TUTDoR

The implications of virus diversity within the SAT 2 serotype for control of foot-and-mouth disease in sub-Saharan Africa.

Item Type	Article
Authors	Bastos, A. D. S.;Haydon, D. T.;Sangare, O.;Boshoff, C. I.;Edrich, J. L.;Thomson, G. R.
DOI	10.1099/vir.0.18859-0
Publisher	Microbiology Society Publisher
Rights	Attribution-NonCommercial-ShareAlike 4.0 International
Download date	2026-05-12 22:13:14
Item License	http://creativecommons.org/licenses/by-nc-sa/4.0/
Link to Item	https://hdl.handle.net/20.500.14519/892

The implications of virus diversity within the SAT 2 serotype for control of foot-and-mouth disease in sub-Saharan Africa

A. D. S. Bastos,[1,2] D. T. Haydon,[3] O. Sangaré,[2,4] C. I. Boshoff,[2] J. L. Edrich[1] and G. R. Thomson[2,5]

[1]Mammal Research Institute, Department of Zoology & Entomology, University of Pretoria, Pretoria 0002, South Africa

[2]ARC-Onderstepoort Veterinary Institute, Exotic Diseases Division, Private Bag X5, Onderstepoort 0110, South Africa

[3]Centre for Tropical Veterinary Medicine, University of Edinburgh, Edinburgh EH25 9RG, UK

[4]Laboratoire Central Veterinaire, BP 2295, Bamako, Mali

[5]Organization of African Unity/Inter-African Bureau for Animal Resources (OAU-IBR), PO Box 30786, Nairobi, Kenya

Correspondence
A. Bastos (at University of Pretoria)
ADBastos@zoology.up.ac.za

SAT 2 is the serotype most often associated with outbreaks of foot-and-mouth disease (FMD) in livestock in southern and western Africa and is the only SAT type to have been recorded outside the African continent in the last decade. Its epidemiology is complicated by the presence of African buffalo (*Syncerus caffer*), which play an important role in virus maintenance and transmission. To assess the level of genetic complexity of this serotype among viruses associated with both domestic livestock and wildlife, complete VP1 gene sequences of 53 viruses from 17 countries and three different host species were analysed. Phylogenetic analysis revealed eleven virus lineages, differing from each other by at least 20 % in pairwise nucleotide comparisons, four of which fall within the southern African region, two in West Africa and the remaining five in central and East Africa. No evidence of recombination between these lineages was detected, and thus we conclude that these are independently evolving virus lineages which occur primarily in discrete geographical localities in accordance with the FMD virus topotype concept. Applied to the whole phylogeny, rates of nucleotide substitution are significantly different between topotypes, but most individual topotypes evolve in accordance with a molecular clock at an average rate of approximately 0·002 substitutions per site per year. This study provides an indication of the intratypic complexity of the SAT 2 serotype at the continental level and emphasizes the value of molecular characterization of diverse FMD field strains for tracing the origin of outbreaks.

Received 26 September 2002
Accepted 5 February 2003

INTRODUCTION

Foot-and-mouth disease virus is a member of the genus *Aphthovirus* within the family *Picornaviridae*. These single-stranded, positive-sense RNA viruses occur as seven immunologically distinct serotypes, of which three, the South African Territories (SAT) types 1–3, are endemic to sub-Saharan Africa. SAT serotypes differ from each other with respect to distribution, outbreak incidence in domestic livestock and infection rates in wildlife. SAT 2 is the most widely distributed serotype throughout sub-Saharan Africa (Brooksby, 1972; Ferris & Donaldson, 1992) and has made incursions into the Middle East on two separate occasions, once in 1990 and again in 2000 (Ferris & Donaldson, 1992; records of the OIE). SAT 2 is also the virus type most frequently associated with outbreaks of the disease in livestock in southern and West Africa (Thomson, 1994; Sangaré, 2002) and with clinical cases of the disease in wildlife (Keet *et al*., 1996; Bastos *et al*., 2000). Despite the regular involvement of this serotype in outbreaks and evidence that buffalo are the primary source of infection for other cloven-hoofed species in southern Africa (Dawe *et al*., 1994a; Bastos *et al*., 2000), it is not the serotype most frequently recovered from foot-and-mouth disease (FMD)-infected buffalo populations in South Africa (records of the Onderstepoort Veterinary Institute). Consequently, it has been proposed that the different SAT virus types may have differential abilities in crossing species barriers and that SAT 2 appears to be the most efficient in doing so (Bastos, 2001).

African buffalo (*Syncerus caffer*) are particularly efficient maintenance hosts and play a central role in the epidemiology of the disease due to their ability to maintain and

transmit FMD virus (reviewed by Thomson, 1996). Results indicate that SAT-type viruses can persist in an individual animal and in an isolated herd for at least 5 and 24 years, respectively (Condy *et al.*, 1985) and that numerous antigenic and genetic variants are generated in an individual animal during persistence (Vosloo *et al.*, 1996). Transmission of SAT-type virus from persistently infected African buffalo to cattle under experimental and natural conditions has been unequivocally demonstrated (Dawe *et al.*, 1994a, b; Vosloo *et al.*, 1996; Bastos *et al.*, 2000). Separation of buffalo and livestock by fencing and vaccination of cattle with viruses that are antigenically closely related to those carried by nearby buffalo form an integral part of FMD control in southern African countries (Hunter, 1998).

Overall antigenic diversity as measured by ELISA and virus neutralization test appears to be reflected genetically in the VP1 gene (Hunter *et al.*, 1996; Araújo *et al.*, 2002), which is one of four structural proteins encoded within the P1 region of the genome. VP1 contains the major neutralizing sites of the virus (Crowther *et al.*, 1993) as well as the integrin cell-attachment site 'RGD', which is located on a highly flexible loop that protrudes from the surface of the virion (Acharya *et al.*, 1990). The multiplicity of roles that VP1 has in receptor-recognition, neutralization and antigenic diversity together with the valuable molecular epidemiological insights obtained from analysis of nucleotide sequences (reviewed by Sobrino *et al.*, 2001) has made this the most intensively studied gene of the FMD viral genome. However, most of these efforts have been directed at the European serotypes, O, A and C, with complete VP1 gene data being restricted to just three SAT 2-type viruses (van Rensburg & Nel, 1999).

The benefits of extensive genetic characterization of the SAT-type viruses are perhaps best illustrated by reports from southern Africa where the central role played by buffalo in the epidemiological situation has been described (Dawe *et al.*, 1994a; Vosloo *et al.*, 1995; Bastos *et al.*, 2000, 2001). Here, the establishment of a partial VP1 gene sequence database for southern African SAT-type buffalo viruses was shown to be invaluable for accurately tracing the source of SAT 1-type outbreaks in this region (Bastos *et al.*, 2001) and for pinpointing the origin of illegally moved buffalo infected with SAT 3-type viruses (Bastos, 2001; Vosloo *et al.*, 2001).

Of the six serotypes occurring in Africa, only serotype A has been studied at the continental scale. Analysis of partial VP1 gene nucleotide sequences revealed the presence of six regionally and genetically distinct virus lineages (Knowles *et al.*, 1998) that conform to the topotype classification (Samuel & Knowles, 2001a). The identification of regionally distinct viruses ensures that the origin of viruses imported trans-regionally and trans-continentally can be accurately determined (Knowles *et al.*, 1998; Sangaré *et al.*, 2001; Sangaré, 2002). Studies similar to that of type A (Knowles *et al.*, 1998) applied to the remaining five African serotypes are beneficial not only for the African continent, but for all countries with FMD-free status, as the disease security of these countries is threatened by the importation of exotic viruses, as occurred in 2001 in the UK (Samuel & Knowles, 2001b). Past records of multiple SAT-type virus incursions into the Middle East and known transmission routes from the latter region to Europe (Kitching, 1998) make it imperative to characterize viruses occurring in FMD endemic areas.

Genetic characterization of SAT 2 viruses from African buffalo maintenance host populations throughout Africa has not been adequately addressed, nor has any attempt been made to determine genetic variability at the continental scale. The former has implications for accurate determination of the source of SAT 2-type infections whilst the latter provides some measure of the complexity of disease control through vaccination. It is with these objectives in mind that SAT 2-type viruses from different regions and species, collected over a 52 year period have been selected for genetic characterization. Specifically, we examine these complete VP1 gene sequences for evidence of inter-lineage recombination, we relate the topological structure of the estimated phylogeny to the geographical location of the virus isolates, we examine the applicability of a molecular clock to describe the evolution of this gene both within and between lineages, and we report the distribution of substitutions along the gene.

METHODS

Genetic characterization of SAT 2 viruses. Fifty SAT 2-type FMD viruses from 17 countries in Africa were selected for genetic characterization and are summarized in Table 1. Viral RNA extraction was performed directly on cell culture isolates using a modified silica-based method (Boom *et al.*, 1990) and cDNA was synthesized as previously described (Bastos, 1998). A forward primer binding within 1C (VP3), termed VP3-AB (5′-CACTGCTACCACTCRGA-GTG-3′) was designed specifically for this study. When combined with a reverse primer P1 (5′-GAAGGGCCCAGGGTTGGACTC-3′) complementary to the highly conserved 2A/B junction site (Beck & Strohmaier, 1987), a product of approximately 880 bp was obtained.

Nucleotide sequences were generated manually using a radionucleotide in the presence of 10 % DMSO according to previously described methods (Bastos, 1998). Alternatively, fluorescent dye deoxyterminator cycle sequencing (Perkin Elmer) was performed on an ABI Prism 377 DNA sequencer (Applied Biosystems). Four sequencing reactions were performed per isolate using the two external primers P1 and VP3-AB, and two internal primers, VP1Ub (5′-CCACGT-ACTACTTYTCTGACCTGGA-3′) (Bastos, 1998) and SAT2-D (5′-GGTGCGCCGTTGGGTTGCCA-3′). Two independent genomic amplification reactions were performed per isolate, with each amplicon being sequenced with both internal and external primers.

Phylogenetic analysis. Nucleotide sequences were aligned to published reference strains (Table 1) using DAPSA (Harley, 2001) and trimmed so that a homologous region of 648 nt, corresponding to the complete VP1 gene, was obtained. All data were submitted to GenBank under the accession numbers provided in Table 1 and where grid references were available these were used to accurately plot geographical distribution.

Phylogenetic reconstruction was performed using a minimum

Table 1. Summary of SAT 2-type FMD viruses used in this study

GR, game reserve; NP, National Park; SA, safari area; NK, not known.

Virus name	Country of origin	Year of sampling	Place of origin	Grid reference	Species of origin	GenBank accession no.
RHO/1/48	Zambia	1948	NK	NK	Bovine	AJ251475*
KEN/3/57	Kenya	1957	Wamba	37° 19′ E 0° 59′ N	Bovine	AJ251473*
ANG/4/74	Angola	1974	NK	NK	NK	AF479417
MAL/3/75	Malawi	1975	NK	NK	NK	AF367099
NIG/2/75	Nigeria	1975	NK	NK	Bovine	AF367139
SEN/5/75	Sénégal	1975	NK	NK	Bovine	AF367140
GAM/8/79	Gambia	1979	NK	NK	Bovine	AF479410
GAM/9/79	Gambia	1979	NK	NK	Bovine	AF479411
MOZ/1/79	Mozambique	1979	NK	NK	NK	AF367137
SEN/7/79	Sénégal	1979	NK	NK	Bovine	AF479412
ZAI/1/82	DRC	1982	Bibatama, Nord Kivu	29° 0′ E 01° 18′ S	Bovine	AF367100
MOZ/4/83	Mozambique	1983	NK	NK	Bovine	AF367101
PAL/5/83	South Africa	1983	Phalaborwa	31° 0′ E 23° 58′ S	Bovine	AF367102
SEN/3/83	Sénégal	1983	NK	NK	Bovine	AF479413
SEN/7/83	Sénégal	1983	NK	NK	NK	AF479414
ZIM/7/83	Zimbabwe	1983	Nyamandhlovu	28° 05′ E 19° 45′ S	Bovine	AF136607†
KNP/7/88	South Africa	1988	Kruger NP	31° 58′ E 24° 54′ S	Buffalo	AF367103
KNP/16/88	South Africa	1988	Kruger NP	31° 36′ E 24° 27′ S	Impala	AF367104
KNP/17/88	South Africa	1988	Kruger NP	31° 24′ E 24° 28′ S	Impala	AF367105
KNP/18/88	South Africa	1988	Kruger NP	31° 24′ E 24° 28′ S	Impala	AF367138
KNP/19/88	South Africa	1988	Kruger NP	31° 30′ E 24° 27′ S	Impala	AF367106
KNP/20/88	South Africa	1988	Kruger NP	31° 28′ E 24° 26′ S	Impala	AF367107
ZIM/1/88	Zimbabwe	1988	Hwange NP	27° 0′ E 19° 0′ S	Buffalo	AF367108
KNP/2/89	South Africa	1989	Kruger NP	31° 43′ E 24° 13′ S	Impala	AF367109
KNP/19/89	South Africa	1989	Kruger NP	31° 37′ E 24° 44′ S	Buffalo	AF367110
GHA/2/90	Ghana	1990	NK	NK	Bovine	AF479415
BUN/1/91	Burundi	1991	Bujumbura province	29° 19′ E 3° 22′ S	Bovine	AF367111
GHA/8/91	Ghana	1991	Tamale province	0° 50′ W 09° 24′ N	Bovine	AF479416
KNP/183/91	South Africa	1991	Kruger NP	31° 56′ E 25° 08′ S	Buffalo	AF367112
ZIM/Gn10/91	Zimbabwe	1991	Gonarezhou NP	32° 0′ E 21° 30′ S	Buffalo	AF367113
KNP/1/92	South Africa	1992	Kruger NP	31° 43′ E 24° 23′ S	Impala	AF367114
KNP/32/92	South Africa	1992	Kruger NP	31° 17′ E 22° 54′ S	Buffalo	AF367115
ZAM/9/93	Zambia	1993	Kafue NP	26° 0′ E 16° 30′ S	Buffalo	AF367116
ZAM/10/93	Zambia	1993	Kafue NP	26° 0′ E 16° 30′ S	Buffalo	AF367117
KNP/18/95	South Africa	1995	Kruger NP	31° 38′ E 24° 34′ S	Buffalo	AF367118
KNP/31/95	South Africa	1995	Kruger NP	31° 38′ E 24° 34′ S	Buffalo	AF367119
ZAM/7/96	Zambia	1996	Mulanga	25° 30′ E 17° 10′ S	Buffalo	AF367120
ZAM/10/96	Zambia	1996	Mulanga	25° 30′ E 17° 10′ S	Buffalo	AF367121
BOT/1/98	Botswana	1998	Nxaraga	23° 15′ E 19° 40′ S	Buffalo	AF367122
BOT/18/98	Botswana	1998	Nxaraga	23° 15′ E 19° 40′ S	Buffalo	AF367123
BOT/29/98	Botswana	1998	Vumbura	23° 35′ E 20° 15′ S	Buffalo	AF367124
BOT/31/98	Botswana	1998	Vumbura	23° 35′ E 20° 15′ S	Buffalo	AF367125
ERI/12/98	Eritrea	1998	Erythrea	38° 45′ E 16° 23′ N	Bovine	AF367126
NAM/286/98	Namibia	1998	East Caprivi GR	23° 20′ E 17° 52′ S	Buffalo	AF367127
NAM/292/98	Namibia	1998	East Caprivi GR	23° 20′ E 17° 52′ S	Buffalo	AF367128
NAM/304/98	Namibia	1998	West Caprivi GR	21° 50′ E 18° 15′ S	Buffalo	AF367129
ZIM/267/98	Zimbabwe	1998	Chizarira	28° 0′ E 17° 47′ S	Buffalo	AF367130
KEN/5/99	Kenya	1999	Athi river, Machakos	36° 59′ E 1° 27′ S	Bovine	AF367131
KEN/7/99	Kenya	1999	Kikuya, Kiambu	36° 39′ E 1° 13′ S	Bovine	AF367132
KEN/9/99	Kenya	1999	Kaloleni, Kilifi	39° 51′ E 3° 36′ S	Bovine	AF367133
RWA/1/00	Rwanda	2000	Gishwati district	29° 25′ E 01° 47′ S	Bovine	AF367134

Table 1. (*cont.*)

Virus name	Country of origin	Year of sampling	Place of origin	Grid reference	Species of origin	GenBank accession no.
SAU/6/00	Saudi Arabia	2000	Al Kahrj, Riyadh	47° 50′ E 24° 10′ N	Bovine	AF367135
ZIM/1/00	Zimbabwe	2000	Tengwe Farm	29° 30′ E 17° 12′ S	Buffalo	AF367136

*Unpublished.
†van Rensburg & Nel (1999).

evolution algorithm implemented in PAUPSTAR (v. 4.0b8; Swofford, 2001). Pairwise distances were computed using the HKY85 model of nucleotide substitution (Hasegawa *et al.*, 1985), with rate heterogeneity estimated using maximum-likelihood estimated over a topologically identical tree. 1000 bootstrap resamples were undertaken using the neighbor-joining method implemented in PAUPSTAR and assuming the same substitution model. Reconstruction was also conducted using maximum-parsimony, neighbor-joining and maximum-likelihood, in order to check for any substantial sensitivities to method of reconstruction. The program PLATO (Grassly & Holmes, 1997) was used to examine whether or not there were areas of the alignment in which support for the reconstructed phylogeny was significantly reduced. This method works by examining the likelihood of the phylogeny within windows of different sizes that are moved sequentially along the alignment. The program TIPDATE (Rambaut, 2000) was used to examine the average annual rate of nucleotide substitution over the phylogeny and to test the hypothesis that the phylogeny and various constitutive clades within it evolved in accordance with a molecular clock. This method uses a likelihood ratio test (LRT) to determine whether the reduction in likelihood of a tree with branch lengths estimated assuming a molecular clock is significantly reduced relative to the maximum-likelihood of a phylogeny assuming no clock. An LRT assumes that twice the difference in likelihood between two nested models is chi-squared distributed with degrees of freedom equal to the number of parameters by which the two models differ [see Rambaut (2000) for further details of this test]. Several different plausible positions for the roots of the phylogeny were selected to test this hypothesis. As this program cannot handle identical sequences within the same phylogeny duplicate sequences were first removed.

The average number of synonymous and non-synonymous substitutions per synonymous and non-synonymous site along the gene was calculated in DNASP (Rozas & Rozas, 1999). The occurrence of synonymous and non-synonymous substitutions along the VP1 gene was examined using the average number of pairwise synonymous and non-synonymous substitutions per codon position. A permutation test was used to test for significant clustering in the distribution of both types of substitutions, using a sliding window of 5 codon positions in length.

RESULTS

Phylogenetic analysis

The minimum evolution phylogeny is shown in Fig. 1 together with the species from which virus was recovered. Analysis using PLATO indicated that there was no subset of the alignment that was significantly inconsistent with this phylogeny, and thus no evidence for genetically dissimilar recombinants within the data set. Phylogenetic resolution of VP1 gene sequences revealed the presence of eleven distinct virus lineages (labelled A–K, Fig. 1) that have high bootstrap support (>90 %) and uncorrected sequence divergence values in excess of 20 %. This value coincides with the topotype divergence value previously defined for SAT 1 (Bastos *et al.*, 2001), which is applicable across all SAT types (Knowles & Samuel, 2003). Most of the topotypes clustered according to sampling locality, in accordance with the FMD topotype concept (Samuel & Knowles, 2001a), and were shown to constitute six regionally distinct virus lineages (I–VI; Fig. 2).

Within the southern African region, one of the four virus topotypes (A) had the same geographical distribution as that previously identified for SAT 1 (Bastos *et al.*, 2001). However, in contrast to the southern African SAT 1 topotypes, the western and northern Zimbabwe SAT 2 isolates were not sufficiently distinct to warrant placement of these viruses in two distinct topotype classes. Furthermore, two of the SAT 2 topotypes (C and D) overlapped slightly in their distributional range in northern Botswana (Fig. 2), a feature not observed for SAT 1. The fourth southern African topotype (K) was genetically so distinct from all other southern African topotypes that a different virus lineage (VI) was assigned.

In West Africa two regionally distinct topotypes (E and F) having an overlapping distributional range in Sénégal were identified (Fig. 2). As sequence divergence values of viruses for topotypes E and F equalled that observed for viruses from different virus lineages, two West African virus lineages (II and III) were defined. All West African viruses comprising topotype F had a unique VP1 protein length of 215 instead of 214 due to an amino acid insertion at position 29.

The north-east African virus lineage (IV) comprised two viruses, ERI/12/98 and SAU/6/00, which were recovered from outbreaks occurring in Eritrea in 1998 and in Saudi Arabia in 2000, respectively. The 100 % bootstrap support for this node, together with the high level of sequence identity (>90 %), indicate that the north-eastern African region was the most likely source of the virus introduced to the Middle East.

In East and central Africa four topotypes were discerned, three of which (H–J) constitute lineage IV and are represented by a single virus each from Rwanda, the Democratic

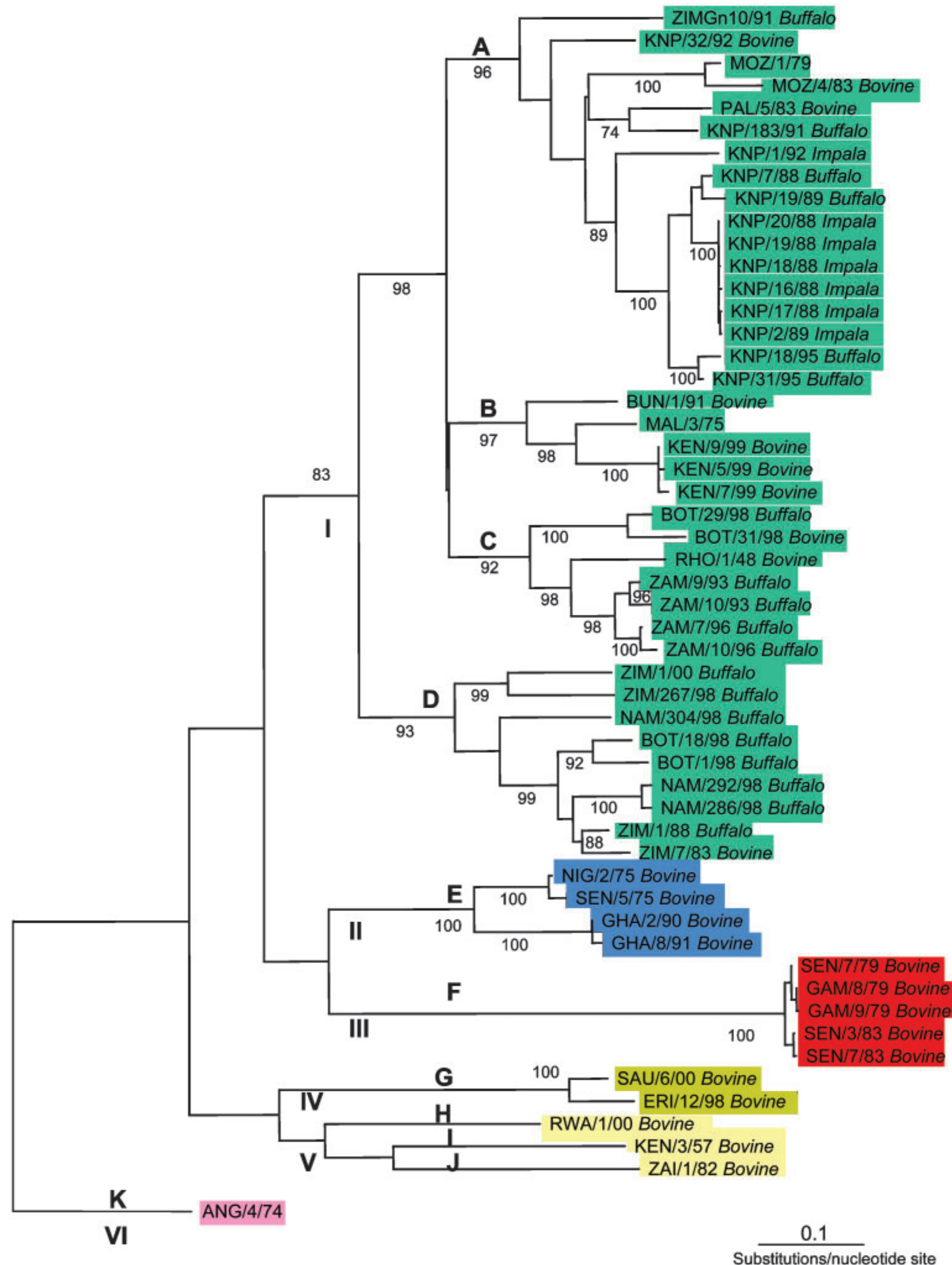

Fig. 1. Phylogenetic tree of the 53 VP1 sequences included in the study. The tree is constructed using a minimum evolution algorithm with the HKY85 model of nucleotide substitution and rate heterogeneity. Bootstrap values were generated using the same model of substitution and a neighbor-joining algorithm; values over 70 % are shown. The species from which virus was recovered is indicated in the tip labels. The major virus lineages are indicated I–VI and topotypes are designated by the letters A–K.

Republic of the Congo (DRC) and Kenya. The fourth topotype (B), represented by viruses from southern Kenya, Malawi and Burundi, was more closely related to topotypes occurring in the southern African region (topotypes A, C and D) than to those constituting the north-east African lineage (IV).

Molecular clock

The hypothesis that the rate of nucleotide substitution is constant over the whole phylogeny was strongly rejected ($P<0{\cdot}0001$) irrespective of the position of the putative root. After removal of some of the more obvious temporally aberrant sequences (KEN/3/57, MAL/3/75 and RHO/1/48),

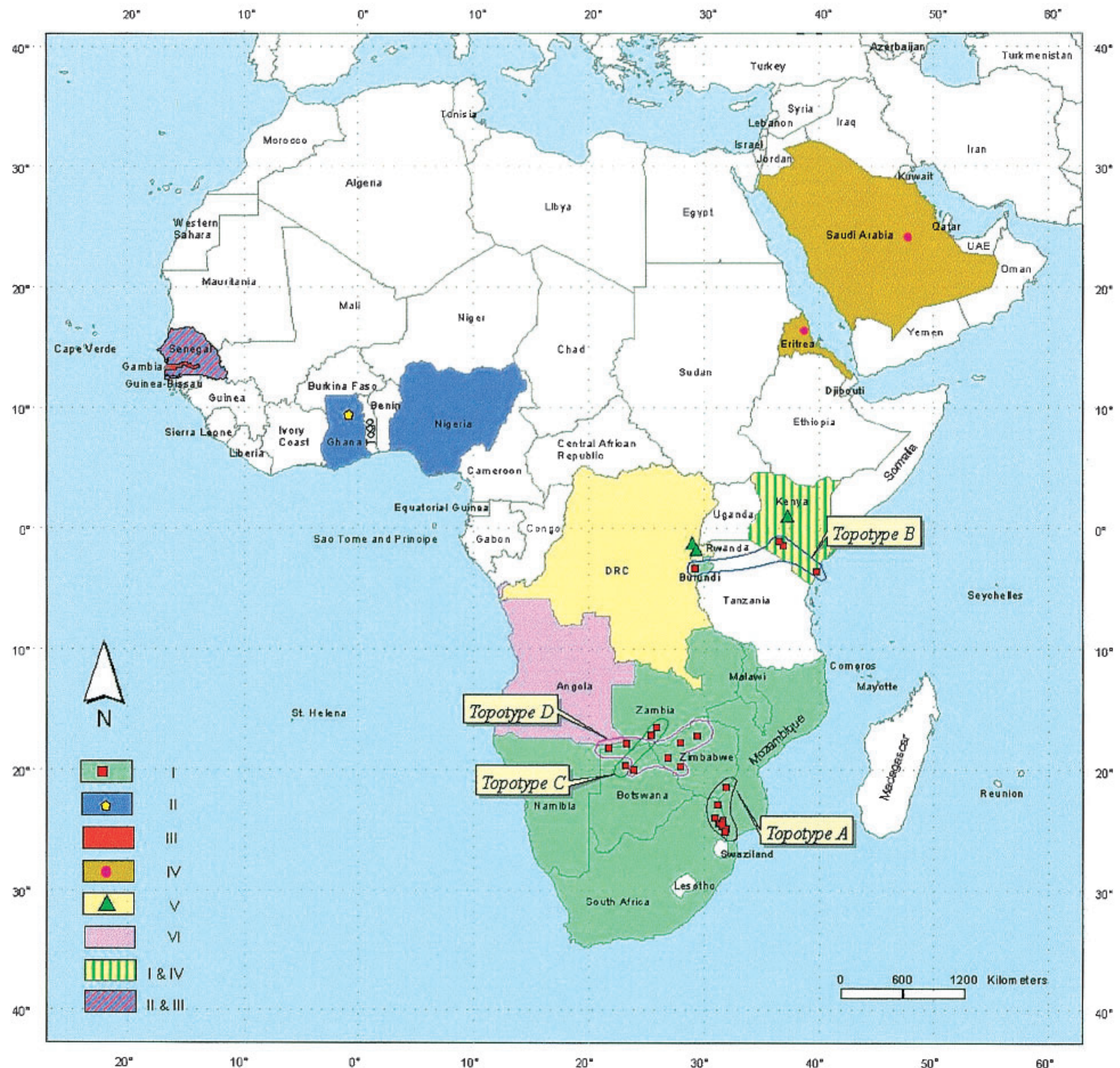

Fig. 2. Geographical distribution of the six major SAT 2-type virus lineages (I–VI) and the southern African topotype distributions (topotypes A–D) based on grid reference data.

the clock hypothesis was still strongly rejected ($P<0{\cdot}0001$). Applying the test to each of the groups A–F in Fig. 1 (each rooted on a sequence in a neighbouring topotype) yielded an average annual rate of nucleotide substitutions per site of 0·0023 (range 0·0002–0·0047). Statistics from the LRT applied to each of these six topotypes and the corresponding uncorrected P-values are reported in Table 2. However, since six hypotheses are being tested, the critical P-value for rejection of the null hypothesis of constant rates within topotypes might be lowered from 0·05 to $0{\cdot}05/6=0{\cdot}0083$, which would lead to the conclusion that a molecular clock could be applied to five of the six topotypes examined here.

Distribution of mutations

Of the 651 nucleotide sites characterized in this study, only 250 (38·4 %) were invariant across all taxa. At the amino acid level this translated to 92 of a total of 215 amino acid sites (42·4 %) being completely conserved (Fig. 3). These invariable sites include the cysteine residue at the base of the G–H loop and the cell-attachment site 'RGD' within the loop. Residues flanking the RGD were completely conserved at the -1 and $+4$ positions, whilst a positively charged arginine was present at the $+1$ position in all but one virus, NAM/292/98, that had a non-polar proline (P) instead. Three neutralization sites, previously identified by mAb studies at positions $+2$ and $+3$, positions $+10$ and

Table 2. Results of the analysis testing for the presence of a molecular clock

Topotype*	Rate†	χ^2‡	Degrees of freedom	*P*-value§
A	0·0014	28·2	14	0·013
B	0·0024	9·72	3	0·021
C	0·0002	13·5	5	0·019
D	0·0017	19·4	7	0·007
E	0·0047	1·86	2	0·395
F	0·0037	6·2	3	0·102

*Topotypes for which there exist more than two sequences.

†Maximum-likelihood estimate of rate of nucleotide substitutions per site per year.

‡Chi-squared value equals twice the difference in likelihoods between models fitted with and without a molecular clock.

§Critical *P*-value = 0·0083 (Bonferroni correction).

+12 of the G–H loop region downstream of the 'RGD' (Crowther *et al.*, 1993), showed moderate levels of variation at the +2 and +3 positions and high levels of variation at the +10 and +12 positions. At the +10 position the G observed in the reference strain (RHO 1/48) was replaced with S/D/N/E, whilst either K/R/Q/T occurred instead of S at the +12 position. At the VP1/2A cleavage site amino acid sequences KQ/LL and RQ/TL predominated, with the latter being unique to viruses of Eritrean and Saudi Arabian origin. The region upstream of αZ (amino acid positions 21–30) was not only hypervariable, but also displayed size variation due to the presence of a tryptophan residue in all West African topotype F viruses (Figs. 1 and 3).

The average pairwise synonymous (π_s) and non-synonymous (π_{ns}) nucleotide diversity per synonymous and non-synonymous nucleotide site was 0·618 and 0·117 respectively. Fig. 4 shows the distribution of nucleotide substitutions along the VP1 gene. The synonymous distributions show no evidence of any clustering regardless of the window size, whilst the non-synonymous profile exhibits a single significant cluster with window size 5 within codons 138–142 ($P=0·022$).

DISCUSSION

Phylogenetic analysis based on 53 complete VP1 gene sequences of diverse SAT 2-type field strains revealed the presence of 27 genotypes which could be assigned to one of eleven topotypes (A–K), corresponding to six regionally distinct evolutionary lineages (I–VI). The lack of any evidence for inter-regional recombination suggests that SAT 2 viruses evolve independently in these six geographical areas of Africa. Nineteen of the 27 genotypes occurred within lineage I which incorporates countries in southern and East Africa. The remaining lineages (II–VI) were genotypically not as well represented, but were nonetheless topotype rich, with seven different topotypes apparent. The resolution of virus relationships at genotype level (viruses that differ from each other by no more than 10 % across the complete VP1 gene) is essential for accurately tracing the source of temporally unrelated viruses (Bastos *et al.*, 2000). From the results presented here, it is clear that outbreaks occurring in the lineage I region (Fig. 2) should be traceable to specific countries, game parks and even to specific regions within game parks, due to the high genotype-richness and locality-specific groupings identified. Further identification and genetic characterization of genotypes within the other evolutionary lineages, particularly lineages IV–VI, should be expanded to ensure similar levels of resolution and assessment of annual rates of nucleotide substitution at the topotype level.

Overlap in the distributional range of topotypes was observed in restricted areas in southern, eastern and western Africa. In southern Africa topotypes C and D display overlap in distributional range at the border region where Botswana, Angola, Namibia, Zambia and Zimbabwe converge. Within this area of convergence numerous disease outbreaks have been recorded which are attributed to intensive livestock trade (Mweene *et al.*, 1996). Thus, the blurring of topotype boundaries arising from transboundary movements that occurs when a 'foreign' topotype is introduced from a neighbouring country explains the presence of the predominantly Zambian topotype (topotype C) in buffalo from northern Botswana. Other countries containing multiple virus lineages (and topotypes) within their borders include Kenya (lineage I and IV) in East Africa, where cattle trading occurs across all borders (Ndiritu, 1984), and Sénégal (lineage II and III) in West Africa, where the practice of transhumance is a contributing factor (Sangaré, 2002). These three virus exchange 'hotspots' reveal that disease control would likely benefit through enforcing stricter border controls. The remaining seven topotypes show no evidence of overlap in geographical range.

From the amino acid sequence alignments it is clear that the VP1 gene of SAT 2-type viruses has features in common with serotype O. In both serotypes there is a conserved cysteine residue at the base of the G–H loop that may account for the presence of conformational neutralizing epitopes in these serotypes (Xie *et al.*, 1987; Crowther *et al.*, 1993). Complete conservation of this cysteine residue occurs in all SAT serotypes (Bastos, 2001; Bastos *et al.*, 2001). However, the presence of the corresponding cysteine in VP2, with which Cys-134 of VP1 forms a heterodimer in type O viruses, has only been identified in two SAT 2-type viruses (van Rensburg & Nel, 1999). A feature unique to SAT 2-type viruses is the presence of a highly conserved positively charged arginine residue within the G–H loop flanking the carboxyl-end of the 'RGD'. In all other serotypes, a non-polar leucine predominates at this position which together with a highly conserved leucine at the +4 position has been shown to be critical for modulating interactions with integrin receptors (Mateu *et al.*, 1996).

```
                                                  αZ         βB         βC              αA         βD
                   10         20         30         40         50         60         70         80
RHO/1/48    TTSVGEGADV VTTDPSTHGG SVIEKKRM-H TDVAFVLDRF THVHTSKTTF NVDLMDTKEK TLVGALLRAS TYYFCDLEIA
KEN/3/57    ...A....E. ...N.T.... K.TTPS.V-. .....L...S .....DT.A. V......... A....I..SA ........V.
ANG/4/74    ...A...... .......... A.TSV..K-. .E....M... .....N.... EIN....ND. ..I.....SA ..........
MAL/3/75    ...A...... .......... ..V..R..-. .......... .....N.... ........QQ A......... ..........
SEN/5/75    ...A..S... .....T.... A.TNPR.K-. .....L...S .....G.... E.N....... A....V...A ..........
ZAI/1/82    ...A...... .....T.... ..QTPR.V-. .....L...S ...F.D..S. A......... A....I..SA ..........
MOZ/4/83    ...A....E. .......... Q.E..R..-. .......... ....KN..S. .......... .......... ..........
SEN/3/83    ...A...... .......... V.DT...RW. ...S..M..A .....N.... A...L...D. .........T ....A.....
PAL/5/83    ...A...... .S........ R.E..R..-. .......... .......... .......... .......... ..........
ZIM/7/83    ...S...... .......... A.T....V-. ......M... ...L.NR.A. A......N.. ....G....A ..........
KNP/20/88   ...A...... .......... Q.V..R..-. .......... .....N.... ........D. .......... ..........
KNP/19/89   ...A...... .......... Q.V..R..-. .......... .....N.... ........D. .......... ..........
GHA/2/90    ...A...... .....T.... A.ANPR.K-. ...G.L...S .....G.... E.N....... A....V...A .........T
BUN/1/91    ...A...... .......... M.T..R..-. .......... .....K.... ........PQ A......... ..........
KNP/183/91  ...A...... .......... Q.M..R..-. .......... .....N.... .......... .......... ..........
ZIMGn10/91  ..TA...... ..S....... Q.S..R..-. ...T?..... .....N..S. ........D. A......... ..........
KNP/1/92    ...A...... .......... Q.AD.R..-. .......... .....N.... S......... .......... ..........
KNP/32/92   ...A...... .......... Q.V..R..-. .......... .....N.... ...I...... .......... ..........
KNP/18/95   ...A...... ?......... Q.V..R..-. .......... .....N.... ........D. .......... ..........
ZAM/10/96   ...A...... ...V...... R.V..R..-. .......... .......... .......... .......... ..........
BOT/18/98   ...S...... .......... A.T..R..-. ......M... ...L.N..A. A......N.. .........A ..........
BOT/29/98   ...A...... .......... ..V..R..-. .......... .......... .....E..D. .......... ..........
ERI/12/98   ...A...... .......... N.Q.GR.K-. .E...L...S .....N..S. V.......G. A....I.... ..........
NAM/286/98  ...S...... .......... A.T..R..-. ......M... ..AL.N..A. A......N.. ..I......A ..........
NAM/304/98  ...A...... .......... V.V..R..-. ......M... ...L.D.SA. A......N.. .......... ..........
ZIM/267/98  ...S...... .......... ..T..R..-. ......M... .....N..ST VI.....N.. .......... ......M...
KEN/5/99    ...A...... ......S... ..V..R..-. .......... .......... ....L...QH A......... ..........
RWA/1/00    ...A...... .....T.... HPNAAR.K-. ..I..L...S .....N..S. A......... A....V..SA ........V.
SAU/6/00    ...A..S... .......... N.Q.GR.K-. .E...L...S .....N..S. V......... A....I.... ..........
ZIM/1/00    ...S...... I......... ..V..R..-. .......... .....N..S. V......N.R .......... ..........

            βD     βE       αB         βF         βG1        βG2                  G-H loop
                   90         100        110        120        130        140        150        160
RHO/1/48    CVGEHSRVFW QPNGAPRTTQ LGDNPMVFSH NGVARFAIPY TAPHRLLATR YNGECKYKQE AKAIRGDRAV LAAKYAGTSH
KEN/3/57    ...K.K.... .......... .......... .N.T.....F .......S.V .....E.TKT VT......E. ..Q..SSAK.
ANG/4/74    ...D.Q.... .......... .R......A. .N.T...L.. .R.......H .....T.TKK VVP......A .R....D.K.
MAL/3/75    ...T.K..Y. .........V .........A. ...T...... .........T .......TDR VS........ ......DSR.
SEN/5/75    ...D.K.... .......A.. .......... .K.T.....F .......S.V .....N.STS VTP....... ......S.K.
ZAI/1/82    ...D.T.... .........K ........YA. .H.T.....F .......S.? .....E.RTN VT......Q. ..Q...S.K.
MOZ/4/83    ...T.RH.Y. .......... .........K ...T...V.. .......S.V .....E..TA VT........ ......D.R.
SEN/3/83    ...D.K..Y. .......... .......... .R.T...... .......S.V .....V.ADT VAP.....Q. .....SNRK.
PAL/5/83    ...D.T.... .......... .........N ...T...V.. .......S.V .....R.ETP VT........ .....S.KQ.
ZIM/7/83    .L...E..W. .........T .R........ .N.T...V.. .......S.. .......T.Q ST........ ......N.K.
KNP/20/88   ...D.K..Y. .........E .........N K..T...V.. .......S.V .......ETP VT........ .....SN.K.
KNP/19/89   ...D.Q..Y. ........RE .........N KR.T...V.. .......S.V .......ETP VT........ .....SNIK.
GHA/2/90    .....A.... .......... .......... .N.T...... .......S.V .......NTT VQP....... ..N...NRK.
BUN/1/91    ...D.K..Y. .......... .......... G..T...V.. .S........ .......TDK VS........ ......D.R.
KNP/183/91  ...T.A.... .......... .........N ...T...V.. .......S.V .....V.NTP VT........ .....S.RQ.
ZIMGn10/91  ...K.T.... ........SE .......... .A.T...V.. .......S.V .....R.TER VT........ ......N.R.
KNP/1/92    .....T..S. .......... .........R D..T...V.. .......S.V .....S.QVP VT........ .....SN.K.
KNP/32/92   ...D.T.... .......... .........R ...T...V.. .......S.V .....E..TR VT........ ..S...DAR.
KNP/18/95   ...D.K..Y. .........E .........S ...T...V.. .......S.V .....E.RTP VA........ .....SNIK.
ZAM/10/96   .....A..Y. .......... .......... .K.T...... .......... .......T.. .R........ .......AK.
BOT/18/98   .L...T..W. .........T .R........ .N.T...V.. .......S.. .......T.Q SA.......A ..QT..N.R.
BOT/29/98   .....K.... .........V .......... .K.T...... .......S.V .....R.T.K TQ........ ..Q...N.R.
ERI/12/98   ...D.T.... .......... .......YAK G..T.....F .......S.V .....T.AKT .T.......A ......ASV.
NAM/286/98  .L...T..W. H...S....T .RH....... .K.T...V.. .......S.. .......T.R ST........ ......N.K.
NAM/304/98  .I...A..W. .......... .R........ .N.T...V.. .......S.. .......T.Q ST........ ..S..SNIK.
ZIM/267/98  ...T.T..W. .......... .R........ .N.T...L.F .......S.. .......T.T ST........ ......NVR.
KEN/5/99    ...T.K..Y. .........T .........A. ...T...... .........V .......TDR VS......LL ......ESR.
RWA/1/00    ...D.A.... .......... .........A. .K.T...... .......S.V .....E.TKT VS........ ......SGK.
SAU/6/00    ...D?T.A.. .......... ........AK G..T.....F .......S.V .....V..KT PT.......A ......DST.
ZIM/1/00    .L.K.A..W. .......... .R........ .N.T...L.. .......S.. .......T.T SG........ ......N.K.
```

```
                       βH            βI                  C-terminus        ↓   2A
                   170        180        190        200        210   215
RHO/1/48    ALPSTFNFGH VTADKAVDVY YRMKRAELYC PRPLLPAYDH AGRDRFDAPI GVEKQ   LF
KEN/3/57    S........F .....P.... .......... ..A.....T. ..G....... ..A..   .L
ANG/4/74    S........F .....P.... F....T.... ..A....... .HN....... .....   .C
MAL/3/75    .......... ....EP.... .......... ......T.Q. NN........ .....   .C
SEN/5/75    T........Y ....AP.... .....S.... .......... QSH....... .....   .C
ZAI/1/82    .........F ..T..P.... .......... .......... QN........ .....   .C
MOZ/4/83    T......... ...EQP.... .......... .DHSF..... .......... .....   .C
SEN/3/83    Q......Y.Y ....EP.... .......F.. ..A....... HS.....S.. .....   .C
PAL/5/83    T......... .......... .......... .......... .D........ .....   .C
ZIM/7/83    K......... .....P.... ......AV.. ......G... .D.....S.. .....   .C
KNP/20/88   T......... ....SS.... .......... .......... .D........ .....   .C
KNP/19/89   T......... ....NS.... .......... .......... .S........ .....   .C
GHA/2/90    T........Y .....P.... .......... .......... QS........ .....   .C
BUN/1/91    S......... ....AP.... .......... ........Q. EN........ .....   .C
KNP/183/91  T......... ....A..... .......... .......... GN........ .....   .C
ZIMGn10/91  .......... .....P.... .......... .......... .N.....S.. .....   .C
KNP/1/92    T......... I....Q.... .......... .......... .D........ .....   .C
KNP/32/92   T..P...... .....S.... .......... .......... .N........ .....   .C
KNP/18/95   T......... .....S.... .......... ........N. VD........ .....   .C
ZAM/10/96   .......... ....A..... .......... ........E. LS.....G.. .....   .C
BOT/18/98   E........Y .....P.... .......... ......G... .N.....S.. .....   .C
BOT/29/98   S......... ....E..... .......... .....?..E. ND...Y.G.. .....   .C
ERI/12/98   T..Q.....F ..V..P.... .......... .......... .S........ ...R.   TL
NAM/286/98  E........Y .....P.... .......... ......G.E. .D.....S.. .....   .C
NAM/304/98  E........Y .....P.S.. .....T.... .......... .D.G...... .....   ?C
ZIM/267/98  E........Y ..V.NP..I. .......... ..A....... QN........ .....   .C
KEN/5/99    T......... ....QP.... .......... ........Q. GN........ .....   .C
RWA/1/00    T........F .....P.N.. .......... ...F...... .S........ .....   .C
SAU/6/00    T........F ..V..P.... .......... ........E. T.G....... ...R.   TL
ZIM/1/00    E........F ....EP..A. .......... .......... GS........ .....   .C
```

Fig. 3. Amino acid alignment of 30 SAT 2 FMD viruses representative of the major genotypes identified in this study. A '.' indicates an amino acid site identical to that of the sequence RHO/1/48 and '?' denotes ambiguous sites. Regions corresponding to beta-sheets are shaded in green, regions corresponding to alpha-helices are shaded in yellow, the major immunodominant sites of the protein (the G–H loop and carboxyl-terminus region) are shaded in red and the conserved cysteine residue at the base of the G–H loop is indicated in blue. Neutralization sites downstream of the embolded 'RGD' cell-attachment site previously identified by mAb studies (Crowther *et al.*, 1993) are underlined.

The steady rate of nucleotide substitution in the phenotypically plastic capsid genes of picornaviruses has been noted many times previously (e.g. Martinez *et al.*, 1992; Vosloo *et al.*, 1996; Zhang *et al.*, 1999; Haydon *et al.*, 2001b; Jenkins *et al.*, 2002). Usually the rate of substitution can only be estimated within monophyletic groups rooted with a virus strain isolated close to the date of the root of the tree. The approach implemented in TIPDATE uses an LRT in the same spirit that Felsenstein (1993) proposed a statistical test to distinguish between ultrametric clock-like trees and non-ultrametric trees. Felsenstein's original approach could not be validly applied to RNA virus phylogenies of strains isolated at different times as the rate of evolution is sufficiently high that even clock-like modes of substitution could not be expected to yield ultrametric phylogenies. It is not unusual to observe that the rates of estimated nucleotide substitution within the 1D genes become progressively less as they are measured over longer time-scales. For example, measured over just a few years, annual rates of between 0·5 and 1·5 % of sites incurring substitutions are not uncommon (Vosloo *et al.*, 1996; Haydon *et al.*, 2001b); however this rate cannot apply over longer time-periods and in the more clock-like clades of the phylogeny reconstructed here, rates were always estimated to be between 0·02 and 0·5 % per year. The determinants of longer term rates of evolution in RNA genomes remains a contentious subject (see for example Korber *et al.*, 1998). Interestingly, the highest rates of nucleotide substitution were observed in virus groups recovered from cattle (topotypes B, E and F) and the lowest rates observed in clades containing virus recovered from wildlife (topotypes A, C and D). The differences in rates between topotypes may reflect the different levels of epidemiological activity associated with different topotypes (with higher numbers of cases resulting in more virus replication and more opportunity for divergence). However, the sample sizes within each topotype do not permit these rates to be distinguished statistically.

The average ratio of synonymous to non-synonymous substitution across the whole VP1 gene was 5·0 : 1. Low ratios of synonymous to non-synonymous substitutions (less than one) measured across parts of the gene encoding viral epitopes has been interpreted as evidence of positive

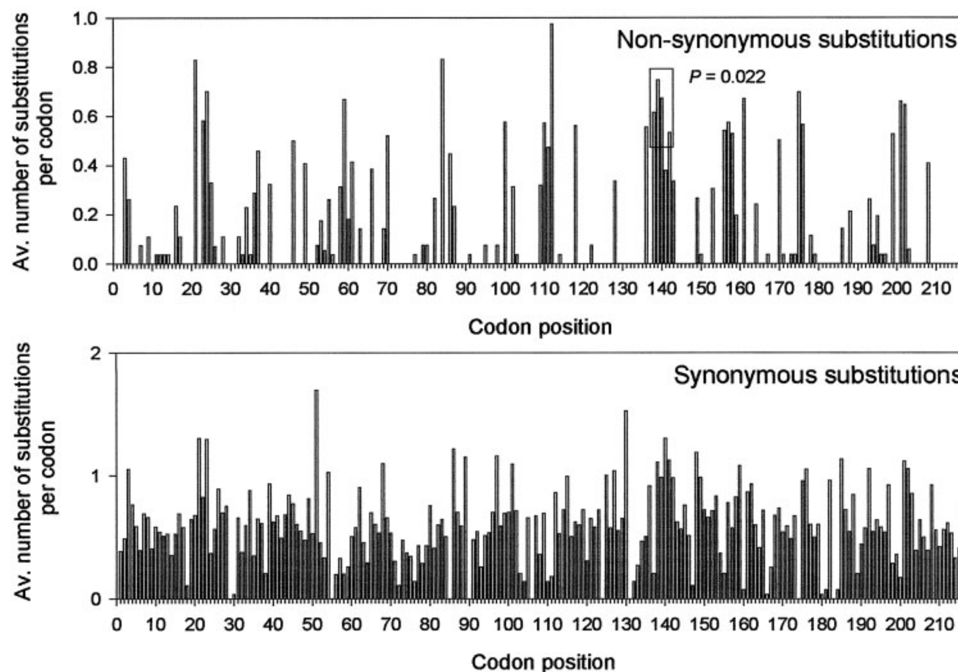

Fig. 4. Distribution of nucleotide substitutions along the VP1 genes sequenced. The *y*-axis indicates the number of synonymous and non-synonymous nucleotide substitutions per codon position averaged over every possible pairwise sequence comparison.

selection (Fares *et al.*, 2001; Haydon *et al.*, 2001a), but this ratio is expected to rise rapidly as the more genetically constrained parts of the gene are included to estimate the average ratio. If virus infections induce a more vigorous immune response in cattle than in buffalo and impala (which may be the case given the difference in disease severity that infection produces), then we might expect to observe higher levels of positive selection and lower ratios of synonymous to non-synonymous substitutions in cattle than in wild hosts. However, we found that when we fitted a model that permitted different ratios of synonymous to non-synonymous substitutions to clades that contained only cattle (see Fig. 1) no significant differences in selective pressure were indicated (results not shown). Synonymous substitutions occur evenly between codons along the gene, showing no sign of clustering, and thus do not indicate any evidence for patterns of larger scale constraints on synonymous substitutions. Surprisingly perhaps, non-synonymous substitutions also showed little evidence of spatial clustering along the gene, with just one weakly significant hotspot apparent around codon 140.

This study has provided some indication of the level of genetic diversity of SAT 2-type FMD in individual countries on the African continent. Of the 17 African countries included in this study, 13 have a single SAT 2 topotype within their borders. However within some countries, notably Botswana, Zimbabwe and Sénégal, there are at least two topotypes present. This translates to within-country genetic variation levels of between 24 and 27 % on nucleotide level and between 16·6 and 19·8 % on amino acid level, when pairwise comparisons are performed for two viruses from the same country but from different topotypes. By far the most heterogeneous is Kenya, which has viruses representative of two of the six major clusters (I and IV) and has the highest within-country levels of genetic diversity. Kenyan strains differed from each other by more than 30 % at nucleotide level and by more than 23·5 % at amino acid level. These within-country levels approximate the maximum observed for the continent where sequence divergence is 31·6 % and 24·1 % at nucleotide and amino acid levels, respectively.

The low level of nucleotide sequence conservation (only 38·4 % of sites are completely conserved across all SAT 2 viruses analysed in this study) has diagnostic implications, as the development of a single set of serotype-specific primers that will be suitable for detection of all SAT 2 field variants is unlikely. Through the generation of complete VP1 gene sequences detailed here, it may now be possible to design efficient serotype-specific primers, something that has eluded most researchers (Reid *et al.*, 2001). However, given high levels of intratypic variation it is unlikely that a single primer set will permit unequivocal identification of all SAT 2 field strains and regional-specific primer sets may be more appropriate.

These high levels of genetic diversity within the SAT 2 serotype are likely to be reflected antigenically and have

implications for control of the disease through vaccination. To be effective, vaccines should not only be stable and have a high antigen load, but should also be immunologically related to viruses circulating in the field at the time (Hunter, 1998). The r_1 value provides a measure of antigenic similarity and is defined as the ratio of the serum titre of heterologous virus (e.g. an outbreak strain) to the serum titre of homologous virus (i.e. the vaccine strain) (Rweyemamu *et al.*, 1977). Serological comparisons of SAT 2 viruses from Kenya revealed that there is extensive antigenic variation in this country, where two of the six major SAT 2 virus clusters occur (Ndiritu *et al.*, 1983). In a comparative study of the antigenic and genetic diversity within SAT 1 (Hunter *et al.*, 1996), it was shown that viruses from different topotypes, i.e. those that differ from each other by more than 20 % on nucleotide sequence level (Bastos *et al.*, 2001), generally have r_1 values below 0·4 and are therefore antigenically poorly related (Esterhuysen, 1994; Hunter *et al.*, 1996). Similar results have been obtained for serotype A viruses where viruses differing from each other by $\geqslant$15 % and therefore of different topotypes (Knowles & Samuel, 2003) were shown to be antigenically poorly related (Araújo *et al.*, 2002). Clearly, the identification of eleven discernable SAT 2 topotypes indicates that multiple region-specific vaccines may have to be developed to address this field heterogeneity. Given the regional diversity within the SAT 2 serotype and that documented for four other serotypes (Knowles *et al.*, 1998; Sangaré *et al.*, 2001; Bastos, 2001; Bastos *et al.*, 2001), controlling the disease through vaccination will be extremely difficult. Furthermore, the costs arising from the administration of multivalent vaccines on a systematic basis places this means of disease control out of reach of most developing countries in sub-Saharan Africa. The twin problems presented by virus diversity and difficulty of establishing effective movement control in western, central and eastern Africa present an apparently insuperable difficulty for effective FMD control.

ACKNOWLEDGEMENTS

We thank N. Ferris, B. Newman and J. J. Esterhuysen for providing information on isolates used in this study. The FMD World Reference Lab (Pirbright, UK) and veterinary personnel from different African laboratories are thanked for providing isolates and samples. D. T. H. was supported by the Wellcome Trust and O. S. by the Winrock Foundation.

REFERENCES

Acharya, A., Fry, E., Stuart, D., Fox, G., Rowlands, D. & Brown, F. (1990). The structure of foot-and-mouth disease virus: implications for its physical and biological properties. *Vet Microbiol* **23**, 21–34.

Araújo, J. P., Jr, Montassier, H. J. & Pinto, A. A. (2002). Extensive antigenic and genetic variation among foot-and-mouth disease type A viruses isolated from the 1994 and 1995 foci in São Paulo, Brazil. *Vet Microbiol* **84**, 15–27.

Bastos, A. D. S. (1998). Detection and characterization of foot-and-mouth disease virus in sub-Saharan Africa. *Onderstepoort J Vet Res* **65**, 37–47.

Bastos, A. D. S. (2001). *Molecular epidemiology and diagnosis of SAT type foot-and-mouth disease in southern Africa*. PhD thesis, University of Pretoria, South Africa.

Bastos, A. D. S., Boshoff, C. I., Keet, D. F., Bengis, R. G. & Thomson, G. R. (2000). Natural transmission of foot-and-mouth disease virus between African buffalo (*Syncerus caffer*) and impala (*Aepyceros melampus*) in the Kruger National Park, South Africa. *Epidemiol Infect* **124**, 591–598.

Bastos, A. D. S, Haydon, D. T., Forsberg, R., Knowles, N. J., Anderson, E. C., Bengis, R. G., Nel, L. H. & Thomson, G. R. (2001). Genetic heterogeneity of SAT-1 type foot-and-mouth disease viruses in southern Africa. *Arch Virol* **146**, 1537–1551.

Beck, E. & Strohmaier, K. (1987). Subtyping of European FMDV outbreak strains by nucleotide sequence determination. *J Virol* **61**, 1621–1629.

Boom, R., Sol, C. J., Salimans, M. M. M., Jansen, C. L., Wertheim-van Dillen, P. M. E. & van der Noordaa, J. (1990). Rapid and simple method for purification of nucleic acids. *J Clin Microbiol* **28**, 495–503.

Brooksby, J. B. (1972). Epizootiology of foot-and-mouth disease in developing countries. *World Anim Rev* **3**, 10–13.

Condy, J. B., Hedger, R. S., Hamblin, C. & Barnett, I. T. R. (1985). The duration of the foot-and-mouth disease carrier state in African buffalo (i) in the individual animal and (ii) in a free-living herd. *Comp Immunol Microbiol Infect Dis* **8**, 259–265.

Crowther, J. R., Rowe, C. A. & Butcher, R. (1993). Characterization of monoclonal antibodies against a type SAT-2 foot-and-mouth disease virus. *Epidemiol Infect* **111**, 391–406.

Dawe, P. S., Flanagan, F. O., Madekurozwa, R. L., Sorenson, K. J., Anderson, E. C., Foggin, C. M., Ferris, N. P. & Knowles, N. J. (1994a). Natural transmission of foot-and-mouth disease from African buffalo (*Syncerus caffer*) to cattle in a wildlife area of Zimbabwe. *Vet Rec* **134**, 230–232.

Dawe, P. S., Sorensen, K., Ferris, N. P., Barnett, I. T. R., Armstrong, R. M. & Knowles, N. J. (1994b). Experimental transmission of foot-and-mouth disease virus from carrier African buffalo (*Syncerus caffer*) to cattle in Zimbabwe. *Vet Rec* **134**, 211–215.

Esterhuysen, J. J. (1994). *The antigenic variation of foot-and-mouth disease viruses and its significance in the epidemiology of the disease in southern Africa*. MSc dissertation. University of Pretoria, South Africa.

Fares, M. A., Moya, A., Escarmis, C., Baranowski, E., Domingo, E. & Barrio, E. (2001). Evidence for positive selection in the capsid protein-coding region of the foot-and-mouth disease virus (FMDV) subjected to experimental passage regimens. *Mol Biol Evol* **18**, 10–21.

Felsenstein, J. (1993). PHYLIP (phylogeny inference package), version 3.5c. Department of Genetics, University of Washington, Seattle.

Ferris, N. P. & Donaldson, A. I. (1992). The World Reference Laboratory for foot-and-mouth disease: a review of thirty-three years of activity (1958–1991). *Rev Sci Tech* **11**, 657–684.

Grassly, N. C. & Holmes, E. C. (1997). A likelihood method for the detection of selection and recombination using nucleotide sequences. *Mol Biol Evol* **14**, 239–247.

Harley, E. H. (2001). DAPSA. *DNA and protein sequence analysis*, version 4.91, Department of Chemical Pathology, University of Cape Town, South Africa.

Hasegawa, M., Kishino, H. & Yano, T. (1985). Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* **21**, 160–174.

Haydon, D. T., Bastos, A. D., Knowles, N. J. & Samuel, A. R. (2001a). Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* **137**, 7–15.

Haydon, D. T., Samuel, A. R. & Knowles, N. J. (2001b). The generation and persistence of genetic variation in foot-and-mouth disease virus. *Prev Vet Med* **51**, 111–124.

Hunter, P. (1998). Vaccination as a means of control of foot-and-mouth disease in sub-Saharan Africa. *Vaccine* **16**, 261–264.

Hunter, P., Bastos, A. D. S., Esterhuysen, J. J. & van Vuuren, C. de W. J. (1996). Appropriate foot-and-mouth disease vaccines for southern Africa. *All Africa Conference on Animal Agriculture, Pretoria, South Africa* 2.2.7, 1–4.

Jenkins, G. M., Rambaut, A., Pybus, O. G. & Holmes, E. C. (2002). Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J Mol Evol* **54**, 156–165.

Keet, D. F., Hunter, P., Bengis, R. G., Bastos, A. D. & Thomson, G. R. (1996). The 1992 foot-and-mouth disease epizootic in the Kruger National Park. *J South African Vet Assoc* **67**, 83–87.

Kitching, R. P. (1998). A recent history of foot-and-mouth disease. *J Comp Pathol* **118**, 89–108.

Knowles, N. J. & Samuel, A. R. (2003). Molecular epidemiology of foot-and-mouth disease virus. *Virus Res* **91**, 65–80.

Knowles, N. J., Ansell, D. M. & Samuel, A. R. (1998). Molecular comparison of recent foot-and-mouth disease type A viruses from West Africa with historical and reference strains. Paper presented to the session of the research group of the standing technical committee of the European Commission for the control of foot-and-mouth disease, Pirbright, UK, 14–18 September, pp. 1–6.

Korber, B., Theiler, J. & Wolinsky, S. (1998). Limitations of a molecular clock applied to considerations of the origin of HIV-1. *Science* **280**, 1868–1871.

Martinez, M. A., Dopazo, J., Hernandez, J., Mateu, M. G., Sobrino, F., Domingo, E. & Knowles, N. J. (1992). Evolution of capsid protein genes of foot-and-mouth disease virus: antigenic variation without accumulation of amino acid substitutions over six decades. *J Virol* **66**, 3557–3565.

Mateu, M. G., Valero, M. L., Andreu, D. & Domingo, E. (1996). Systematic replacement of amino acid residues within an Arg-Gly-Asp-containing loop of foot-and-mouth disease virus and effect on cell recognition. *J Biol Chem* **271**, 12814–12819.

Mweene, A. S., Pandey, G. S., Sinyangwe, P., Nambota, A., Samui, K. & Kida, H. (1996). Viral diseases of livestock in Zambia. *Jpn J Vet Res* **44**, 89–105.

Ndiritu, C. G. (1984). Foot and mouth disease virus antigenic variation and its implication on vaccines. *Kenya Vet* **8**, 14–19.

Ndiritu, C. G., Ouldridge, E. J., Head, M. & Rweyemamu, M. M. (1983). A serological evaluation of 1979–1982 Kenyan foot-and-mouth disease type SAT 2 viruses. *J Hyg* **91**, 335–341.

Rambaut, A. (2000). Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics* **16**, 395–399.

Reid, S. M., Ferris, N. P., Hutchings, G. H., De Clercq, K., Newman, B. J., Knowles, N. J. & Samuel, A. R. (2001). Diagnosis of foot-and-mouth disease by RT-PCR: use of phylogenetic data to evaluate primers for the typing of viral RNA in clinical samples. *Arch Virol* **146**, 2421–2434.

Rozas, J. & Rozas, R. (1999). DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**, 174–175.

Rweyemamu, M. M., Pay, T. W. F. & Parker, M. J. (1977). Serological differentiation of foot-and-mouth disease virus strains in relation to selection of suitable vaccine viruses. *Dev Biol Stand* **35**, 205–214.

Samuel, A. R. & Knowles, N. J. (2001a). Foot-and-mouth disease type O viruses exhibit genetically and geographically distinct evolutionary lineages (topotypes). *J Gen Virol* **82**, 609–621.

Samuel, A. R. & Knowles, N. J. (2001b). Foot-and-mouth disease virus: cause of the recent crisis for the UK livestock industry. *Trends Genet* **17**, 421–424.

Sangaré, O. (2002). *Molecular epidemiology of foot-and-mouth disease virus in West Africa*. PhD thesis, University of Pretoria, Pretoria, South Africa.

Sangaré, O., Bastos, A. D. S., Marquardt, O., Venter, E. H., Vosloo, W. & Thomson, G. R. (2001). Molecular epidemiology of serotype O foot-and-mouth disease virus with emphasis on West and South Africa. *Virus Genes* **22**, 345–351.

Sobrino, F., Saíz, M., Jiménez-Clavero Núñez, J. I., Rosas, M. F., Baranowski, E. & Ley, V. (2001). Foot-and-mouth disease virus: a long known virus, but current threat. *Vet Res* **32**, 1–30.

Swofford, D. L. (2001). PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.0b8. Sunderland, MA: Sinauer Associates.

Thomson, G. R. (1994). Foot-and-mouth disease. In *Infectious Diseases of Livestock with Special Reference to Southern Africa*, pp. 825–952. Edited by J. A. W. Coetzer, G. R. Thomson & R. C. Tustin. Cape Town: Oxford University Press.

Thomson, G. R. (1996). The role of carrier animals in the transmission of foot and mouth disease. *O.I.E. Comprehensive Reports on Technical Items Presented to the International Committee or to Regional Commissions*, pp. 87–103.

van Rensburg, H. G. & Nel, L. H. (1999). Characterization of the structural-protein-coding region of SAT2 type foot-and-mouth disease virus. *Virus Genes* **19**, 229–233.

Vosloo, W., Kirkbride, E., Bengis, R. G., Keet, D. F. & Thomson, G. R. (1995). Genome variation in the SAT types of foot-and-mouth disease viruses prevalent in buffalo (*Syncerus caffer*) in the Kruger National Park and other regions of southern Africa, 1986–1993. *Epidemiol Infect* **114**, 203–218.

Vosloo, W., Bastos, A. D., Kirkbride, E., Esterhuysen, J. J., Janse van Rensburg, D., Bengis, R. G., Keet, D. F. & Thomson, G. R. (1996). Persistent infection of African buffalo (*Syncerus caffer*) with SAT-type foot-and-mouth disease viruses: rate of fixation of mutations, antigenic change and interspecies transmission. *J Gen Virol* **77**, 1457–1467.

Vosloo, W., Bastos, A. D. S., Michel, A. & Thomson, G. R. (2001). Tracing movement of African buffalo in southern Africa. *Rev Sci Tech* **20**, 630–639.

Xie, Q.-C., McCahon, D., Crowther, J. R., Belsham, G. J. & McCullough, K. C. (1987). Neutralization of foot-and-mouth disease virus can be mediated through any of at least three separate antigenic sites. *J Gen Virol* **68**, 1637–1647.

Zhang, G., Haydon, D. T., Knowles, N. J. & McCauley, J. W. (1999). Molecular evolution of swine vesicular disease virus. *J Gen Virol* **80**, 639–651.