Molecular epidemiology of the African horse sickness virus S10 gene

Melvyn Quan,1,2 Moritz van Vuuren,2 Peter G. Howell,2 Daleen Groenewald2 and Alan J. Guthrie1

1Equine Research Centre, Faculty of Veterinary Science, University of Pretoria, Private Bag X04, Onderstepoort 0110, South Africa
2Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, Private Bag X04, Onderstepoort 0110, South Africa

Between 2004 and 2006, 145 African horse sickness viruses (AHSV) were isolated from blood and organ samples submitted from South Africa to the Faculty of Veterinary Science, University of Pretoria. All nine serotypes were represented, with a range of 3–60 isolates per serotype. The RNA small segment 10 (S10) nucleotide sequences of these isolates were determined and the phylogeny investigated. AHSV, bluetongue virus (BTV) and equine encephalosis virus (EEV) all formed monophyletic groups and BTV was genetically closer to AHSV than EEV. This study confirmed the presence of three distinct S10 phylogenetic clades (a, b and c). Some serotypes (6, 8 and 9 in a; 3 and 7 in b; 2 in c) were restricted to a single clade, while other serotypes (1, 4 and 5) clustered into both the a and c clades. Strong purifying selection was evident and a constant molecular clock was inappropriate. The S10 gene is the second most variable gene of the AHSV genome and the use of S10 in molecular epidemiology was illustrated by an AHS outbreak in the Western Cape in 2004. It was shown that two separate AHSV were circulating in the area, even though AHSV serotype 1 was the only isolate from the outbreak. The small size of the gene (755–764 bp) and conserved terminal regions facilitate easy and quick sequencing. The establishment of an S10 sequence database is important for characterizing outbreaks of AHS. It will be an essential resource for elucidating the epidemiology of AHS.

INTRODUCTION

African horse sickness virus (AHSV) is the aetiological agent of African horse sickness (AHS), an Office International des Epizooties (OIE) listed disease. AHS is non-contagious but nevertheless it is a serious disease of equids that may cause up to 95% mortality in naïve horses. The disease is endemic to sub-Saharan Africa (Coetzer & Guthrie, 2004) and is transmitted by midges (Culicoides spp.) (DuToit, 1944) under favourable climatic conditions. The extension of the insect’s range due to climate change makes the international spread of AHS a very real possibility (Purse et al., 2005).

AHSV is a member of the genus Orbivirus, the family Reoviridae of which Bluetongue virus (BTV) is the type species. The virus consists of 10 segments of double-stranded RNA that encodes seven structural proteins (VP1–VP7) and four non-structural proteins (NS1, NS2, NS3 and NS3a). The genome is surrounded by an inner capsid, which forms the virus core and has two distinct layers of proteins. The inner layer or ‘subcore’ is composed of VP3 and associated with it are three minor structural proteins (the core-associated enzymes VP1, VP4 and VP6). The outer surface of the core particle is composed of VP7. Surrounding the inner capsid is the outer capsid composed of VP2 and VP5 proteins. VP2 contains the major serotype-specific, neutralizing antigens and is primarily involved in cell attachment and penetration (Stone-Marschat et al., 1996; Burrage et al., 1993). The RNA small segment 10 (S10) gene is the smallest segment of the AHSV genome, between 755 and 764 bp long. It contains two overlapping in-phase open reading frames that encode two non-structural proteins, NS3 and NS3a (van Staden & Huismans, 1991; Mertens et al., 1984). NS3 is larger than NS3a as a result of an additional 10 aa at the N-terminal end. NS3 is an integral membrane protein (Bansal et al., 1998; Wu et al., 1992; Hyatt et al., 1991) and it has been suggested that the protein is involved in the final stages of viral morphogenesis and release of virions (Wirblich et al., 2006; Martin et al., 1998; Stoltz et al., 1996; Hyatt et al., 1993). It is cytotoxic to in vitro insect cells (van Niekerk et al., 2001a; van Staden et al., 1995) and it may
play a role in the determination of virulence (O’Hara et al., 1998).

NS3 is the second most variable AHSV protein after the outer capsid protein VP2, and it can vary 37 % across serotypes and 28 % within serotypes (van Niekerk et al., 2001b). The AHSV NS3 is not as well conserved as the BTV NS3, which varies by only 7–10 % (van Niekerk et al., 2001b) and the equine encephalitis virus (EEV) NS3, which varies by 17 % (van Niekerk et al., 2003). Conserved regions within AHSV NS3 have been reported to be (i) the initiation codon for NS3a, (ii) a proline-rich area between residues 22 and 34, (iii) a sequence of highly conserved amino acids from residues 46 to 90 and (iv) two hydrophobic regions (residues 116–137 and 154–170, respectively) predicted to form transmembrane helices (van Staden et al., 1998).

Between 2004 and 2006, 145 AHSV were isolated from blood and organ samples submitted from South Africa, Namibia and Zimbabwe to the Department of Veterinary Tropical Diseases (DVTD), Faculty of Veterinary Science, University of Pretoria. The S10 nucleotide sequences of these isolates were determined and the phylogeny of S10 was investigated, first by comparing AHSV S10 to other orbivirus S10 genes and then by determining the relationship of all AHSV S10 sequences to each other, the three distinct phylogenetic clades (α, β and γ) of AHSV S10 described previously (van Niekerk et al., 2003; Martin et al., 1998; Saillieu et al., 1997), the established nine AHSV serotypes (Coetzter & Guthrie, 2004) and the geographical location where they were isolated. The phylogeny of AHSV S10 has never been described using the number of isolates indicated in this paper.

Another novel aspect of this study was the use of the S10 data to characterize AHS outbreaks. The high sequence variation in S10 makes the identification of various AHSV variants possible. An AHS outbreak in 2004 in the Western Cape, a region usually free of the disease, was examined, as good epidemiological data and field samples were available (Sinclair, 2006). Linking epidemiological and sequence data are described and we examine whether the molecular data provide additional information to the epidemiology of the outbreak.

METHODS

Virus isolation and identification. Virus was recovered from heparinized blood, spleen or lung samples submitted to the DVTD. Blood in heparin was centrifuged at 440 g for 10–15 min. The leukocytes (buffy coat) were harvested, dispensed into two sterile cryotubes (Nunc) and frozen at −80 °C for future reference and the remainder was filtered through a 0.22 μm membrane filter (Microsep).

Processed blood (0.2 ml) or 0.5 ml organ sample filtrate was inoculated onto a confluent monolayer of BHK-21 cells in 25 cm² tissue culture flasks. Cell cultures were incubated at 37 °C and observed daily for any cytopathic effects (CPE). Blood cells were washed off after 3 days and the medium replaced. Cultures showing no CPE after 10–14 days were passaged by inoculating 0.2–0.5 ml of the culture onto a freshly prepared BHK-21 cell monolayer. When 100 % of the cell monolayer showed CPE or had detached from the surface due to degeneration, the cells and supernatant were harvested and identified as AHSV by competitive ELISA (Hamblin et al., 1990). Cultures showing no CPE after three passages were classified as negative for AHSV.

Plaque inhibition neutralization test. The serotype of AHSV recovered from the sample was determined by a modified plaque inhibition test, after the technique described by Porterfield (1960), in which electrical insulating fish-spine beads filled with type-specific antiserum were used to indicate homologous virus-antibody neutralization. A set of three confluent monolayers of Vero cells prepared in 35 mm diameter cluster plates was inoculated with a dilution of the virus sample. After adsorption for 1 h at 37 °C in a 5 % CO₂ gassed and humidified incubator, the inoculum was removed and replaced with 5 ml of an agarose overlay containing 1:20 000 neutral red. The fish-spine beads, containing the type-specific antisera, were placed in strict sequence on the surface of the overlay, after which the plates were returned to the incubator. Homologous neutralization was clearly evident around the periphery of a bead by the absence of plaques after 5–6 days incubation.

RNA extractions. The contents of a tissue culture flask (≤ three passages) that showed an AHSV-positive ELISA were agitated and 1 ml transferred to a 1.5 ml Eppendorf tube. Samples were spun at 3000 r.p.m. for 10 min in a 5417C centrifuge (Eppendorf) and the supernatant was discarded. RNA extractions from the cell pellet were performed with either TRizol (Invitrogen) or the RNeasy Mini kit (Qiagen) according to the manufacturer’s instructions. RNA was dissolved in either 50 μl TE buffer (10 mM Tris, 1 mM EDTA, pH 7:0; Ambion) or 50 μl nuclease-free H₂O (Ambion) and stored at −80 °C. To denature the RNA, 1 μl 0.2 M methymercury (II) hydroxide (Alfa Aesar) was added to every 5 μl of thawed RNA. The mixture was incubated for 10 min at room temperature before reduction with 1 μl 1 M 2-mercaptoethanol (Sigma) per 5 μl RNA.

RT-PCR. A one-step RT-PCR was performed with the GeneAmp Gold RNA PCR core kit (Applied Biosystems) according to the manufacturer’s instructions. Two 25 μl reactions were performed per sample, containing 2 μl denatured RNA, 0.5 μl 20 nM MQ.AHS.NS3.1-22F (5’-GTTTTAAATTATCCCTGTCTAG-3’) and 0.5 μl 20 nM MQ.AHS.NS3.749-769R (5’-GTAAGTCGTTATC-CGGGCCTC-3’) terminal primers per reaction. Cycling conditions for RT-PCR were 42 °C for 12 min, 95 °C for 10 min; 40 cycles of 94 °C for 20 s, 57 °C for 30 s, 72 °C for 30 s and 72 °C for 7 min, and a hold at 4 °C on a GeneAmp PCR System 9700 (Applied Biosystems). Five microlitres PCR product mixed with 1 μl loading dye (Fermentas) was visualized by UV trans-illumination of a 1.5 % agarose gel electrophoresed at 110 V.

The two RT-PCR reactions per sample were combined and the mixture cleaned up using a QiAquick PCR purification kit and a QiAprep 65 vacuum manifold (Qiagen) according to the manufacturer’s instructions. Samples were eluted in 60 μl EB buffer (Qiagen).

Sequencing. A BigDye Terminator v3.1 cycle sequencing kit was used for sequencing. At least two forward and two 1/2 (2 μl) reverse reactions (4–7 μl DNA, total volume per reaction=20 μl) were
performed per sample according to the manufacturer’s instructions. Cycling conditions and clean-up of sequencing products were as described in the Applied Biosystems protocol. Samples were analysed with an ABI 3130xl Genetic Analyzer (Applied Biosystems).

Phylogenetic analyses

Orbivirus phylogeny. Orbivirus S10 sequence data were obtained from the GenBank website (http://www.ncbi.nlm.nih.gov/sites/ entrez?db=PubMed). Protein sequences were aligned with MAFFT (v5.8) (Katoh et al., 2002, 2005) using the L-INS-i and MAFFT homologues options, then back-translated into nucleotide sequences by the translalign programme in EMBoss (Rice et al., 2000).

A general time reversible model with a proportion of invariable sites and a gamma-shaped distribution of rates across sites (GTR+I+Γ) substitution model was determined by Modeltest v3.7 (Posada & Crandall, 1998) and MrModeltest2 (Nylander, 2004). PAUP* v4b10 (Swofford, 2003) was used to explore distance, parsimony (using a heuristic search and tree bisection reconnection) and maximum-likelihood phylogenetic methods for both nucleotide and amino acid sequences; MRBAYES v3.1.1 (Ronquist & Huelsenbeck, 2003) was used to explore Bayesian substitution model was determined by Modeltest v3.7) and PAUP* v4b10.

Saturation was evaluated by plotting transitions and transversions against K80 distance using DAMBE (Xia & Xie, 2001). The likelihood ratio test (Felsenstein, 1981) was used to test the molecular clock using values obtained from Modeltest v3.7. The relative abundance of synonymous and non-synonymous substitutions within the S10 gene was calculated with the Nei–Gojobori method (Nei & Gojobori, 1986) in MEGA v4, and the Z-test used to test for neutrality.

S10 phylogeny. Phred (Ewing et al., 1998; Ewing & Green, 1998) and the Staden package (Staden et al., 2000; Staden, 1996) were used for base calling and sequence assembly. A sequence alignment was done with MAFFT using the G-INS-i option.

A cladogram was generated with MRBAYES v3.1.1 using a GTR+Γ substitution model (as determined by MrModeltest2) and three heated chains. The consensus tree was constructed in MEGA v4. Genetic distances were calculated between 56 AHSV isolates with unique S10 sequences, using a transversional model and a gamma-shaped distribution of rates across sites (TVM+Γ) substitution model (as determined using Modeltest v3.7) and PAUP* v4b10.

Conserved regions, hydrophilicity and transmembrane domains of the AHSV NS3 gene were determined using the plotcon, pepwindowall programmes within EMBoss.

RESULTS

Orbivirus S10 phylogeny

The relationship of AHSV with other orbivirus S10 genes was explored by phylogenetic analysis. Although a good phylogenetic signal was present ($g_1=−0.59$, $P<0.01$), the phylogeny of the S10 gene of the genus Orbivirus could not be resolved with certainty due to conflicting results and weak bootstrap values from the different phylogenetic analyses. A plot of transitions and transversions against the K80 distance between sequences showed a high degree of saturation, especially of the third codon position (results not shown).

A maximum-parsimony cladogram based on amino acid sequences and a Bayesian phylogram based on nucleotide sequences are shown in Fig. 1. Although there were differences between the trees, areas of agreement existed. AHSV formed a monophyletic group separate from other orbiviruses. The closest relation to AHSV appeared to be a group including BTV, although the support for this association was weak (bootstrap value of 58%). The latter group was made up of BTV, which was closely related to epizootic hemorrhagic disease virus (EHDV), Ibaraki virus (IBAV) and only weakly linked to Chuzan virus (CHUV) of the Palyam group (in the protein analysis, CHUV was not part of this group). Peruvian horse virus (PHV) was closely related to Yunnan orbivirus, while the closest relation to Broadhaven virus (part of the Great Island virus group) was St Croix River virus. EEV formed a separate clade from other orbiviruses.

No conclusions regarding the evolution of the S10 gene could be made. A likelihood ratio test showed that rates of substitution varied amongst the branches of the tree and that a constant molecular clock was inappropriate ($\chi^2=1301.13$, d.f. $=40$, $P<0.001$).

A codon-based Z-test that tested the relative abundance of synonymous and non-synonymous substitutions within the orbivirus S10 gene indicated that strong purifying selection was operating on this gene ($P<0.001$).

AHSV S10 phylogeny

One hundred and forty-five AHSV were isolated from samples submitted to the Faculty of Veterinary Science, University of Pretoria between 2004 and 2006 (Fig. 2). All nine serotypes were represented, with a range of 3–60 isolates (and 2–15 isolates with unique sequences) per serotype. Most of the samples were submitted from the Gauteng Province (GP, 53 %), followed by the KwaZulu-Natal Province (KZ, 18 %) and then the Western Cape Province (WP, 14 %) but they did not represent the prevalence of disease. The majority of AHSV serotypes submitted from GP were either serotype 7 (42 %) or serotype 5 (22 %). Outbreaks of serotype 1 occurred in the WP in 2004 and serotype 7 in KZ in 2005, which account for the majority of serotypes submitted from those provinces.

Sequencing of the virus tissue culture isolates confirmed the presence of three distinct S10 phylogenetic clades ($\alpha$, $\beta$ and $\gamma$) (Fig. 3), as previously reported (van Niekerk et al., 2003; Martin et al., 1998; Sailleau et al., 1997). Some serotypes (6, 8 and 9 in $\alpha$; 3 and 7 in $\beta$; 2 in $\gamma$) were restricted to a single clade, while other serotypes (1, 4 and 5) clustered into both the $\alpha$ and $\gamma$ clades.

The genetic distances between unique AHSV isolates are summarized in Fig. 4. The mean S10 distance within a serotype ($\bar{x}=0.36$) was significantly less than between heterologous serotypes ($\bar{x}=0.66$, $P<0.001$). The S10 genes of serotypes 5, 4 and 1 showed the greatest distance within
a serotype (x̄=0.60, 0.42 and 0.33, respectively), while the least distance was within serotype 7 (x̄=0.01). Between serotypes, the greatest distance was between serotypes 2 and 3 (x̄=1.19), 1 (x̄=1.11) and 3 (x̄=1.11). The least distance was between serotypes 3 and 7 (x̄=0.06), 6 and 8 (x̄=0.10) and 8 and 9 (x̄=0.14).

**Fig. 1.** Orbivirus S10 phylogeny of (a) protein (rooted 50% majority-rule consensus cladogram calculated by maximum-parsimony and bootstrap estimates from 1000 replicates, tree length=1441, consistency index=0.79 and retention index=0.89) and (b) nucleotide (rooted phylogram calculated by Bayesian inference, GTR+I+Γ substitution model and 5×10⁶ iterations) sequences. AHSV, African horse sickness virus; BRDV, Broadhaven virus (Great Island virus group); BTV, bluetongue virus; CHUV, Chuzan virus (Palyam virus group); EEV, equine encephalosis virus; EHDV, epizootic hemorrhagic disease virus; IBAV, Ibaraki virus; PHV, Peruvian horse virus; RV-A, rotavirus A; SCRV, St Croix River virus; YUOV, Yunnan orbivirus. Virus names are followed by the serotype (if applicable), gene and GenBank accession number. Bootstrap values and posterior probabilities are shown on the nodes of the tree. The NSP4 gene of RV-A (family Reoviridae) was used as an outgroup.

**Fig. 2.** Classification of AHSV isolates collected between 2004 and 2006 by province and serotype. Pie chart in key is included for size comparison.
Fig. 3. An unrooted cladogram of AHSV S10 sequences from isolates obtained between 2004 and 2006, determined with Bayesian inference and 1 x 10^6 iterations. Three distinct clades (\(\alpha\), \(\beta\) and \(\gamma\)), supported by high posterior probabilities (indicated on cladogram), are evident. Isolate names reflect serotype (x, not classified), year of isolation and province/country (EC, Eastern Cape; GP, Gauteng; KZ, KwaZulu-Natal; LP, Limpopo; MP, Mpumalanga; NA, Namibia; NW, North-West; WP, Western Cape; ZM, Zimbabwe) where isolated. Jane isolates refer to reference virulent field strains. Pairwise comparisons of isolates marked with similar symbols (either \(\Delta\), \(\triangledown\) or \(\times\)) by a codon-based Z-test showed a higher abundance of non-synonymous to synonymous changes (\(P<0.05\)), indicating positive selection. Underlined isolates refer to samples collected during the AHS outbreak in the Western Cape Province in 2004.

http://vir.sgmjournals.org 1163
A likelihood ratio test showed that rates of substitution varied amongst the branches of the tree and that a constant molecular clock was inappropriate ($\chi^2=280.84$, d.f. = 152, $P<0.001$). Overall, a strong purifying selection was operating on the S10 gene ($P<0.001$) but a pairwise comparison identified pairs of isolates for which positive selection was evident ($P<0.05$) (Fig. 3). Amino acid differences between these isolates are indicated in Fig. 5.

**AHSV NS3 protein**

Analysis of all the AHSV NS3 sequences showed a well conserved proline-rich region between residues 20 and 28, as described by van Staden et al. (1995). This region was surrounded by hydrophilic, variable residues under positive-selective pressure (Fig. 5) and was followed by a well conserved region from residues 47 to 104. A transmembrane region between residues 108 and 135 (consensus sequence of RRKAILIVIFMSGCVTMATSMVGGLT) was identified with tmap and was strongly hydrophobic on a hydropathy plot. The plot identified another hydrophobic region, between residues 160 and 174. This compares favourably with residues 116 to 137 and 154 to 170, reported by van Staden et al. (1995) to be transmembrane helices. Between the two transmembrane regions was a highly variable, hydrophilic region from residues 139 to 158, under positive-selective pressure. The region from residues 159 to 200 was well conserved, while the last 20 residues of NS3 were variable and showed evidence of positive selection.

**AHS outbreak, Western Cape, 2004**

The index case was a horse at Elsenburg Agricultural Research Farm that died on the 31 January 2004 (Fig. 6 and Table 1). No viral isolations were performed from the first two horses that died in the outbreak, as annual rye grass toxicity was diagnosed and AHS was not suspected. Three horses died subsequently (22–24 February 2004) and one horse was clinically ill from AHS (26 February 2004). AHSV serotype 1 was isolated from all the samples submitted. S10 characterization showed that all these isolates belonged to the $c$ clade (underlined isolates, Fig. 3).
After the deaths at Elsenburg, horses died on the Troughend, Oakhill and Daktari (26–28 February 2004) farms, about 15 km south of Elsenburg. AHSV serotype 1 was again isolated from these samples and a point introduction for the outbreak was suspected. Sequencing, however, identified these isolates as belonging to the α clade and not the γ clade. The α clade isolates were only 65% similar to the γ clade and introduction of AHS to these farms was, therefore, from a separate source compared with that from Elsenburg. AHSV isolated from the remainder of the outbreak samples all belonged to the α clade. No difference in virulence between the twoviruses was observed. The outbreak resulted in the death of 17 horses, with the last death on the 28 March 2004 at Kalbaskraal. S10 characterization showed that in this outbreak there were at least two separate introductions of AHSV to the Western Cape.

Isolates from the α clade showed sequence variation within the S10 gene, which may have been useful for tracking the spread of the virus and determining a chain of infection if these mutations were acquired during the course of the outbreak. The data showed that this was not the case and that a set of quasispecies was present from the beginning of the outbreak. Quasispecies can be described as a population of viruses without a defined unique sequence structure, but rather as a weighted average of a large number of different individual sequences (Biebricher & Eigen, 2006; Domingo et al., 1978). Quasispecies αi, αii and αiii were all present at Troughend, Oakhill and Daktari farms at the beginning of the α clade outbreak and within 3 days of each other. In addition, αii and αiii isolates were obtained from the same horse (Bow Street Belle, Table 1), indicating quasispecies in one horse. Rather than a sequential acquisition of mutations during the outbreak, a genetically diverse population of viruses was present from the start of the outbreak and therefore no inference could be made about a chain of infection using the genetic data.

The source of the serotype 1 AHSV strains responsible for the Western Cape outbreak could not be determined. The

![Fig. 6. Phylogenetic analysis of AHSV serotype 1 suggests two separate introductions of the virus (α and γ clade) into the Western Cape Province in 2004. Quasispecies within the α clade indicated with subscript (αi, αii and αiii).](http://vir.sgmjournals.org)

### Table 1. Isolates from the AHS outbreak in the Western Cape Province in 2004 from which the S10 gene was sequenced and typed

<table>
<thead>
<tr>
<th>Date</th>
<th>Locality</th>
<th>Horse name</th>
<th>Case no.</th>
<th>S10 clade</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 February 2004</td>
<td>Elsenburg Agricultural Research</td>
<td>Michelle</td>
<td>E02004</td>
<td>γ</td>
</tr>
<tr>
<td>22 February 2004</td>
<td>Elsenburg Agricultural Research</td>
<td>Laura</td>
<td>E01904</td>
<td>γ</td>
</tr>
<tr>
<td>24 February 2004</td>
<td>Elsenburg Agricultural Research</td>
<td>Mara</td>
<td>E02104</td>
<td>γ</td>
</tr>
<tr>
<td>26 February 2004*</td>
<td>Elsenburg Agricultural Research</td>
<td>Mandy†</td>
<td>E03404</td>
<td>γ</td>
</tr>
<tr>
<td>26 February 2004</td>
<td>Troughend</td>
<td>SA saddle horse 1</td>
<td>E03104</td>
<td>αi</td>
</tr>
<tr>
<td>27 February 2004</td>
<td>Oakhill</td>
<td>Murphy</td>
<td>E03304</td>
<td>αii</td>
</tr>
<tr>
<td>28 February 2004</td>
<td>Daktari</td>
<td>Murphy Brown</td>
<td>E02904</td>
<td>αii</td>
</tr>
<tr>
<td>03 March 2004*</td>
<td>Troughend</td>
<td>Melody Fire†</td>
<td>E03904</td>
<td>αiii</td>
</tr>
<tr>
<td>04 March 2004*</td>
<td>Avontuur</td>
<td>Bow Street Belle (blood)</td>
<td>E04304</td>
<td>αii</td>
</tr>
<tr>
<td>09 March 2004</td>
<td>Avontuur</td>
<td>Bow Street Belle (organs)</td>
<td>E04804</td>
<td>αiii</td>
</tr>
<tr>
<td>13 March 2004</td>
<td>Daktari</td>
<td>Calypso</td>
<td>E05804</td>
<td>αiii</td>
</tr>
<tr>
<td>15 March 2004*</td>
<td>Vredenheim</td>
<td>Bruin Perd (blood)</td>
<td>E06204</td>
<td>αiii</td>
</tr>
<tr>
<td>16 March 2004*</td>
<td>Avontuur</td>
<td>Special Edition (blood)</td>
<td>E06104</td>
<td>αi</td>
</tr>
<tr>
<td>17 March 2004</td>
<td>Vredenheim</td>
<td>Bruin Perd (organs)</td>
<td>E06504</td>
<td>αiii</td>
</tr>
<tr>
<td>17 March 2004</td>
<td>Troughend</td>
<td>SA saddle horse 2</td>
<td>E06404</td>
<td>αii</td>
</tr>
<tr>
<td>18 March 2004</td>
<td>Avontuur</td>
<td>Special Edition (organs)</td>
<td>E06604</td>
<td>αi</td>
</tr>
<tr>
<td>24 March 2004</td>
<td>Avontuur</td>
<td>2003 Stars and Stripes</td>
<td>E08104</td>
<td>αi</td>
</tr>
<tr>
<td>29 March 2004</td>
<td>Kalbaskraal</td>
<td>Amber</td>
<td>E08604</td>
<td>αiii</td>
</tr>
</tbody>
</table>

†Horse clinically ill, did not die.
DISCUSSION

The phylogeny of the S10 gene of the genus Orbivirus was examined. The high mutation rate of RNA viruses and high degree of saturation of the sequences masked the phylogeny, and the evolution of the gene could not be determined with any certainty. AHSV, BTV and EEV all formed monophyletic groups and BTV was closer to AHSV than EEV in terms of evolutionary distance.

One hundred and forty-five AHSV were isolated and sequenced from blood and organ samples submitted to our laboratory between 2004 and 2006. During that period, the virus collection was not representative of the AHS situation in the country, as samples submitted to the DVTD were biased towards those samples submitted from the formal equine sector, due to the costs involved in veterinary examinations and diagnostic tests. Very few samples were submitted from the Eastern Cape, yet it has one of the highest unvaccinated horse populations in the country (livestock values were obtained from the Department of Agriculture).

The phylogenetic analysis confirmed the presence of three distinct S10 phylogenetic clades (α, β and γ), as previously reported (van Niekerk et al., 2003; Martin et al., 1998; Sailleau et al., 1997). In our analysis, serotype 8 clustered only in the α clade. In contrast, van Niekerk et al. (2003) found that serotype 8 clustered in both the α and β clades, while Martin et al. (1998) found that it clustered in the γ clade. Similar results were obtained when their data (GenBank accession nos AJ007307 and AF276685) were included in our analyses (see Supplementary material available in JGV Online). With our data, the β clade was made up exclusively of serotypes 3 and 7.

Serotypes differed in the mean evolutionary distance of S10 within a serotype. Serotype 7 isolates showed very little variation in their S10 sequences. Although an AHS epidemic, caused by serotype 7, occurred in KZ in 2005, samples were also submitted from the NW, GP, KZ and EC. On the other hand, the S10 sequences from serotype 5 isolates were very diverse, appearing in the α and γ clades and in different clusters within a clade. The differences in variation between serotypes may be a reflection of the disease characteristics at the time, an epidemic resulting in a viral population of fairly low genetic diversity, due to a rapidly replicating population under little selective pressure. On the other hand, endemic stability would favour a viral population of higher genetic diversity, due to a long period of selective pressure.

A large S10 distance within a serotype may be an indication of reassortment (exchange of segments of the AHSV segmented genomes between different AHSV infecting the same midge or horse). There are possible examples of reassortment, e.g. AHSV serotype 4 (Jane4) shares the identical S10 sequence to some of the serotype 1 isolates from the Western Cape (Fig. 3).

Although strong purifying selection of the S10 gene was evident, there were isolates for which positive selection was detected. The positions of the amino acid changes of those isolates were plotted in Fig. 5. They corresponded to hydrophilic regions that were not well conserved, suggesting that regions under selective pressure are extra/intracellular. The most variable region of the gene was between residues 139 and 158 and surrounded by two hydrophobic regions (van Staden et al., 1995). Purifying or positive selection of the virus during cell culture amplification could not be accounted for, but the effect was minimized by sequencing the isolates obtained from three or fewer passages.

The use of molecular epidemiology of the S10 gene was illustrated with the AHS outbreak in the Western Cape in 2004, by showing that there were two separate AHSV circulating in the area. In contrast to the serotype 1 γ clade viruses, which shared the same S10 sequence, the serotype 1 α clade isolates showed some genetic variation, which indicated a population of AHSV with differing genome sequences (quasispecies), rather than representing an accumulation of mutations over time. It was therefore not possible to identify a chain of infection for the outbreak in the Western Cape.

The significance of reassortment of field viruses with vaccine strains could not be established, as S10 sequence data of vaccine strains are unavailable. It is possible that the serotype 1 α or γ clade might represent a vaccine strain. Widespread vaccination is practised in the area with 70% of horses vaccinated against AHSV (Sinclair, 2006). In addition, concern was expressed that the closest related 2004 serotype 1 γ clade viruses (E25304 and E29904) to the Western Cape isolates were as a result of vaccination, as both horses became ill 3 weeks after vaccination against AHSV. However, the S10 sequence from E25304 and E29904 was exactly the same as that obtained from Jane1, a documented virulent field strain. Likewise, the Jane4 sequence is exactly the same as the serotype 1 α clade viruses from the Western Cape and is also virulent. The phylogeny suggests that none of the Western Cape isolates represents vaccine strains.
The S10 gene is the second most variable gene of the AHSV genome after VP2 and is therefore well suited for use in characterizing AHSV strains. In addition, the small size of the gene (755–764 bp) and conserved terminal regions facilitate easy and quick sequencing. It is not clear why there is such a large variation in the S10 gene. NS3 is membrane-associated and therefore may be under immunological pressure, unlike other AHSV non-structural proteins that are intracellular and much more conserved than NS3 (M. Quan, unpublished data).

The establishment of an S10 sequence database is important for characterizing outbreaks of AHS. The sequence data of the AHSV vaccine strains represent a gap in the current S10 sequence database. In addition, it would be desirable to sequence VP2, or a portion thereof, and to combine these data in the phylogenetic analysis. The database will be an essential resource for elucidating the epidemiology of AHS, but needs to be backed up with a sensitive and accurate surveillance system and submission of samples for viral isolation and characterization.

ACKNOWLEDGEMENTS

We would like to thank N. James MacLachlan for useful comments about the manuscript. This research was funded by the Equine Research Centre, University of Pretoria and the Technology and Human Resources for Industry Programme (THRIP) of the Department of Trade and Industry and the National Research Foundation.

REFERENCES


