGCTAG  --AAT-GACCTCCATTTGTGGGAGATCCACA-GTGTTGG-GTGGATC
AHSV-4  GGCGATAGCTAG  --TAC-GACCTCCACAAGCGGAAAATCCA-TCGTGTTGGA-TGGATG
AHSV-9  GGCGAAAGCTAG  --TAT-GACCTCCATGAGCGGAAAATCCA-TCGTGTTGGA-TGGATG
AHSV-1  GGAGAC---TAG  CGGATAGACCTCCAAAAGCGGGA-TTCCACTGGTGTTGCAGTGGAT-
AHSV-8  GGAGAC---TAG  CGGATAGACCTCCAAAAGCGGGA-TTCCGCTGGTGTTGCAGTGGAT-

                                                          782
                                                           |
PALYAM  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
AHSV-3  AAAC-ACCTAGATCGTTTTCTAGGGAGCCGGGATAACGACTTAC
AHSV-4  GAAC-GCCTAGATCGTTTTCTAGGGAGCCGGCATAACGACTTAC
AHSV-9  GAAC-GCCTAGATCGTTTTCTAGCNNNNNNNNNNNNNNNNNNN
AHSV-1  -AACCGCCTAGATCGTGTTCTAGGGAGCCGGGATAACGACTTAC
AHSV-8  -AACCGCCTAGATCGTGTTCTAGGGAGCCGGGATAACGACTTAC
```

Fig. 1 (continued).

into *Eco*RI/*Bam*HI digested pBluescript II KS vector (Stratagene), and the mixture used to transform competent HB101 cells.

Plasmids, with the appropriate sized insert, were sequenced by the dideoxy chain termination method (Sanger et al., 1977) using Sequenase Version 2.0 (U.S. Biochemicals). Purified PCR products were sequenced following a modified procedure (Winship, 1989). The 5′ and 3′ primers listed above, internal primers AAAAGCAGCGTTCGGTGCTATGGCGGAA (bases 213–240), ATTCCATGACATACTCGCGCAGATACGTAA (bases 620–590) specific for AHSV-4 and internal primers CCAAACTTAAAAACAC (bases 298–314), and CCGGATCTCTCAACGCC (bases 263–247) specific for AHSV-1, and forward and reverse universal primers which hybridize on either side of the polylinker region of the vector were used for sequencing. Both the PCR products and individual clones from each serotype were sequenced. The IBI-Pustell sequence analysis package (Kodak) was used for data entry and sequence analysis. Sequence data for serotypes 1, 4, and 8 has been deposited with GenBank under accession numbers U02711, U02712, and U02713, respectively.

Using RT-PCR we synthesized segment 10 from 8 of the 9 AHSV serotypes. We were unable to synthesize segment 10 from serotype 6 even when the annealing temperature during the PCR was reduced to 45°C (data not shown). The PCR products from serotypes 1, 4 and 8 were cloned and sequenced. Segment 10 from AHSV-4 is 758 nucleotides, while serotypes 1 and 8 segment 10 are 764 nucleotides (Fig. 1). These correspond to full-length segment 10 clones as indicated by the presence of 5′ and 3′ terminal consensus sequences (Fig. 1). No significant difference was found in the base composition among serotypes. As previously reported for segment 10 of AHSV-3 and 9 (Van Staden and Huismans, 1991) and for segment 10 of other orbiviruses (Hwang et al., 1992; Gould, 1988; Lee and Roy, 1986; Moss et al., 1992), segment 10 of AHSV-1, 4 and 8 possesses two putative initiation codons. The first initiation codon is located at nucleotide positions 20–22 in all AHSV serotypes so far analyzed, except AHSV-9 in which it is at nucleotide 19–21 as in Palyam virus. The second initiation codon is slightly more variable, being at position 50–52 in AHSV-1, 3, and 8, at nucleotide 52–54 in AHSV-9 and Palyam virus, and at 53–55 in AHSV-4. Neither AUG is in the optimal context for initiation of translation, but because of its position relative to the 5′ end of the mRNA the first codon is the major initiation site in eucaryotic cells (Kozak, 1991).

As with other orbivirus segment 10 genes, there is a relatively long 3′ untranslated region of 88 bases for serotypes 3 and 4 and 91 bases for serotypes 1 and 8 (Hwang et al., 1992; Moss et al., 1992; Roy, 1989; Van Staden and Huismans, 1991). There is also considerable sequence conservation at the 3′ end, with only 2 differences extending 17 bases upstream of the primer sequence.

Amino acid sequence for each serotype was deduced from the nucleotide sequence. The coding region was aligned using the amino acid sequences with the alignment subroutines of the IBI/Pustell sequence analysis software (Pustell and Kafatos, 1986) (Fig. 2). Gaps (interpreted as insertions or deletions in one or more taxa) were introduced manually in the sequences to increase their aligned similarity. The amino acid alignment was then reverse translated into the nucleotide

```
PALYAM  M-LARSLNEY KAMRSESEMS VVPYQPPAY- PTAPVGLRSE -KDLSGISLG
AHSV-3  MSLATIAENY -MMHNGNQRA IVPYVPPPYA YANAPTLGGQ AGEMESMSLG
AHSV-4  MNLAAIAKNY -SMHNGESGA IVPYVPPPYN FASAPTFSQR TSQMESVSLG
AHSV-9  MNLAAIAENY -SMHNGESGA IVPYVPPPYN FASAPTFSQR TSQMESVSLG
AHSV-1  MNLASISQSY -MSHNENERS IVPYIPPPY- HPTAPALAVS ASQMETMSLG
AHSV-8  MNLASISQSY -MSHNENERS IVPYIPPPY- HPTAPALAVS ASQMETMSLG

PALYAM  VLNNAMNDTT AATQAEREEK VAYASFAEAL RDPMCVREIK KRVSSRTIVA
AHSV-3  ILNQAMSSTT GASRALKDEK AAFGAMAEAL RDPEPIRQIK KHVGLRTLKH
AHSV-4  ILNQAMSSTT GASGALKDEK AAFGAMAEAL RDPEPLRQIK KQVGIRTLKN
AHSV-9  ILNQAMSSTT GASGALKDEK AAFGAMAEAL RDPEPLRQIK KQVGIRTLKN
AHSV-1  ILNQAMSSSA GASGALKDEK AAYGAVAEAL RDPEPIRKIK RQVGIQTLKT
AHSV-8  ILNQAMSSSA GASGALKDEK AAFGAVAEAL RDPEPIRKIK RQVGIQTLKT

PALYAM  LEKEYRHQKR IYDFVRLILF IMSVIAVVTS SLSAAIVVIP ETKNLRD---
AHSV-3  LKIELASMRR RYAILRVVIF MSGCVTMATS MAGGLTIIDN EIYEDLS---
AHSV-4  LKMELATMRR KKSALKITIL ISGCVTLATS MVGGLSIVDN EIFEDYK---
AHSV-9  LKMELATMRR KKSALKIMIF ISGCVTLATS MVGGLSIVDD QILDDYK---
AHSV-1  LKVELSGMRR KKLILLIIMF ICANVTMATS LVGGMSIVDE DIAKHLAFDG
AHSV-8  LKVELSGMRR KKLILLIIMF ICANVTMATS LVGGMSIVDE DIAKHLAFDG

PALYAM  --D-WINILV HVGNLFSTGV YVGMSKYAEK LDSFLKRTRK EIVKKGIIY*
AHSV-3  -GDGWLSKTI HGLNLLCTTM LLAAGKISDK IQEEISRTKR DIAKRESYVS
AHSV-4  KND-WLMKAI HGLNLLCTTV LLAAGKISDK IQEEISRTKR DIAKRESYVS
AHSV-9  KND-WLMKTI HGLNLLCTTV LLAAGKISDK IQEEISRTKR DIAKRESYVS
AHSV-1  KGD-WVSKTV HGLNLLCTTM LLAANKISEK VREEIARTKR DIAKRQSYVS
AHSV-8  KGD-WVSKTV HGLNLLCTTM LLAANKISEK VREEIARTKR DIAKRQSYVS

PALYAM  ********** *************
AHSV-3  AASMSWSGDT SV-LLKEVKYGDS
AHSV-4  AASMSWNGDT EV-LLQGIKYGDS
AHSV-9  AASMSWNGDT EM-LLQGTKYGES
AHSV-1  AATMSWDGD- SVTQLRDVKYGD-
AHSV-8  AATMSWDGD- SVTLLRDVKCGD-
```

Fig. 2. Comparison of the predicted amino acid sequence of AHSV and Palyam segment 10 gene products. The amino acid sequence of segment 10 from AHSV serotypes 1, 3, 4, 8, and 9 and Palyam were determined from the nucleic acid sequence. A dash indicates a gap inserted to obtain maximum alignment. An asterisk indicates that the amino acid sequence in this region has not been obtained.

sequences respecting the gaps introduced during the protein alignment. When gaps were introduced, they were treated in two different ways. First, gaps were rendered uninformative by setting the gap symbol equal to missing data, and secondly each gap was treated as a single event (independently of its length) and coded separately as present or absent (Swofford, 1993). The Palyam sequence was used as an outgroup and characters were run unordered.

The analysis of the NS3 protein coding region (673 nucleotide positions) provided the shortest tree (length = 379, C.I. = 0.868), indicating a close relationship between AHSV-4 and AHSV-9, and between AHSV-1 and AHSV-8; AHSV-3 is placed closer to AHSV-4/AHSV-9 than to the AHSV-1/AHSV-8 group (Fig. 3). All branches were significantly supported (100%) on the basis of their presence in all 100 bootstrap iterations. The analyses of the coding region when gaps were treated as independent events, as well as when transversion substitutions were weighted over transition substitutions resulted in the same tree. The 5′ non-coding

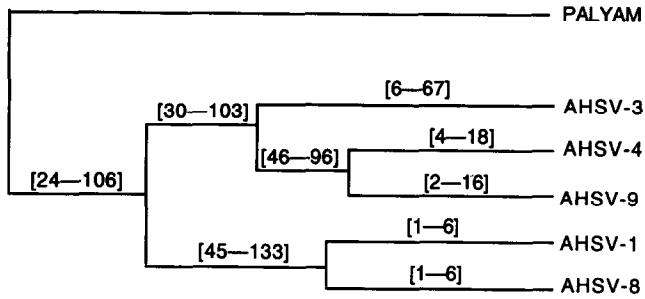

Fig. 3. Evolutionary relationships among AHSV serotypes. Tree obtained from PAUP analysis of segment 10 sequences ignoring uninformative characters. Tree length = 379, consistency index = 0.868, values above branches correspond to branch lengths under all character optimization.

region was not analyzed because its entire length consisted of the primer sequence used for cloning. An analysis performed using the 3′ non-coding region, with the exclusion of the 3′ primer sequence, agreed with the previous analyses. Finally, analyses were performed combining both data sets. Each of these analyses resulted in the shortest tree that supports identical relationships among the serotypes.

Using the first AUG codon, segment 10 from AHSV-3, 4 and 9 codes for a protein of 217 amino acids, while AHSV-1 and 8 code for a 218 amino acid protein (Fig. 2). All five proteins have a similar composition of non-polar, polar, acidic and basic amino acid residues. The AHSV-8 protein contains 3 cysteine residues, while the other 4 serotypes contain 2 cysteine residues similar to the equivalent BTV protein. The cysteine at residue 162 is conserved in all 5 serotypes, while serotypes 1 and 8 share a cysteine at residue 120 and serotypes 3, 4 and 9 share a cysteine at residue 122. Serotypes 4 and 9 segment 10 gene products are 95% identical, 1 and 8 are 99% identical, serotype 3 is 75–76% identical to 4 and 9, but only 65% identical to serotypes 1 and 8. Although serotype 1 and 8 segment 10 gene products have only 65–66% identity to the segment 10 gene product from serotypes 3, 4 and 9, they have very similar hydrophobic profiles (data not shown). Also, as previously reported by Van Staden and Huismans (1991) for serotypes 3 and 9, there are regions of very high sequence similarity for all five serotypes (Fig. 2).

Variation in segment 10 among AHSV serotypes is greater than the variation exhibited by segment 10 of BTV (Hwang et al., 1992), equine encephalosis virus (Viljoen and Huismans, 1989), and Palyam virus (Bodkin and Knudson, 1985). The latter two studies utilized either partial cDNA clones of segment 10 or segment 10 dsRNA as hybridization probes. The six segment 10 genes of BTV that have been sequenced can be aligned without inclusion of gaps (Hwang et al., 1992). Full-length clones of segment 10 from 4 AHSV serotypes vary from 758 to 764 nucleotides, whereas full-length copies of all known BTV serotypes are 822 nucleotides in length (Hwang et al., 1992). The inter-serogroup size variation increases the alignment ambiguity and we have not yet been able to satisfactorily align segment 10 of BTV with AHSV.

AHSV segment 10 codes for 2 non-structural proteins NS3/NS3a (Van Staden

and Huismans, 1991). These proteins as yet have no known function. Immunoelectron microscopic studies by Hyatt et al. (1991) have shown that BTV NS3/NS3a are associated with intracellular smooth-surfaced vesicles and the plasma membrane, suggesting that they may be involved in the final stages of BTV morphogenesis. Wu et al. (1992) have recently shown that BTV NS3/NS3a are glycosylated and are probably integral membrane proteins, although N-linked glycosylation is not required for their intracellular transport. BTV NS3/NS3a has two potential glycosylation sites that are conserved in all six segment 10 serotypes sequenced (Hwang et al., 1992). AHSV-1 and 8 NS3/NS3a have a potential glycosylation site at residues 122–124 (NVT), AHSV-4 and 9 share a potential site at residues 9–11 (NYS), while AHSV-3 NS3/NS3a contain no potential N-linked glycosylation site.

The variation in the segment 10 gene of AHSV is surprising, since the cognate gene of BTV is highly conserved among the six serotypes sequenced and the BTV gene product has been implicated in a function that one might expect to be conserved. However, comparison of AHSV NS3/NS3a proteins reveals regions of high conservation, even among the more distantly related groups. Furthermore, a comparison of the hydrophobic profiles of AHSV and BTV NS3/NS3a shows a very similar pattern, including the conservation of 2 hydrophobic domains. It will be of interest to determine if these conserved regions are involved in the function NS3/NS3a has in the virus' replication cycle.

Acknowledgements

We thank Drs. Carol House and James House for seed stocks of the viruses used in this study and an anonymous reviewer for meticulous and helpful comments.

References

Bodkin, D.K. and Knudson, D.L. (1985) Sequence relatedness of Palyam virus genes to cognates of the Palyam serogroup viruses by RNA-RNA blot hybridization. Virology 143, 55–62.

Bremer, C.W., Huismans, H. and Van Dijk, A.A. (1990) Characterization and cloning of the African horsesickness virus genome. J. Gen. Virol. 71, 793–799.

Gorman, B.M. and Taylor, J. (1985) Orbiviruses. In: B.N. Fields et al. (Eds), Virology pp. 907–925. Raven Press, New York.

Gould, A.R. (1988) Nucleotide sequence of the Australian bluetongue virus serotype 1 RNA segment 10. J. Gen. Virol. 69, 945–949.

Grubman, M.J. and Lewis, S.A. (1992) Identification and characterization of the structural and non-structural proteins of African horsesickness virus and determination of the genome coding assignments. Virology 186, 444–451.

Howell, P.G. (1962) The isolation and identification of further antigenic types of African horsesickness virus. Onderstepoort J. Vet. Res. 29, 139–149.

Hwang, G.-Y, Yang, Y.-Y, Chiou, J.-F. and Li, J.K.K. (1992) Sequence conservation among the cognate non-structural NS3/3A protein genes of six bluetongue viruses. Virus Res. 23, 151–161.

Hyatt, A.D., Gould, A.R., Coupar, B. and Eaton, B.T. (1991) Localization of the non-structural protein NS3 in bluetongue virus-infected cells. J. Gen. Virol. 72, 2263–2267.

Kozak, M. (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. 266, 19867–19870.

Lee, J.W. and Roy, P. (1986) Nucleotide sequence of a cDNA clone of RNA segment 10 of bluetongue virus (serotype 10). J. Gen. Virol. 67, 2833–2837.

Mizukoshi, N., Sakamota, K., Iwata, A., Tsuchiya, T., Ueda, S., Watanabe, T., Kamada, M. and Fukusho, A. (1992) The complete sequence of African horsesickness virus serotype 4 (vaccine strain) RNA segment 5 and its predicted polypeptide compared with NS1 of bluetongue virus. J. Gen. Virol. 73, 2425–2428.

Moss, S.R., Jones, L.D. and Nuttall, P.A. (1992) Comparison of the nonstructural protein, NS3, of tick-borne and insect-borne orbiviruses. Virology 187, 841–844.

Oellermann, R.A., Els, H.J. and Erasmus, B.J. (1970) Characterization of African horsesickness virus. Arch. Gesam. Virusforsch. 29, 163–174.

Pustell, J. and Kafatos, F.C. (1986) A convenient and adaptable microcomputer environment for DNA and protein manipulation and analysis. Nucleic Acids Res. 14, 479–488.

Roy, P. (1989) Bluetongue virus genetics and genome structure. Virus Res. 13, 179–206.

Roy, P., Hirasawa, T., Fernandez, M., Blinov, V.M. and Sanchez-Vixcain Rodrique, J.M. (1991) The complete sequence of the group-specific antigen VP7, of African horsesickness disease virus serotype 4 reveals a close relationship to bluetongue virus. J. Gen. Virol. 72, 1237–1241.

Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Swofford, D.L. (1993) PAUP: Phylogenetic analysis using parsimony, Version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign, IL.

Van Staden, V. and Huismans, H. (1991) A comparison of the genes which encode non-structural protein NS3 of different orbiviruses. J. Gen. Virol. 72, 1073–1079.

Van Staden, V. Theron, J., Greyling, B.J., Huismans, H. and Nel, L.H. (1991). A comparison of the nucleotide sequence of cognate NS2 genes of three different orbiviruses. Virology 185, 500–504.

Viljoen, G.J. and Huismans, H. (1989) The characterization of equine encephalosis virus and the development of genomic probes. J. Gen. Virol. 70, 2007–2015.

Winship, P.R. (1989) An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. Nucleic Acids Res. 17, 1266.

Wu, X., Chen, S.-Y., Iwata, H., Compans, R.W. and Roy, P. (1992) Multiple glycoproteins synthesized by the smallest RNA segment (S10) of bluetongue virus. J. Virol. 66, 7104–7112.