Chapter 4

Natural transmission of foot-and-mouth disease virus between African buffalo (*Syncerus caffer*) and impala (*Aepyceros melampus*) in the Kruger National Park, South Africa

Summary

VP1 gene sequences of SAT-2 type foot-and-mouth disease (FMD) virus isolated from impala and African buffalo in the Kruger National Park (KNP), were used to determine intra- and interspecies relationships of viruses circulating in these wildlife populations. In this way, five distinct lineages of SAT-2 virus were identified in routine sampling of oesophageo-pharyngeal epithelium from buffalo between 1988 and 1996. The different lineages were each found to be associated with a discrete geographic sampling locality. The results presented here clearly show that over the period 1985-1995, four unrelated epizootics occurred in impala in defined localities within the KNP. Evidence for natural transmission of FMD virus between buffalo and impala is presented for the most recent 1995 outbreak, with data linking the 1985 and 1988/9 impala epizootics to viruses associated with specific buffalo herds. Circumstantial evidence indicates that juvenile buffalo infected for the first time are the most likely source of infection for impala and that conditions driving these species to share habitats are likely to facilitate inter-species transmission.

Parts of the results presented here have been published in *Epidemiology and Infection*, **124**: 591-598 (2000)

4.1 Introduction

The Kruger National Park (KNP) is situated in the north eastern corner of South Africa (between 22°25' and 25°32' latitude south and 30°50' and 32°02' longitude east) and is bordered by Zimbabwe and Mozambique, to the north and east respectively. The KNP and surrounding area is the only region in South Africa where foot-and-mouth disease (FMD) occurs, with the remainder of the country being recognized as free of the disease by the O.I.E. (Office International des Epizooties). Since the last recorded outbreak of FMD in livestock in South Africa in 1983, the disease has been restricted to wildlife within the FMD control area (Records of the Directorate of Animal Health). Clinical disease has primarily involved impala (*Aepyceros melampus*), a medium sized antelope, but it is well recognized that infection with SAT-type viruses is endemic to African buffalo (*Syncerus caffer*) herds within the KNP.

Of the eighteen southern African wildlife species found to have antibodies to FMD (Condy *et al.* 1969; Keet *et al.* 1996), only the African buffalo has been shown to be an efficient maintenance host of the virus (Condy *et al.* 1985) and is considered the primary source of infection for domestic and wild ungulates. This is due to its ability to both maintain and transmit the disease (Thomson 1996), with transmission from carrier buffalo to cattle being demonstrated under both experimental and natural conditions (Hedger & Condy 1985; Dawe *et al.* 1994a; Dawe *et al.* 1994b;Vosloo *et al.* 1996). In contrast, evidence of field transmission between buffalo and other cloven-hoofed wildlife species is lacking despite the co-occurrence of buffalo and a variety of antelope in game parks throughout southern Africa and the known susceptibility of some antelope species to FMD virus (FMDV) infection (Hedger *et al.* 1972; Anderson *et al.* 1975).

In South Africa, the diagnosis of clinical FMD in impala has been regularly recorded since 1944 (Rossiter & Albertyn 1947; Meeser 1962; Vosloo *et al.* 1992; Keet *et al.* 1996). Over 90% of these outbreaks have occurred in the months of June to November within the southern region of the KNP. This area, south of the Olifants river (Fig. 1) coincides with the region of the KNP that is most densely populated by both impala and buffalo (Pienaar *et al.* 1966; Pienaar 1974). In addition, the months of June to August are generally rain-free and therefore a period of water scarcity, which leads to animals congregating at permanent water points (Keet *et al.* 1996). This is believed to facilitate disease transmission and is supported by the observation that some outbreaks in impala have spread along water courses (Vosloo *et al.* 1992). In addition to shared drinking water, mutual grazing of at least three grass species by impala and buffalo (Skinner & Smithers 1990) is another factor promoting selection of common environs.

Despite the high incidence of outbreaks in impala (Keet *et al.* 1996) and the active collection and characterization of viruses obtained from buffalo in recent years (Vosloo *et al.* 1995), direct evidence of natural transmission between these species is lacking. VP1 gene nucleotide sequence analysis has, however, been instrumental in determining the apparent unrelatedness of successive epizootics in impala (Vosloo *et al.* 1992; Keet *et al.* 1996). In a study of FMD outbreaks in impala from 1983 to 1989, Vosloo and co-workers (1992) showed that viruses isolated 6 months apart were derived from the same epizootic. A later study focussing on the 1992/3 epizootic in the KNP identified a potential circulation period of 12.8 months (Keet *et al.* 1996). Although these results infer the possibility of viral persistence in impala populations, other studies have shown experimentally, that infected impala develop clinical lesions and excrete virus, but that most do not become carriers (Hedger *et al.* 1972). This was confirmed in a separate study which showed that virus could not be isolated from any tissues of infected impala, 7 or more days after inoculation (Anderson *et al.* 1975). Given the low antibody levels of impala during inter-epizootic periods and their apparent inability to act as long term carriers, it is unlikely that these antelope are capable of maintaining the disease independently and therefore African buffalo are presumed to be the usual source of infection for impala (Keet *et al.* 1996).

In order to determine whether buffalo are indeed the source of infection for impala, a twelve year study period (1985 to 1996) was identified in which clinical FMD was diagnosed on six occasions in impala in the KNP (Fig. 4.1). The availability of impala isolates from these SAT-2 epizootics and buffalo field strains collected within the same time period, together with improved SAT-type genetic characterization methods (Bastos 1998a), prompted an investigation to determine whether inter-species transmission could be demonstrated.

4.2 Material & Methods

4.2.1 Viruses used in this study

Impala viruses used in this study originated from the southern and central region of the Kruger National Park, South Africa (Fig. 4. 1). Impala viruses were isolated on pig kidney cells (PK) from clinical FMD lesions recovered during the epizootics of 1988, 1989, 1992, 1993 and 1995. Tissue culture isolations of an impala virus from the 1985 outbreak was supplied by the World Reference Laboratory, Pirbright. Oesophageo-pharyngeal specimens (probangs) obtained from buffalo during routine surveys in the KNP from 1988 to 1996 (Fig. 4.2), were used as a source of viruses circulating in buffalo. The geographical and species origin of all viruses included in this study are summarized in Table 4.1.

TABLE 4.1: List of SAT-2 viruses of buffalo and impala origin (1985-1996) from the Kruger NP

Strain	Sampling locality	Grid Reference	Species of origin	Sampling date	Genbank Acc. No
KNP 1/85	Gudzani	31°50'E 24°15'S	Impala	21 NOV 1985	AF136986
KNP 7/88	Rietpan	31°58'E-24°54'S	Buffalo	8 JUL 1988	AF137000
KNP 8/88	Rietpan	31°58'E-24°54'S	Buffalo	8 JUL1988	AF137001
KNP 9/88	Shilolweni	31°50'E-24°48'S	Buffalo	27 MAY 1988	AF137002
KNP 10/88	Shilolweni	31°50'E-24°48'S	Buffalo	27 MAY 1988	AF137003
KNP 14/88	Ripape	31°37'E-24°44'S	Buffalo	19 OCT 1988	AF137004
KNP 16/88	Kingfisherspruit	31°26'E-24°27'S	Impala	10 OCT 1988	AF136987
KNP 17/88	Orpen	31°24'E-24°28'S	Impala	7 OCT 1988	AF136988
KNP 19/88	Rabelais Dam	31°30'E-24°27'S	Impala	11 OCT 1988	AF136989
KNP 20/88	Timbavati River	31°28'E-24°26'S	Impala	25 OCT 1988	AF136990
KNP 2/89	Ngotso	31°43'E-24°13'S	Impala	30 APR 1989	AF136991
KNP 19/89	Ripape	31°37'E-24°44'S	Buffalo	25 OCT 1989	AF137005
KNP 25/89	Matjipiri	31°36'E-24°47'S	Buffalo	24 OCT 1989	AF137006
KNP 40/89	Ripape	31°37'E-24°44'S	Buffalo	25 OCT 1989	AF137007
KNP 5/91	Satara	31°47'E-24°24'S	Buffalo	20 JUN 1991	AF137008
KNP 143/91	Reënvoël Dam	31°20'E-23°58'S	Buffalo	13 JUL 1991	AF137009
KNP 147/91	Reënvoël Dam	31°20'E-23°58'S	Buffalo	13 JUL 1991	AF137010
KNP 160/91	Ndziyospruit	31°34'E-23°59'S	Buffalo	18 JUL 1991	AF137011
KNP 183/91	Water Affairs Weir	31°56'E-25°08'S	Buffalo	24 JUL 1991	AF137012
KNP 1/92	Nsemani Dam	31°43'E-24°23'S	Impala	3 AUG 1992	AF136992
KNP 5/92	Timbavati	31°38'E-24°16'S	Impala	3 AUG 1992	AF136993
KNP 8/92	Shibotwana Dam	31°50'E-24°24'S	Impala	3 AUG 1992	AF136994
KNP 32/92	Boyela Vlakteplaas	31°17'E-22°54'S	Buffalo	22 JUN 1992	AF137013
KNP 9/93	Boyelaspruit	31°20'E-23°01'S	Buffalo	17 AUG 1992	AF137014
KNP 16/93	Capricorn	31°26'E-23°29'S	Buffalo	6 AUG 1992	AF137015
KNP 51/93	Tshokwane	31°51'E-24°47'S	Impala	26 AUG 1993	AF136995
KNP 18/95	Monzweni	31°38'E-24°34'S	Buffalo	16 NOV 1995	AF137016
KNP 31/95	Monzweni	31°38'E-24°34'S	Buffalo	16 NOV 1995	AF137017
KNP 43/95	Mbyamiti Mouth	31°46'E-25°18'S	Impala	29 NOV 1995	AF136996
KNP 44/95	Lwakahle Picket	31°41'E-25°23'S	Impala	2 DEC 1995	AF136997
KNP 49/95	Mbyamiti Mouth	31°46'E-25°18'S	Impala	2 DEC 1995	AF136998
KNP 52/95	Lwakahle Picket	31°41'E-25°23'S	Impala	2 DEC 1995	AF136999
KNP 6/96	Mahlanganzwane Dam	32°01'E-25°14'S	Buffalo	17 APR 1996	AF137018
KNP 24/96	Mulalanespruit	31°18'E-23°57'S	Buffalo	29 JUN 1996	AF137019

Fig. 4.1: Geographical origin of impala viruses isolated between 1985 and 1995

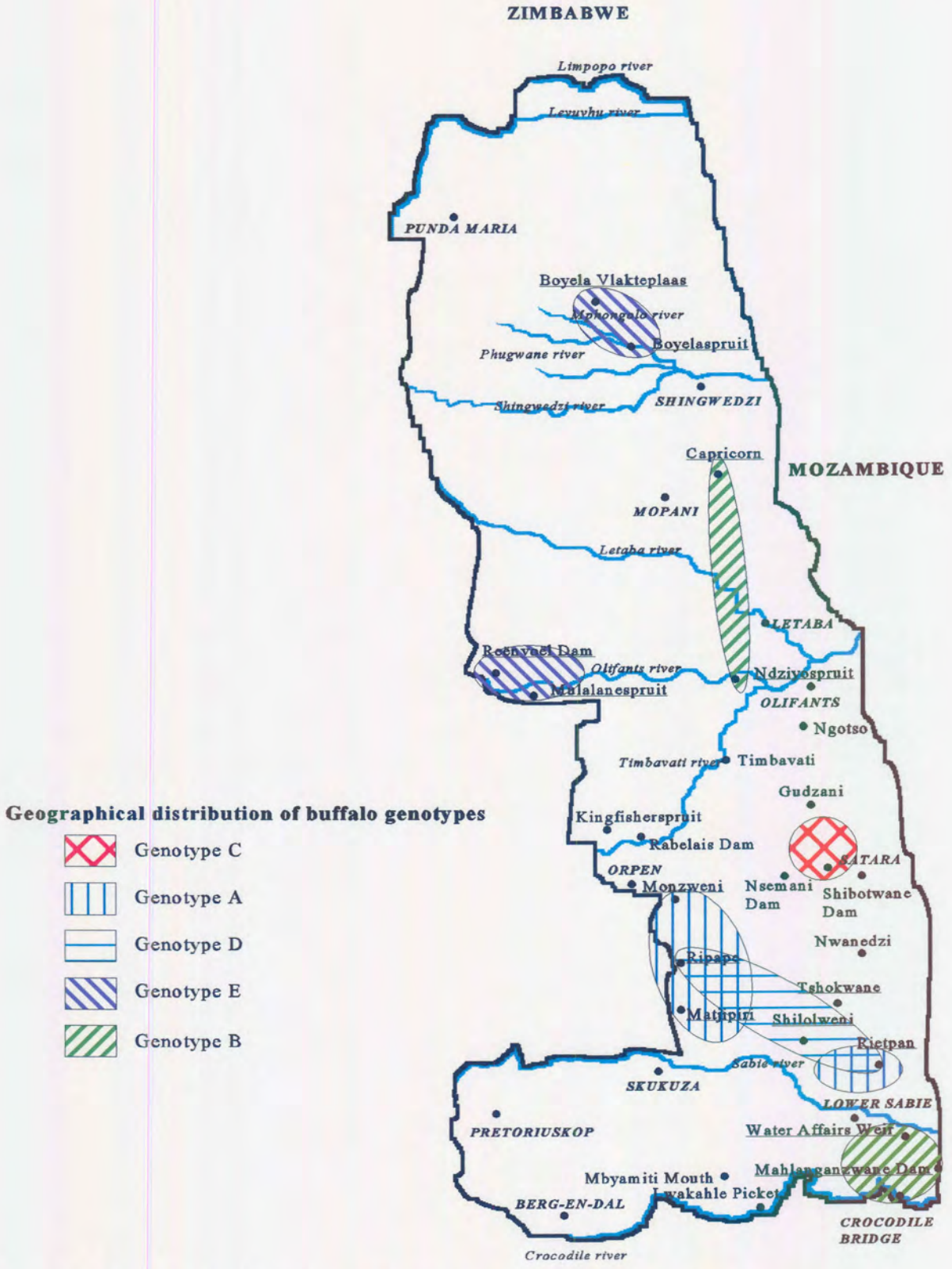

Fig. 4.2: Georgraphical origin and distribution of SAT-2 type buffalo viruses

4.2.2 Genetic characterization

Viral RNA was extracted from clinical specimens and tissue culture isolates of buffalo and impala viruses by a modified guanidinium thyocianate /silica method (Boom *et al.* 1990). A 500 bp fragment corresponding to the C-terminus region of the VP1 gene was amplified, purified and sequenced as described in Chapters 2 and 3. Nucleotide sequences were aligned (Harley 1994) and an homologous 414 nt sequence was identified and selected for further analyses. Pairwise comparisons of amino acid and nucleotide sequences of impala outbreaks strains and field strains of carrier buffalo were determined using the MEGA programme (Kumar *et al.* 1993). Various distance and parsimony methods included in MEGA were also used to construct VP1 gene trees. The cattle outbreak stain, ZIM/7/83 from Zimbabwe (Genbank accession No: AF023523) was selected as an outgroup for the phylogenetic analyses, due to it's distant geographic and genetic relationship to KNP viruses (Bastos 1998a). All nucleotide sequences used in this study have been submitted to Genbank under the accession numbers indicated in Table 4.1.

4.3 Results

4.3.1 Genetic relationships of impala viruses

Four genetically unrelated impala virus lineages were identified by VP1 gene sequence analysis (Fig. 4.3), each from a distinct geographical area indicated in the figure as follows:

(I) Gudzani Area (1985 epizootic)
(II) Orpen Area (1988/9 epizootic)
(III) Satara/Tshokwane Area (1992/3 epizootic)
(IV) Crocodile Bridge Area (1995 epizootic)

Pairwise comparisons of nucleotide and amino acid sequence data reveal that viruses of the same epizootic, sampled within a month of each other are identical or differ by less than 1 %. The 1988/9 and 1992/3 viruses, show marked differences on amino acid level over time. In the 6 months separating the 1989 virus from the 1988 viruses, there was a 1-2 % accumulation in mutations, whilst the 1993 impala virus differed by between 3 and 4 % from the 1992 viruses within a sampling time span of more than 12 months. In the most recent SAT-2 epizootic, the 1995 impala viruses differ from each other by 1 - 4 % on amino acid level, despite being sampled within a five day period. Nucleotide sequence identity (completely conserved sites) for the 14 impala viruses used in this study was 70.5 % across the 414 nt specified in this study. Amino acid sequence identity for the same region was 80.4 %.

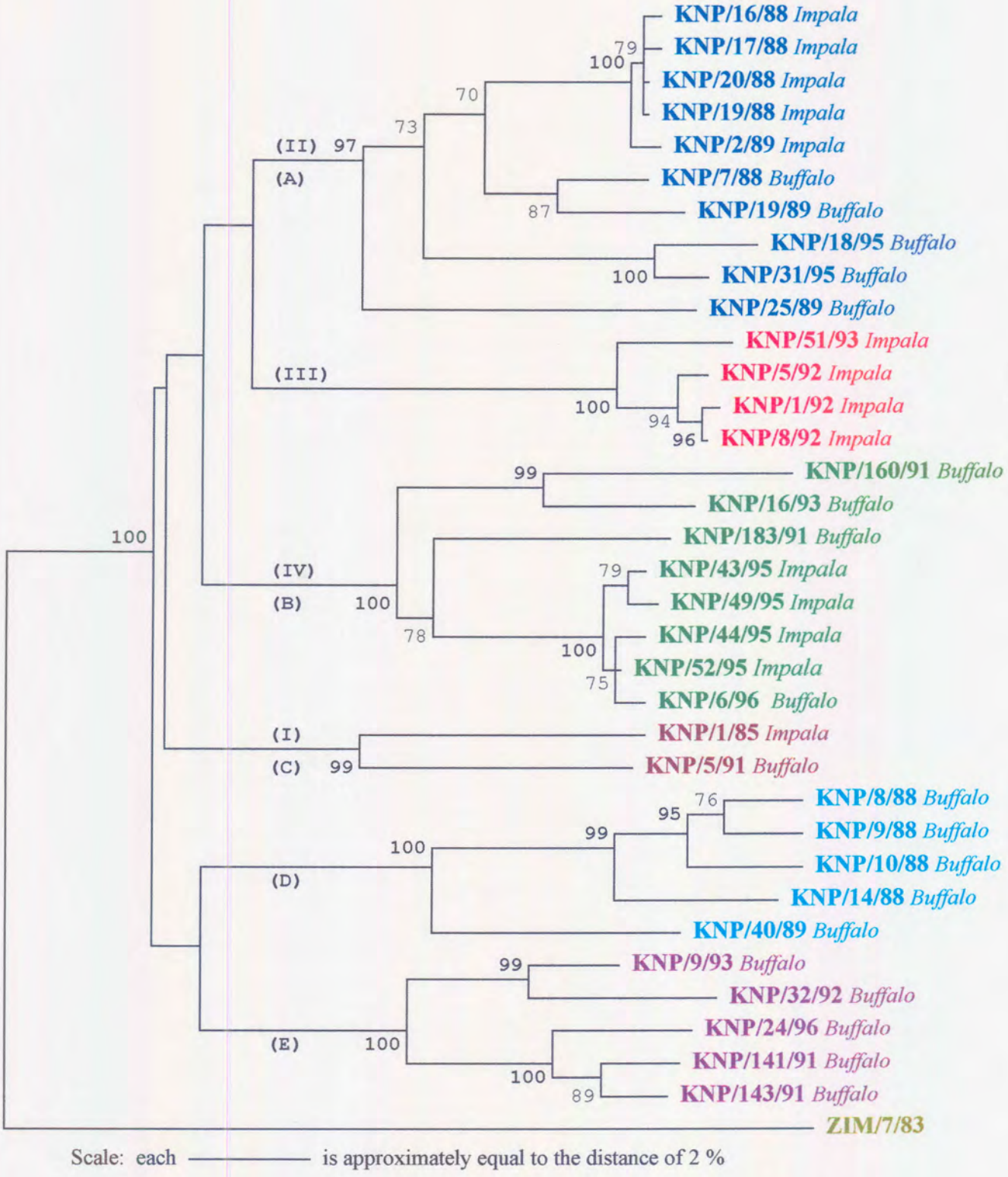

Fig. 4.3: Neighbor-joining tree depicting VP1 gene relationships of buffalo and impala viruses from the Kruger National Park (1985-1996). Bootstrap values indicated as a percentage were determined by 1000 replications.

4.3.2 Genetic relationships of buffalo viruses

VP1 gene trees of SAT-2 type FMDV were used to identify the major buffalo virus lineages within the Kruger National Park. Both distance methods (UPGMA and neighbor-joining) and character-based phylogenetic inference methods produced trees with identical topology (results not shown). The neighbor-joining tree depicted in Fig. 4.3 indicates that there are five distinct buffalo virus lineages (coded A-E) south of the Olifants river (Fig. 4.2), which cluster according to geographical origin in the following manner:

(A) Monzweni (1995), Ripape (1989), Rietpan (1988) and Matjipiri (1989) viruses
(B) Water Affairs Weir (1991) and Mahlanganzwane Dam (1996) viruses
(C) Satara (1991) virus
(D) Rietpan (1988), Ripape (1988 & 1989) & Shilolweni (1988) viruses
(E) Mulalanespruit (1996) virus

The results (Fig. 4.3) also show cases of co-circulation of two different SAT-2 genotypes within a single buffalo herd. This is demonstrated by the grouping of KNP/7/88 and KNP/8/88 with the 'A' and 'D' genotypes respectively, despite having been sampled on the same day and from the same herd. Similarly, the two Ripape viruses KNP/19/89 and KNP/40/89 sampled from the same herd on the 25th of October 1989, fall into two separate clusters. Furthermore, the geographical distribution of buffalo genotypes (Fig. 4.2) reveals that there are shared genotypes across the Olifants river. This is illustrated by the clustering of viruses from the Crocodile Bridge area (KNP/183/91 and KNP/6/96), which is in the southern part of the KNP, with the Capricorn and Ndziyospruit viruses (KNP/160/91 and KNP/16/93) to the north of the Olifants river. In addition, two Reënvoël Dam viruses (KNP/143/91 and KNP/147/91) sampled north of the Olifants river, group with a Mulalanespruit virus (KNP/24/96), sampled south of the river. These form part of a larger cluster (E) containing Boyela viruses (KNP/32/92 and KNP/9/93) sampled in the far north. The five buffalo virus clusters identified in Fig. 4.3 confirm the spacial rather than temporal grouping of FMD viruses, as exemplified by Cluster A (1988-1995) and Cluster B (1991-1996). Overall sequence identity (sites complete conserved across all sequences) for the 20 buffalo viruses used in this study was 59.9 % across the 414 nt regions specified in this study. Amino acid sequence identity for the same region was 74.6 %.

4.3.3 Inter-species relationships of impala and buffalo viruses

The combined buffalo and impala data-set was used to determine inter-species relationships of viruses (Fig. 4.3). Statistically supported phylogenetic clustering for three of the four impala epizootics with specific buffalo lineages, is clearly demonstrated. The

1985 impala strain groups with buffalo virus KNP 5/91 (99% bootstrap support) and the 1988/89 outbreak strains cluster with 1988, 1989 and 1995 buffalo viruses from Rietpan, Ripape and Mondzweni (97% bootstrap support). The 1995 impala viruses from the Crocodile Bridge region are almost identical to a 1996 buffalo virus obtained 4 months later from the same locality (100% bootstrap support). The impala epizootic of 1992/93 did not group with any of the buffalo viruses sampled between 1985 and 1996. Overall nucleotide sequence identity for the 34 Kruger National Park viruses used in this study was 57.2 % across the 414 nt region specified, with amino acid sequence identity being 69.6 %.

4.4 Discussion

The nucleotide sequence data from this study indicate that four unrelated SAT-2 epizootics occurred in impala the KNP from 1985 to 1995. The similarity in nucleotide sequence (>99 %) between impala viruses from the 1995 epizootic and a virus of buffalo origin obtained four months later within the area of the outbreak provided the first evidence for natural transmission between these two wildlife species. Similarly, evidence linking the impala epizootics of 1985 and 1988/89 to distinct buffalo genotypes can be deduced from the phylogenetic analyses of the VP1 genotypes of the viruses involved. One of the impala outbreaks, which occurred in 1992/93 could however not be linked to any buffalo field strains.

The high nucleotide sequence identity (>99%) between the 1995 impala viruses and the 1996 buffalo virus is significant, especially in view of the extent of the genetic variation in the maintenance host species. Amino acid sequence identity of the twenty buffalo field strains used in this study was found to be 74.6 %. Similar intratypic sequence identity determinations for European serotype viruses from diverse geographical origins ranged from 73 % to 83 % (Domingo *et al.* 1990). Clearly the corresponding level of sequence variation (25.4 %) within a buffalo from single game park and country exemplifies the high level of intratypic variation found in the SAT types as it surpasses or almost equals that determined for serotypes A, O and C.

The superimposition of the geographical localities of impala (Fig. 4.1) and buffalo viruses (Fig. 4.2) indicates that shared habitats between these species are likely to be important in transmission, as the origin of impala outbreaks overlap with the geographical range of the buffalo herds to which they can be linked. Although the sampling date of the buffalo viruses did not always

coincide with the year of the impala outbreak to which they were linked, it was shown that buffalo viruses group on the basis of geographic origin, rather than year of isolation, viz. relationships are spatial rather than temporal. Studies of buffalo behaviour have determined that buffalo herds have defined home ranges which they use continuously for several years (Mloszewski 1983; Funston *et al.* 1994). This explains the recovery of related viruses from discrete localities even when separated in time and explains the close genetic relationship between temporally unrelated buffalo viruses from a specific locality. Thus, the clustering of the 1991 buffalo virus with the 1985 impala outbreak strain is not surprising because it is representative of the buffalo herd from which the outbreak strain probably originated. It is known that buffalo home ranges can increase significantly in response to ecological pressures (Funston *et al.* 1994) and may vary from approximately 40 to 1000 square kilometres (Mloszewski 1983; Funston *et al.* 1994), with home range size being dictated by game park boundaries and other buffalo herds. In addition, bulls are known to move between buffalo herds (Mloszewski 1983). These factors would contribute to the findings of this study which demonstrated that some buffalo viruses obtained from the southern and northern regions of the study area are clearly more related than expected.

The occurrence of over 90% of outbreaks of FMD in impala within the time from June to November, is significant in that it coincides with the time at which buffalo calves are likely to become infected for the first time. Strict seasonal breeding of buffalo (Pienaar 1969a) results in synchronized calving, with the majority of births being recorded in January and February in the KNP (Fairall 1968). These buffalo calves, when infected for the first time probably excrete virus in approximately the same quantities and by the same routes as do cattle (Gainaru *et al.* 1986). Studies have shown that this first acute infection is likely to occur at around 3-8 months of age when maternally derived antibodies have waned sufficiently to make them susceptible to infection (Condy & Hedger 1978). On this basis it follows that buffalo calves in the KNP are likely to become infected for the first time from the month of May onwards. It is during this period that they are most likely to be a source of infection for other species within the KNP and therefore notable that most outbreaks in impala have occurred at a slightly later, but overlapping time period.

Factors facilitating inter-species transmission are poorly understood, but as buffalo and impala do not come into direct contact in natural circumstances, it can be assumed that elements driving them to share habitats are important. Shared drinking and grazing localities together with the high population density of both species in the southern part of the KNP are likely to facilitate transmission. It is interesting to note that of the five grazing routes identified by Meeser (1962) to be paths of FMD infection, three overlap with the origin of recent impala epizootics, discussed

here. This lends credence to the possibility that shared grazing may play a role in transmission and is further supported by the observation that impala grazing (as opposed to browsing) peaks at around 85 % in November (Skinner *et al.* 1984), and includes two grass species of preference for the African buffalo, viz. *Panicum maximum* and *Digitaria eriantha* (Pienaar 1969b;Skinner *et al.* 1984).

The role of the impala in the epidemiology of FMD remains unclear, although the historical use of this antelope as an indicator species for predicting both the course of an epizootic and outbreaks in livestock (Meeser 1962), would seem to imply that they may be intermediaries in the transmission of FMD from buffalo to cattle. Conflicting evidence exists in the literature, however, regarding the susceptibility of impala to infection and their ability to transmit the disease. Evidence for transmission of virus from buffalo to impala, where shared drinking and feeding was experimentally enforced, could not be demonstrated (Gainaru *et al.* 1986). Conversely, under experimental conditions where no close contact between the species occurred, it was shown that carrier buffalo inadvertently infected impala (Hedger *et al.* 1972) and acutely infected impala infected carrier buffalo (Vosloo *et al.* 1996). Clearly, a special set of circumstances is required to effect indirect interspecies viral transfer and the role of impala as intermediaries in disease transmission remains to be clarified. Circumstantial evidence does however indicate that infected antelope provided the most likely link between buffalo and cattle in a recent outbreak of FMD in cattle in Zimbabwe. In that instance, genetic characterization of buffalo viruses showed that buffalo were the primary source of infection (Hargreaves *et al.* Unpublished observations).

This study provides the first documented evidence of field transmission between buffalo and impala. As more southern African countries attempt to obtain and retain FMD free zones, it is of fundamental importance to elucidate the role of different wildlife species in the epidemiology of the disease so that appropriate disease control measures are implemented. Clearly, additional factors should be investigated in order to fully understand the epidemiological significance and to permit more accurate risk assessment of this antelope species. These include studies confirming or refuting circumstantial evidence pointing to the role of impala as intermediaries in disease transmission and studies determining the importance of environmental factors in facilitating inter-species transmission of FMD.

Chapter 5

Phylogeographic distribution of SAT-type foot-and-mouth disease viruses in African buffalo populations in southern Africa

Summary

In order to assess the genetic variation and relationships of SAT-type viruses within African buffalo populations in different regions in southern Africa, 30 viruses representative of each the SAT-types were selected for this study. The C-terminus half of the immunogenic VP1 gene was sequenced and relationships were resolved by phylogenetic analysis of the nucleotide data. Independently evolving virus populations were identified for each of the three SAT serotypes endemic to southern Africa, with members of the same genotype being defined as those sharing at least 80 % nucleotide sequence identity across the C-terminus region of the VP1 gene, specified here. Most genotype distributions were shown to correspond with discrete geographic localities in the following manner: Genotype I: South Africa and southern Zimbabwe; Genotype II: Botwana, Namibia and western Zimbabwe; Genotype III: Zambia and Malawi; Genotype IV: northern Zimbabwe.

Intra- and intertypic analyses revealed that regions of hypervariability within the SAT-type viruses correspond with immunogenic sites A (G-H loop) and C (C-terminus region) of the VP1 gene. In addition, SAT-1 and SAT-3 type viruses were shown to accumulate mutations in the H-I and F-H loops, respectively, indicating that structural constraints on VP1 gene evolution probably vary between the different SAT-type viruses. Of importance was the strict conservation of the RGD motif in the G-H loop and the presence of a cysteine residue at the base of the G-H loop in all SAT-type viruses analysed here. In common with serotype O viruses, the longer length of the VP1 protein and high degree of conservation of the cysteine residue upstream of the RGD indicates that conformational epitopes are likely to be important antigenic determinants in the endemic African serotypes.

This study identifies independently evolving virus populations in distinct geographic localities, defines the extent of the genetic heterogeneity among field isolates and provides a powerful reference for determining the origin of outbreaks in livestock and wildlife.

Parts of the results presented here have been accepted for publication in *Archives of Virology*

5.1 Introduction

The epidemiology of foot-and-mouth disease (FMD) in southern Africa is complex due to the circulation of six of the seven known serotypes of FMD virus (FMDV), and the role of wildlife in virus maintenance and disease transmission (Thomson 1996). Due to these epidemiological complexities, emphasis is placed on control rather than eradication of the disease and is achieved by vaccination and restriction on animal movement (Hunter 1998). South African Territories (SAT) type viruses are of particular importance in this region due to the high incidence of these serotypes in FMD outbreaks of livestock and the high infection rates in wildlife (Condy *et al.*. 1969). Of the 350 epizootics reported between 1931 and 1990 in southern Africa, approximately 73 % were caused by SAT-type viruses with European types A, O, C and untyped viruses accounting for the remainder (Thomson 1994).

The African buffalo (*Syncerus caffer*) plays a central role in the epidemiology of the disease due to its ability to maintain and transmit the disease (Thomson 1996). Buffalo populations throughout southern Africa have high infection rates with SAT type viruses (Condy *et al.* 1969; Thomson *et al.* 1992) and the geographical distribution of historical outbreaks of the disease in livestock overlap with the distribution of buffalo populations in southern Africa (Thomson 1994). Furthermore, buffalo are known to be efficient maintenance hosts of SAT-type FMD viruses (Condy *et al.* 1985) and have been shown to be the ultimate source of infection for livestock and wildlife species (Dawe *et al.* 1994a; Bastos *et al.* 2000). Historically, buffalo had a continuous distribution throughout sub-Saharan Africa, but although still widespread, populations are now fragmented due both to natural causes and human intervention. The 1889-1897 rinderpest pandemic which swept through Africa, resulted in an estimated 10 000 fold reduction in buffalo numbers (Sinclair 1977). This may have resulted in some of the southern-most pockets of buffalo becoming free of foot-and-mouth disease (Esterhuysen *et al.* 1985). However, most populations throughout southern Africa retain high infection rates with SAT-type viruses.

The three SAT types (1, 2 and 3) which are endemic to sub-Saharan Africa, differ from each other with regard to geographic distribution, infection rates in wildlife and incidence in FMD outbreaks in livestock. SAT-1 and SAT-2 have extensive distributions throughout sub-Saharan Africa and have made incursions into North Africa and the Middle East (Pereira 1981; www.OIE.org). In contrast, SAT-3 is restricted to five southern African (Brooksby 1972) countries and to Uganda (Hedger *et al.* 1973). SAT-1 has the highest seroprevalence rate in the maintenance host, the African buffalo (Condy *et al.* 1969; Thomson *et al.* 1992). Despite the high infection rate of buffalo with SAT-1 type viruses, this serotype has accounted for only 36 % of the SAT-type

outbreaks of FMD in cattle in southern Africa this century, with most outbreaks (48 %) being caused by SAT-2 type viruses and SAT-3 being responsible for 16 % of outbreaks (Thomson 1994). The reason for this anomaly between SAT-type incidence in cattle outbreaks and seroprevalence rates in the African buffalo maintenance host, is unclear, but may be due to differential abilities of SAT-types to cross species barriers. This possibility is supported by the observation that all clinical cases of FMD in impala antelope in South Africa from 1985 to 1995 were due exclusively to SAT-2 type viruses (Bastos *et al.* 2000), despite the higher recovery rate of SAT-1 type viruses from oesophageo-pharyngeal scrapings of buffalo sampled within the same area and time period (Table 5.1).

The SAT-types are genetically distinct from the European and Asian serotypes, on the basis of RNA hybridization studies (Robson *et al.* 1977) and amino acid sequence comparison of the structural protein genes (Palmenberg 1989). Analysis of complete VP1 gene sequences of European strains has revealed high levels of intratypic variation for the A, O and C serotypes (Domingo *et al.* 1990). Studies based on complete (Brown *et al.* 1989; Van Rensburg & Nel 1999) and partial characterization (Vosloo *et al.* 1995; Bastos 1998a; Bastos *et al.* 2000) of the VP1 gene of SAT-type viruses, although limited, have demonstrated that these serotypes have considerably higher levels of intratypic variation than that documented for other FMDV serotypes. Despite the wide range of variation in the SAT-types it is difficult to identify specific immunological subtypes (Pereira 1977; Esterhuysen 1994).

Genetic characterization of SAT-type viruses of African buffalo has mainly addressed the variation within the Kruger National Park (KNP), South Africa (Vosloo *et al.* 1995; Bastos *et al.* 1999; Bastos *et al.* 2000). Comparisons of SAT-1 viruses from South Africa with viruses from two game parks outside South Africa have however revealed that viruses circulating in buffalo in South Africa appear to be genetically distinct from buffalo viruses of Zambia and Zimbabwe (Vosloo *et al.* 1995; Bastos 1998a). Despite the limited sequence data that is presently available for SAT-type viruses, the genetic characterization of viruses of the maintenance host has been shown to be useful for determining the origin of FMD outbreaks in livestock and wildlife (Dawe *et al.* 1994a; Bastos *et al.* 2000). As SAT type outbreaks of FMD in livestock have been recorded in 8 southern African countries (Thomson 1994), efforts to establish a genetic reference base to assist in determining the origin of outbreaks in southern Africa would require that buffalo viruses of all affected countries be characterized. In this chapter an attempt is made to expand the genetic database of SAT type FMD viruses of the maintenance host species by sequencing representative buffalo viruses from as many southern African countries as possible. This information will be used to establish a regional reference base for determining the geographical origin of FMD outbreaks in livestock and wildlife in southern Africa.

TABLE 5.1 Summary of SAT-type viruses isolated from African buffalo in the Kruger National Park (1986-1996)

Year	No of SAT1 isolates / year	% of SAT1 isolates / year	No of SAT2 isolates / year	% of SAT2 isolates / year	No of SAT3 isolates / year	% of SAT3 isolates / year	Total No of isolations / year
1986	4	40 %	3	30 %	3	30 %	10
1987	3	100 %	0	0 %	0	0 %	3
1988	6	46 %	5	38 %	2	15 %	13
1989	24	73 %	8	24 %	1	3 %	33
1990	3	43 %	2	28 %	2	28 %	7
1991	35	61 %	14	24 %	8	14 %	57
1992	8	89 %	1	11 %	0	0 %	9
1993	5	62 %	2	25 %	1	12 %	8
1994	3	43 %	0	0 %	4	57 %	7
1995	4	67 %	2	33 %	0	0 %	6
1996	4	27 %	1	7 %	10	67 %	15
1986-1996	**99**	**59 %**	**38**	**23 %**	**31**	**18 %**	**168**

5.2 Materials and Methods

5.2.1 Study area

For the purpose of this study, southern African countries are defined as those predominantly occurring south of latitude 10°00'S. This area is inclusive of the following eight countries: South Africa, Namibia, Angola, Botswana, Zimbabwe, Mozambique, Zambia and Malawi.

5.2.2 Viruses used in this study

The viruses used in this study originate from African buffalo occurring within conservancies within 6 of the 8 southern African countries. No buffalo viruses were available for Mozambique and Angola, for historical and logistical reasons. 30 buffalo viruses per SAT serotype, representative of as many southern African countries as possible, were selected for genetic characterization. The geographic origin and strain designation of these viruses is summarized in Table 5.2 (SAT-1 viruses), Table 5.3 (SAT-2 viruses) and Table 5.4 (SAT-3 viruses). FMD viruses were isolated on primary pig kidney (PK) cells inoculated with 10 % suspensions (W/V) of African buffalo (*Syncerus caffer*) oesophageo-pharygeal (probang) samples and propagated further on IBRS2 (Insituto Biologico Rim Suino) cells. Samples and tissue culture isolates that were not available at the Onderstepoort Veterinary Institute and which were obtained from other sources are indicated in Tables 5.2 - 5.4.

3.2.3 RNA extraction, cDNA synthesis and PCR amplification

RNA was extracted from PK isolations or low passage cell culture specimens by a modified guanidinium-based nucleic acid extraction method (Boom *et al.* 1990). The RNA viral template was reverse transcribed using AMV-RT (Promega) in the presence of random hexanucleotides (Boehringer Mannheim) and an antisense primer targeting the 2A/B junction (Beck & Strohmaier 1987). Genomic amplification of the FMD viral genome was performed with primers targeting a 500 bp region corresponding to 1D/2A, as previously described (Bastos 1998a). Reactions were performed in a 50 µl volume in the presence of 0.2 mM dNTP, 0.25 µM of each primer, 1x buffer (DynaZyme) and 1 U of a thermostable DNA polymerase (DynaZyme). After an initial denaturation step at 96°C for 30 s, thirty cycles of denaturation at 96°C for 12 s, annealing at 58°C for 20 s and extension at 70°C for 40 s were performed.

3.3.4 PCR purification and nucleotide sequencing

Amplification of the expected fragment was confirmed by product size estimation against a DNA molecular weight marker (φX174 (*HindIII*); Promega) on a 1.5 % agarose gel. Bands of the correct size were excised from the gel and purified by means of the Cleanmix purification system (Talent). The purified products were sequenced with T7 DNA polymerase (Sequenase Version 2.0, USB) and an [α-^{32}P]dATP radioactive label (Amersham) in the presence of 10% DMSO (Winship 1989). Two or more independent amplification and sequencing reactions were performed per

isolate with each of the sequencing primers.

5.2.5 Genomic region used in this study

Genetic characterization of the carboxy-terminal region of the VP1 gene was performed as described in Chapters 2-4. Amino acid sequence data generated for the three SAT-types were aligned and an homologous region commencing with the conserved amino acid sequence 'WXPNG' in the VP1 protein (Fig. 5.7, Fig. 5.8 and Fig. 5.9) and ending on the C-terminal amino acid was identified for further analyses. This region corresponds to approximately 60 % of the VP1 gene and varies in length between serotypes with SAT-1, SAT-2 and SAT-3 having a sequence length of 132, 128 and 130 respectively. This regions is also inclusive of the two immuno-dominant sites within the VP1 gene, namely the A (G-H loop) and C (C-terminus region) sites.

5.2.6 Analyses of the partial VP1 gene sequencing data

Phylogenetic analyses followed the methods advocated by Martin and co-workers (1995) for the analysis of partial VP1 gene sequences of FMD viruses. Briefly, nucleotide sequences were translated and aligned using the DAPSA programme (Harley 1994) and neighbor-joining trees using the Jukes and Cantor (1969) algorithm were constructed. In order to identify the virus clusters that are well-supported (and repeatedly recovered), 1000 bootstrap replications were performed and a 60 % majority consensus rule was applied.

5.2.7 Amino acid variability and secondary structure prediction

Amino acid sequence data of 30 buffalo viruses of each serotype were aligned and used to plot VP1 gene variability, using the MEGA programme (Kumar *et al.* 1993). Analyses were performed with overlapping windows of 10 amino acids, with windows containing 7 or more variable sites, being considered hypervariable. Secondary structure predictions were performed with the nnpredict programme (Kneller *et al.* 1990), which was available through the BCM search launcher at http://www.cmpharm.ucfs.edu. Tertiary structure could only be inferred by aligning SAT-type reference strains with that of a type O_1BFS (Makoff *et. al.* 1982) whose tertiary structure has been resolved (Acharya *et. al.* 1989).

TABLE 5.2 Species and geographical origin of 30 SAT-1 type viruses from African buffalo populations in southern Africa (1985-1998)

Strain designation	Year of sampling	Country	Game Park / Place of origin	Grid reference
MAL/1/85	1985	Malawi	Kasungu NP	33°30'E-13°00'S
*ZIM/3/88	1988	Zimbabwe	Hwange NP	27°00'E-19°00'S
SWA/2/89	1989	Namibia	Caprivi	23°20'E-17°50'S
*ZIM/HV/3/90	1990	Zimbabwe	Hippo Valley NP	31°37'E-21°07'S
*ZIM/2/90	1990	Zimbabwe	Chirisa SA	28°15'E-18°00'S
*ZIM/Gn/13/91	1991	Zimbabwe	Gonarezhou	32°00'E-21°30'S
KNP/196/91	1991	South Africa	Kruger NP	31°38'E-25°17'S
*ZIM/2/91	1991	Zimbabwe	Bumi Hills	28°22'E-16°49'S
ZAM/2/93	1993	Zambia	Kafue NP	26°00'E-16°30'S
*ZIM/7/93	1993	Zimbabwe	Lone Star	31°54'E-21°07'S
*ZIM/2/94	1994	Zimbabwe	Hwange NP	27°00'E-19°00'S
KNP/8/95	1995	South Africa	Kruger NP	31°38'E-24°34'S
KNP/14/95	1995	South Africa	Kruger NP	31°38'E-24°34'S
KNP/17/95	1995	South Africa	Kruger NP	31°38'E-24°34'S
KNP/41/95	1995	South Africa	Kruger NP	31°38'E-24°34'S
KNP/2/96	1996	South Africa	Kruger NP	31°30'E-24°55'S
KNP/17/96	1996	South Africa	Kruger NP	31°32'E-23°18'S
KNP/22/96	1996	South Africa	Kruger NP	NA
ZAM/18/96	1996	Zambia	Kafue NP	26°25'E-16°17'S
ZAM/29/96	1996	Zambia	Lochinvar NP	27°30'E-15°50'S
*BOT/2/98	1998	Botswana	Nxaraga	23°15'E-19°40'S
*BOT/14/98	1998	Botswana	Nxaraga	23°15'E-19°40'S
*BOT/25/98	1998	Botswana	Vumbura	22°41'E-18°56'S
*BOT/37/98	1998	Botswana	Vumbura	22°41'E-18°56'S
KNP/75/98	1998	South Africa	Kruger NP	31°16'E-22°59'S
NAM/272/98	1998	Namibia	Mamili NP	23°39'E-18°22S
NAM/306/98	1998	Namibia	West Caprivi NP	21°50'E-18°15'S
NAM/307/98	1998	Namibia	West Caprivi NP	21°50'E-18°15'S
NAM/308/98	1998	Namibia	West Caprivi NP	21°50'E-18°15'S
*ZIM/14/98	1998	Zimbabwe	Lubangwa Island	28°20E-16°47'S

SA: Safari Area; NP: National Park; NA: Not available; *Viruses supplied by the World Reference Laboratory, Pirbright, UK

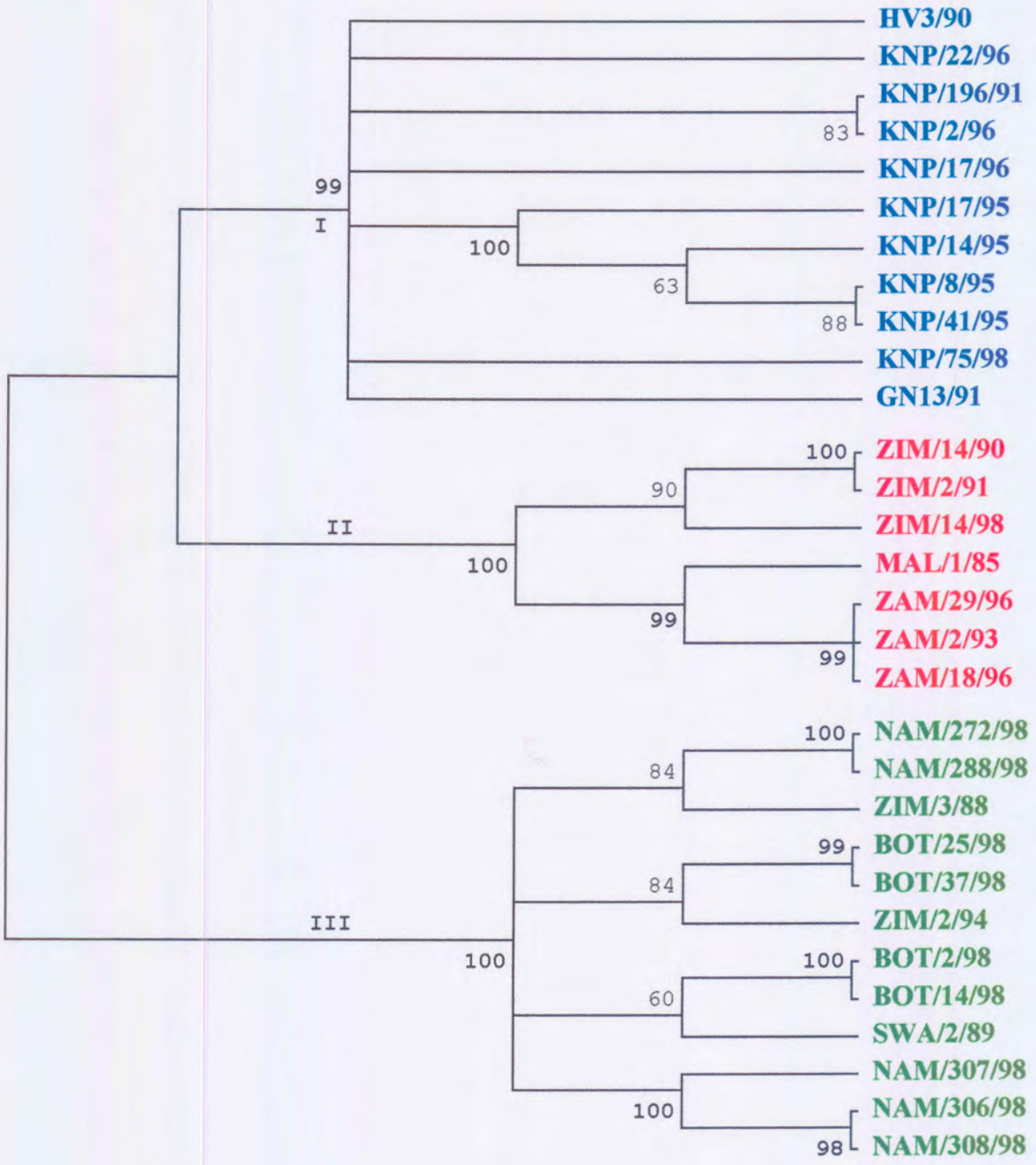

Fig. 5.1: Neighbor-joining tree based on 396 nt of the VP1 gene (amino acid position 87-221) depicting SAT-1 buffalo virus relationships in southern African. Jukes & Cantor algorithm with 1000 bootstrap replications and a 60 % majority consensus rule were applied. Bootstrap values ≥ 95 % confidence level are indicated in bold. Virus clusters (labeled I - III) are colour-coded to coincide with colours used to indicate the geographical distribution of these genotypes in southern Africa (Fig. 5.2).

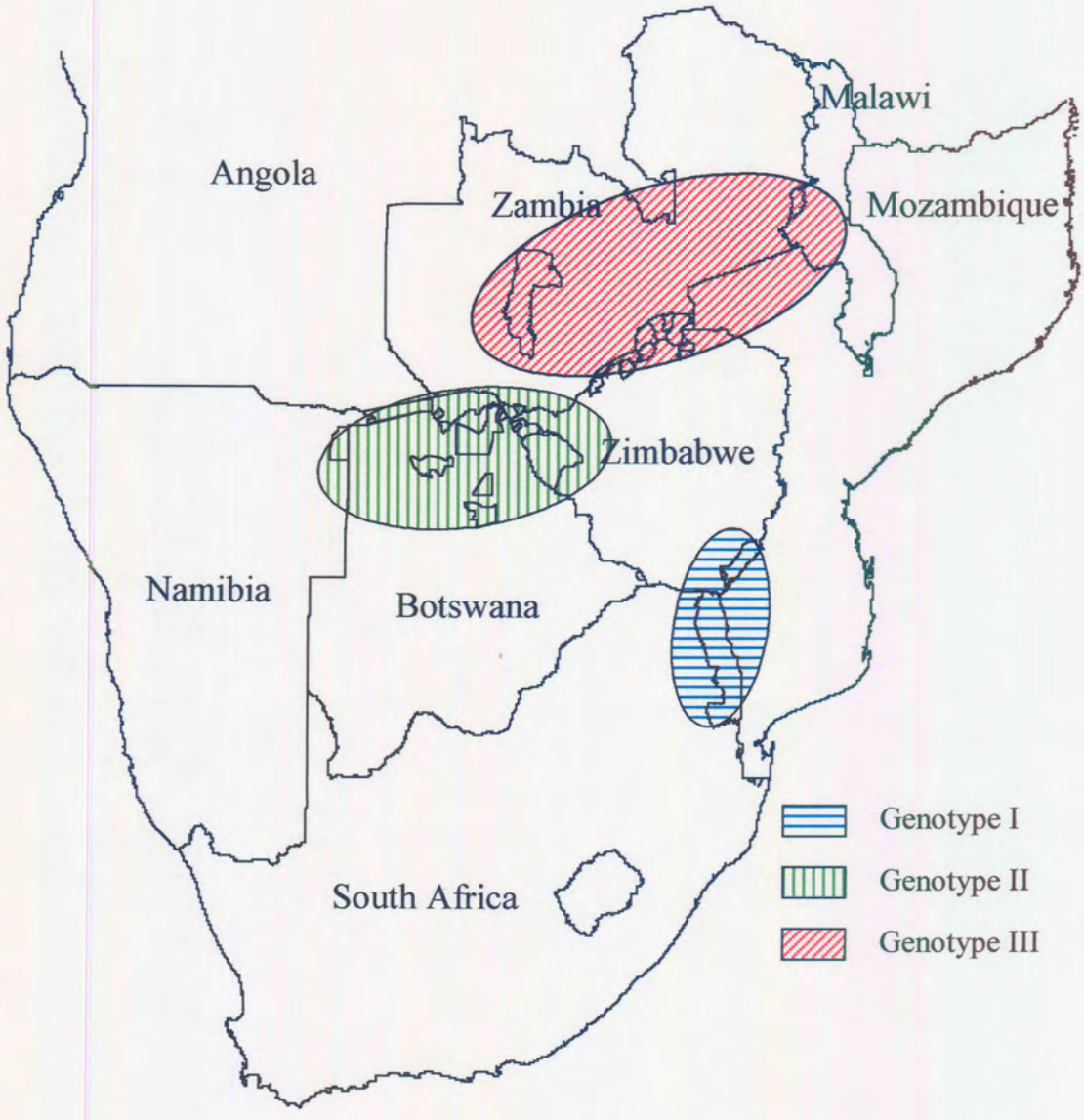

Fig. 5.2: Map of southern African game parks indicating the geographical distribution of SAT-1 buffalo virus genotypes. These regions are colour-coded as follows: blue (genotype I); green (genotype II); red (genotype III)

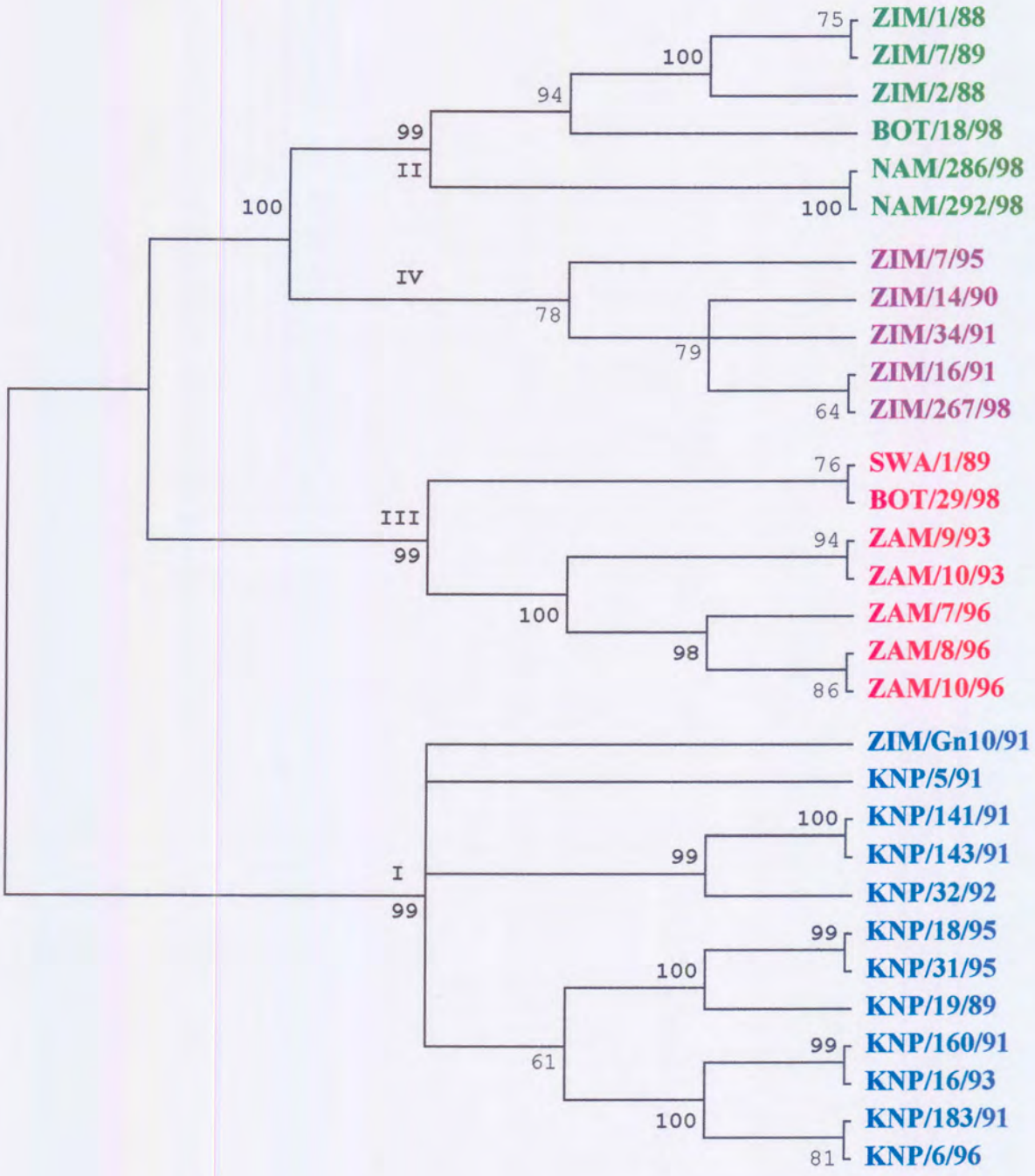

Fig. 5.3: Neighbor-joining tree based on 384 nt of the VP1 gene (amino acid position 89-216) depicting SAT-2 buffalo virus relationships in southern African. Jukes & Cantor algorithm with 1000 bootstrap replications and a 60 % majority consensus rule were applied. Bootstrap values ≥ 95 % confidence level are indicated in bold. Virus clusters (labeled I - IV) are colour-coded to coincide with colours used to indicate the geographical distribution of these genotypes in southern Africa (Fig. 5.4).

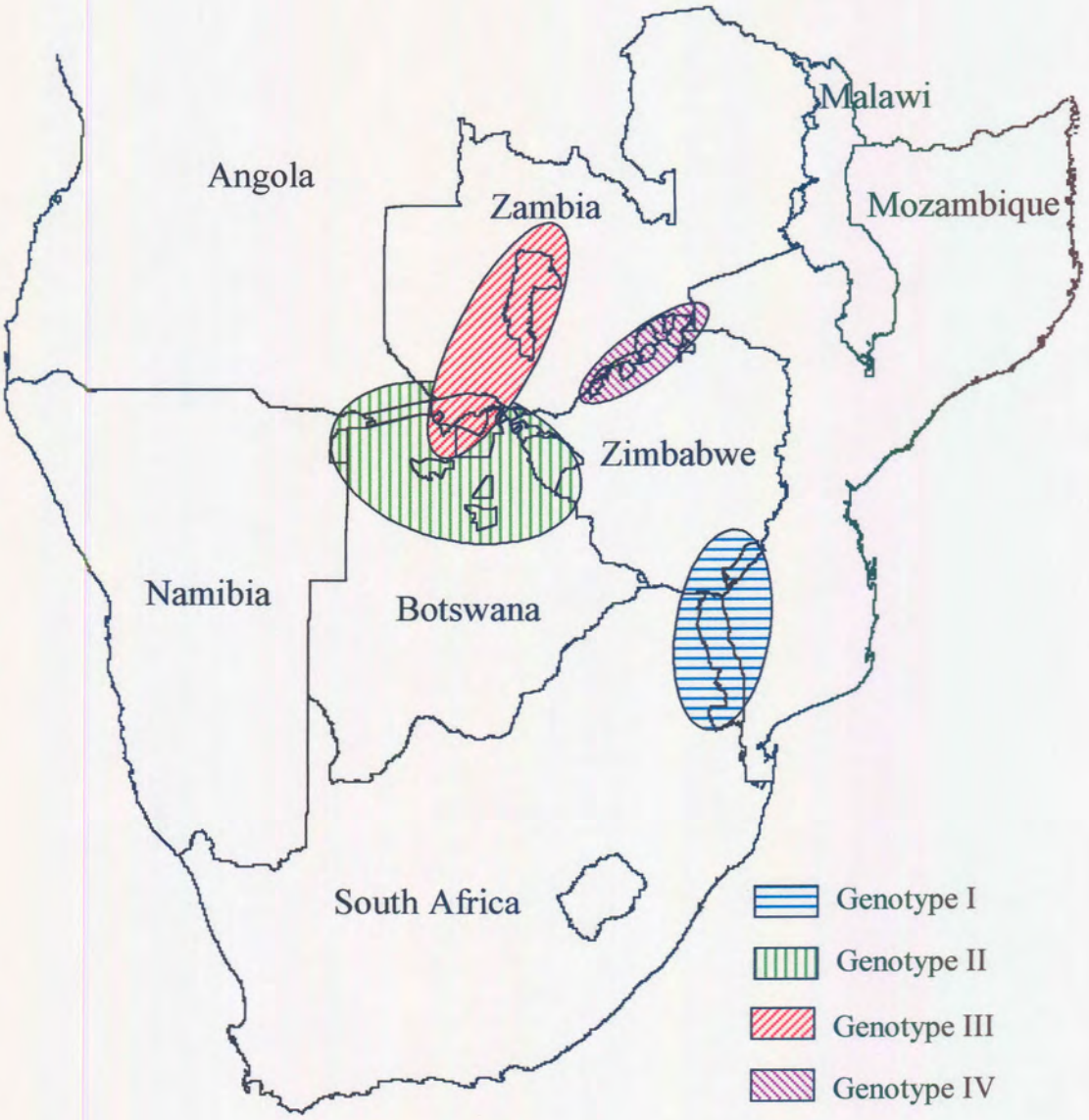

Fig. 5.4: Map of southern African game parks indicating the geographical distribution of SAT-2 buffalo virus genotypes. These regions are colour-coded as follows: blue (genotype I); green (genotype II); red (genotype III); purple (genotype IV)

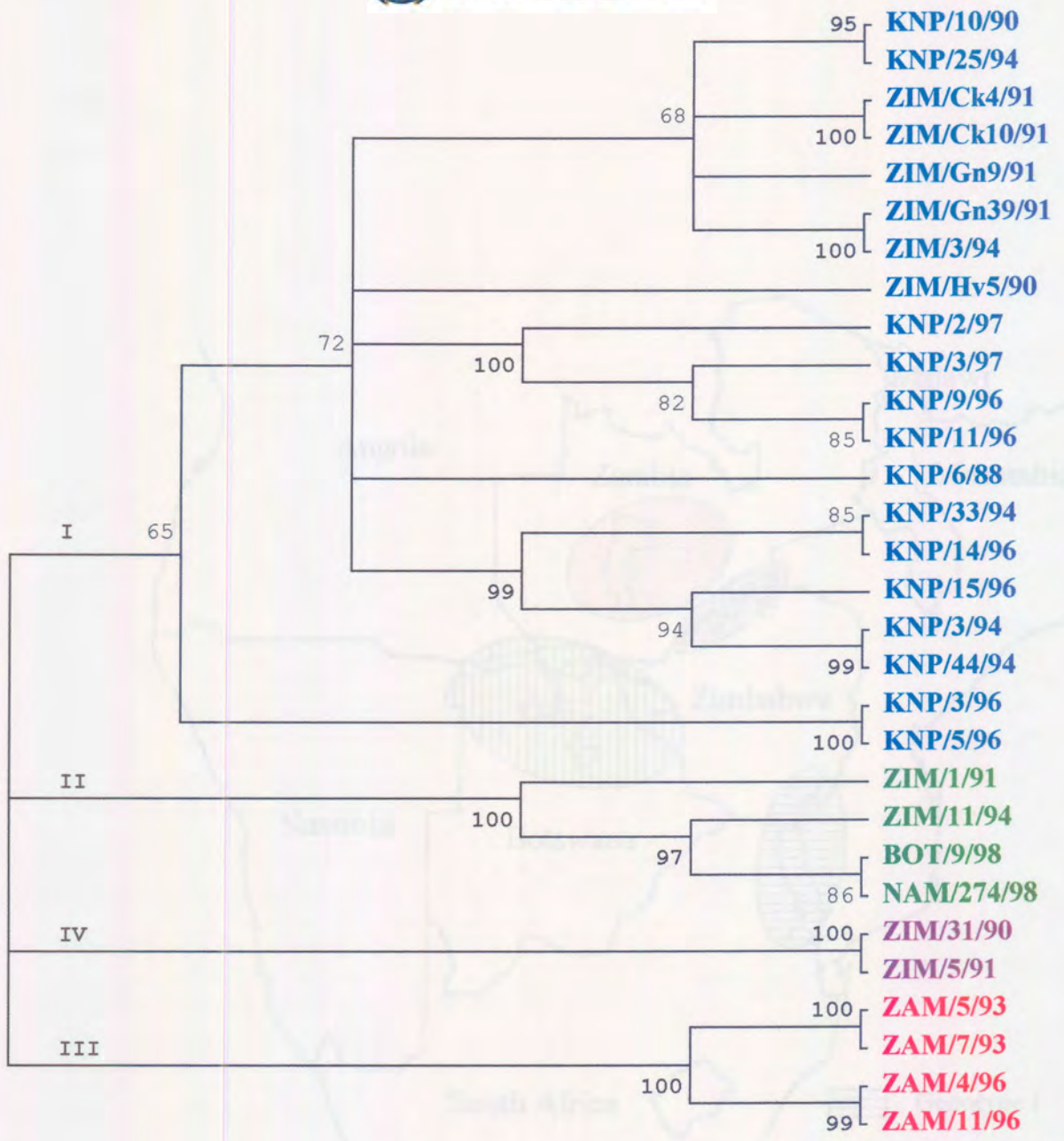

Fig. 5.5: Neighbor-joining tree based on 390 nt of the VP1 gene (amino acid position 88-217) depicting SAT-3 buffalo virus relationships in southern African. Jukes & Cantor algorithm with 1000 bootstrap replications and a 60 % majority consensus rule were applied. Bootstrap values ≥ 95 % confidence level are indicated in bold. Four distinct genotypes (I-IV) are denoted and are colour-coded to correspond to the geographical origin and distribution of SAT-3 viruses, depicted in Fig. 5.6.

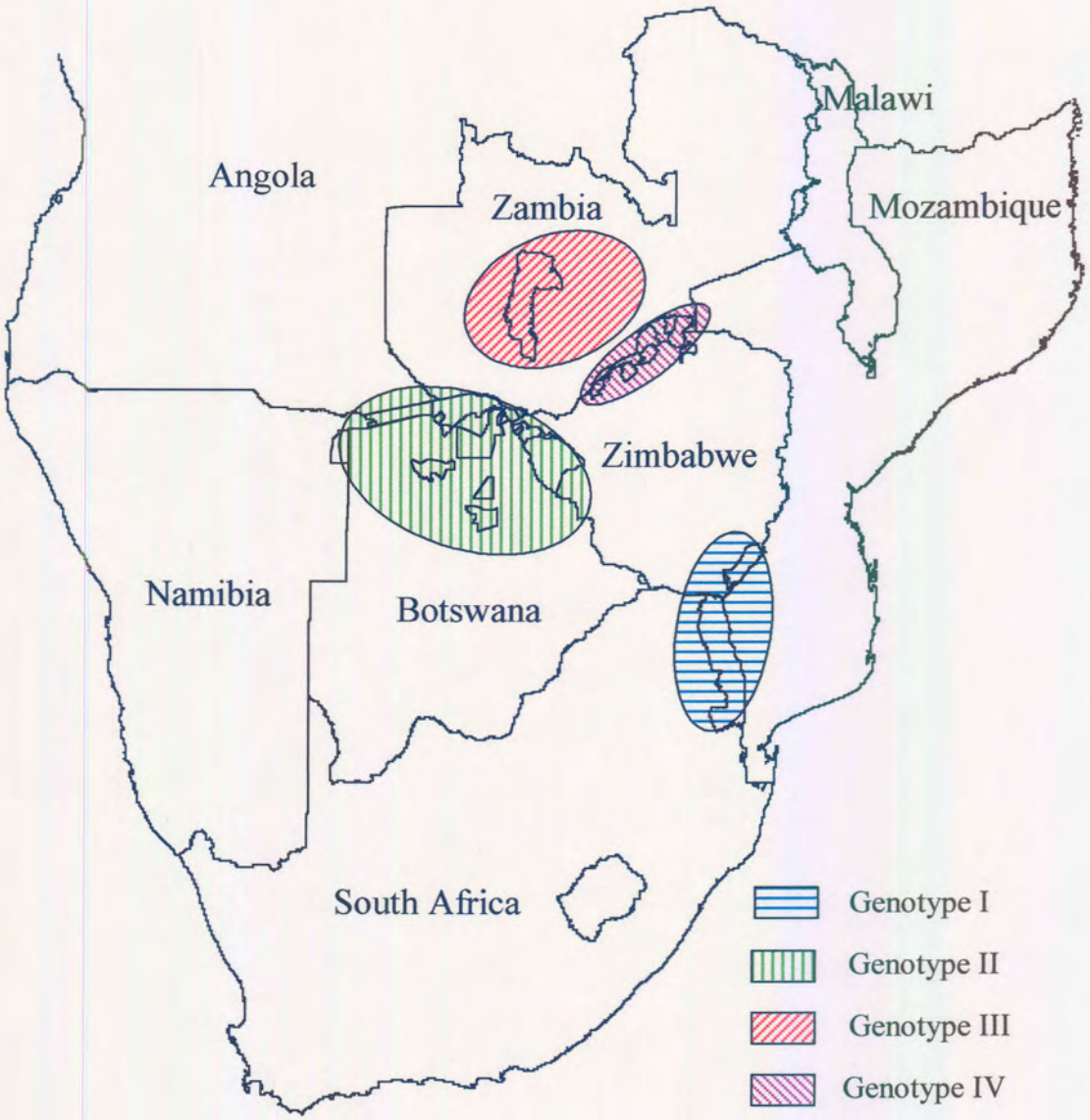

Fig. 5.6: Map of southern African game parks indicating the geographical distribution of SAT-3 buffalo virus genotypes. These regions are colour-coded as follows: blue (genotype I); green (genotype II); red (genotype III); purple (genotype IV)

TABLE 5.3 Species and geographical origin of 30 SAT-2 type viruses from African buffalo populations in southern Africa (1988-1998)

Strain designation	Year of sampling	Country of origin	Game Park / Sampling locality	Grid Reference
*ZIM/1/88	1988	Zimbabwe	Hwange NP	27°E-19°S
*ZIM/2/88	1988	Zimbabwe	Hwange NP	27°E-19°S
*ZIM/4/88	1988	Zimbabwe	Hwange NP	27°E-19°S
KNP/19/89	1988	South Africa	Kruger NP	31°37'E-24°44'S
SWA/1/89	1989	Namibia	Caprivi	23°25'E-14°41'S
*ZIM/7/89	1989	Zimbabwe	Hwange NP	27°E-19°S
*ZIM/14/90	1990	Zimbabwe	Doma SA	30°15'E-16°20'S
KNP/5/91	1991	South Africa	Kruger NP	31°47'E-24°24'S
KNP/143/91	1991	South Africa	Kruger NP	31°20'E-24°00'S
KNP/147/91	1991	South Africa	Kruger NP	31°20'E-24°00'S
KNP/160/91	1991	South Africa	Kruger NP	31°34'E-24°33'S
KNP/183/91	1991	South Africa	Kruger NP	31°56'E-25°08'S
*Gn/10/91	1991	Zimbabwe	Gonarezhou NP	32°00'E-21°30'S
*ZIM/16/91	1991	Zimbabwe	Matusadona NP	28°30'E-16°45'S
*ZIM/34/91	1991	Zimbabwe	Urungwe SA	28°55'E-16°30'S
KNP/32/92	1992	South Africa	Kruger NP	31°17'E-22°54'S
KNP/16/93	1993	South Africa	Kruger NP	31°26'E-23°29'S
ZAM/9/93	1993	Zambia	Kafue NP	26°00'E-16°30'S
ZAM/10/93	1993	Zambia	Kafue NP	26°00'E-16°30'S
*ZIM/8/94	1994	Zimbabwe	Matetsi NP	26°50'E-18°00'S
KNP/18/95	1995	South Africa	Kruger NP	31°38'E-24°34'S
KNP/31/95	1995	South Africa	Kruger NP	31°38'E-24°34'S
*ZIM/7/95	1995	Zimbabwe	Sengwa	28°00'E-17°05'S
KNP/6/96	1996	South Africa	Kruger NP	32°01'E-25°14'S
ZAM/8/96	1996	Zambia	Mulanga	NA
ZAM/10/96	1996	Zambia	Mulanga	NA
*BOT/29/98	1998	Botswana	Vumbura	NA
*BOT/31/98	1998	Botswana	Vumbura	NA
NAM/286/98	1998	Namibia	East Caprivi GR	23°20'E-17°52'S
NAM/292/98	1998	Namibia	East Caprivi GR	23°20'E-17°52'S
ZIM/267/98	1998	Zimbabwe	Chizarira	28°00'E-17°47'S

GR: Game Reserve; NP: National Park; SA: Safari Area; NA: Not available; * Indicates viruses supplied by WRL

TABLE 5.4 Species and geographical origin of 30 SAT-3 viruses from African buffalo populations in southern Africa (1988-1998)

Strain designation	Year of sampling	Country of origin	Game Park / Sampling locality	Grid Reference
KNP/6/88	1988	South Africa	Kruger NP	31°16'E-25°10'S
KNP/10/90	1990	South Africa	Kruger NP	31°20'E-24°00'S
*ZIM HV5/90	1990	Zimbabwe	Hippo Valley NP	31°35'E-21°10'S
*ZIM/31/90	1990	Zimbabwe	Dande SA	30°20'E-15°55'S
*ZIM Ck4/91	1991	Zimbabwe	Chikwarakwara	31°05'E-22°20'S
*ZIM Ck10/91	1991	Zimbabwe	Chikwarakwara	31°05'E-22°20'S
*ZIM Gn9/91	1991	Zimbabwe	Gonarezhou NP	32°00'E-21°30'S
*ZIM Gn39/91	1991	Zimbabwe	Gonarezhou NP	32°00'E-21°30'S
*ZIM/1/91	1991	Zimbabwe	Matusadona NP	28°30'E-16°45'S
*ZIM/5/91	1991	Zimbabwe	Urungwe SA	28°55'E-16°30'S
ZAM/5/93	1993	Zambia	Kafue NP	26°00'E-16°30'S
ZAM/7/93	1993	Zambia	Kafue NP	26°00'E-16°30'S
KNP/3/94	1994	South Africa	Kruger NP	31°55'E-25°15'S
KNP/25/94	1994	South Africa	Kruger NP	31°55'E-25°15'S
KNP/33/94	1994	South Africa	Kruger NP	31°55'E-25°15'S
KNP/44/94	1994	South Africa	Kruger NP	31°55'E-25°15'S
*ZIM/3/94	1994	Zimbabwe	Matetsi	25°50'E-18°06'S
*ZIM/11/94	1994	Zimbabwe	Hwange NP	27°E-19°S
KNP/3/96	1996	South Africa	Kruger NP	31°30'E-24°55'S
KNP/5/96	1996	South Africa	Kruger NP	31°30'E-24°55'S
KNP/9/96	1996	South Africa	Kruger NP	31°23'E-22°52'S
KNP/11/96	1996	South Africa	Kruger NP	31°23'E-22°52'S
KNP/14/96	1996	South Africa	Kruger NP	NA
KNP/15/96	1996	South Africa	Kruger NP	NA
ZAM/4/96	1996	Zambia	Mulanga	NA
ZAM/11/96	1996	Zambia	Kafue NP	26°00'E-16°30'S
KNP/2/97	1997	South Africa	Kruger NP	31°23'E-22°52'S
KNP/3/97	1997	South Africa	Kruger NP	31°23'E-22°52'S
*BOT/9/98	1998	Botswana	Moremi GR?	NA
NAM/274/98	1998	Namibia	Mamili NP	23°39'E-18°22'S

GR: Game Reserve; NP: National Park; SA: Safari Area; NA: Not available; * Indicates viruses supplied by WRL

5.3 Results

5.3.1 Phylogeographic distribution of SAT-1 buffalo viruses

In order to determine the genetic relationships of SAT-1 viruses in southern Africa, an homologous 396 nt region, corresponding to amino acid positions 80-221 of the VP1 protein was identified for phylogenetic analyses. The neighbor-joining method (Saitou & Nei 1987) and Jukes & Cantor (1969) algorithm were used to construct the gene tree as this evolutionary model was shown by Martin and co-workers (1995) to increase the probability of recovering the correct phylogenetic tree when partial VP1 data is used. Alternative distance estimation algorithms produced trees with identical topology as did UPGMA and parsimony methods (results not shown), confirming that the recovered phylogeny is reliable (Kim 1993). In addition, bootstrap support (based on 1000 replications) was consistently above 90 %, irrespective of the type of analysis method or evolutionary model used. Results indicate that there are three independently evolving SAT-1 virus clusters (Fig. 5.1), which correspond to distinct geographical localities (Fig. 5.2) as follows:

Genotype I: South Africa and southern Zimbabwe (99 % bootstrap support)
Genotype II: Botswana, western Zimbabwe and Namibia (100 % bootstrap support)
Genotype III: northern Zimbabwe, Zambia and Malawi (100 % bootstrap support)

As these three viral genotypes correspond to three geographically independent regions, the term 'topotype' is used to reflect this association. This grouping of viruses on the basis of geographical origin has previously been described for serotype O FMDV (Knowles *et al.* 1997) and for numerous unrelated viruses (Deubel *et al.* 1986; Kerschner *et al.* 1986; Monath *et al.* 1986; Calisher *et al.* 1987; Chu *et al.* 1989; Lindsay *et al.* 1993; Yanagihara 1994).

5.3.2 Phylogeographic distribution of SAT-2 buffalo viruses

Four major SAT-2 viral lineages (Fig. 5.3) were similarly obtained when a phylogeny was inferred using partial VP1 sequence data corresponding to amino acid positions 89-216. The geographical distribution of viruses (indicated in Fig. 5.4) comprising each of the four SAT-2 genotypes is as follows:

Genotype I: South Africa and southern Zimbabwe (99 % bootstrap support)
Genotype II: Botswana, Namibia and western Zimbabwe (99 % bootstrap support)
Genotype III: Botswana, Namibia and Zambia (99 % bootstrap support)
Genotype IV: Northern Zimbabwe (78 % bootstrap support)

On the basis of these results, viruses of the same SAT-2 genotype are defined as those sharing at least 80 % sequence identity across the C-terminus region of VP1 specified in this study. In contrast to the genetically and geographically distinct grouping of SAT-1 viruses, co-circulation of 2 distinct genotypes within Botswana and Namibia was observed for the SAT-2 serotype (Fig. 5.4). The map depicting geographic distribution of SAT-2 variants in southern Africa clearly shows that viruses originating from the Okavango Delta (Botswana) and from the eastern tip of the Caprivi strip (Namibia) fall within two sub-clusters of genotypes II and III and have overlapping distributions. In particular, BOT/29/98 and SWA/1/89 are the only viruses originating from Botswana and Namibia that appear to be genetically more closely related to Zambian viruses (Genotype III) than to other viruses originating from Botswana and Namibia (Genotype II).

5.3.3 Phylogeographic distribution of SAT-3 type buffalo viruses

Four viral lineages were identified for SAT-3 type viruses from southern Africa when a phylogeny was inferred using partial VP1 sequence data (Fig. 5.5). These virus clusters were shown to correspond to distinct geographic localities (Fig. 5.6) as follows:

Genotype I: South Africa and southern Zimbabwe (65 % bootstrap support)
Genotype II: Namibia, Botswana and western Zimbabwe (100 % bootstrap support)
Genotype III: Zambia (100 % bootstrap support)
Genotype IV: northern Zimbabwe (100 % bootstrap support)

In contrast to the SAT-1 serotypes, SAT-3 has four genotypes that are phylogeographically distinct from each other. These genotypes therefore concur with the topotype definition. The geographic distribution of the southern genotype (I) is identical to that observed for SAT-1 and SAT-2 as is the distribution of SAT-1 and SAT-3 western genotype (II) viruses. It is noteworthy that genotype III and IV SAT-3 viruses from Zambia and northern Zimbabwe, respectively, form distinct country-specific virus groupings within each of the major genotypes into which they fall.

In 1994, buffalo were translocated from Gonarezhou National Park in southern Zimbabwe to Matetsi National Park in the west. A virus recovered from one of these translocated buffalo, ZIM/3/94 groups with a virus obtained from a Gonarezhou buffalo in 1991 (ZIM/Gn39/91), with 100 % bootstrap support. Clearly, the ZIM/3/94 virus recovered from the Matetsi buffalo reflects its true origins, indicating that VP1 gene characterization may provide a powerful means of identifying the origin of buffalo moved between different genotype localities.

5.3.4 Phylogeographic comparison of the three SAT-type viruses in southern Africa

Sequence identity (percentage of sites that are completely conserved across all the sequences being analysed) within each of the SAT-type buffalo virus genotypes was determined (Table 5.5). Overall variation within the homologous region of the VP1 protein characterized in this study was highest within SAT-3 and lowest in SAT-1 (on amino acid level). In addition, SAT-1 has the lowest number of distinct viral lineages despite having a widespread distribution throughout sub-Saharan Africa. In contrast, SAT-3 which has the most restricted distribution, has the highest level of intratypic variation and occurs as four independently evolving virus clusters. Levels of nucleotide sequence variation by topotype ranged from 32 % to 33 % and from 20 % to 23 % on amino acid level (results not shown).

TABLE 5.5 Comparison of the three SAT-type viruses of the African buffalo (*Syncerus caffer*) maintenance host, in southern Africa

Serotype	No of viruses	No of southern African countries	Nucleotide variation	Amino Acid variation	*No of southern African viral lineages
SAT-1	30	6	52 %	42 %	3
SAT-2	30	5	54 %	43 %	4
SAT-3	30	5	55 %	46 %	4

** Number of lineages obtained with the NJ method with p-distances, 1000 bootstrap replications and a 60% majority consensus rule applied.*

5.3.5 Structurally and immunologically important amino acid sites in SAT-1 viruses

The RGD motif (amino acids 149-151), which is thought to be the cell attachment site of the virus (Geysen *et al.* 1985; Fox *et al.* 1989) and is known to interact with the αv ß3 integrin cellular receptor (Berinstein *et al.* 1995) is completely conserved in the SAT-1 serotype (Fig. 5.7). In addition, leucine at position 152 is also completely conserved in all but one virus (KNP/8/95) as is proline at position 166 (with the exception of one buffalo virus, KNP/17/96). Leu-152 and Pro-166 of SAT-1 type viruses correspond to Leu-148 and Pro-153 of serotype A_{12} which were shown to be important residues in

determining antibody interactions with the highly immunogenic GH-loop (France *et al.* 1994). A conserved cysteine residue at position 135 in the SAT-1 serotype, corresponds to the structurally important Cys-134 in serotype O viruses (Fig.5.10). This cysteine forms a disulphide bond with a cysteine residue in VP2 and is thought to explain why most neutralizing monoclonal antibodies against type O viruses recognize conformational rather than linear epitopes (Acharya *et al.* 1990). The 1D/2A junction site of SAT-1 type viruses is variable, with two distinct forms occurring, viz. 'KQ/MC' in 2/30 buffalo viruses and KQ/LC in the remaining 28 SAT-1 sequences (Fig. 5.7). Of interest is the identification of an additional RGD triplet (position 110-112), upstream of the GH-loop RGD site, in the Namibian virus, designated NAM/307/98.

5.3.6 Structurally and immunologically important amino acid sites in SAT-2 viruses

Amino acid alignment of the 30 SAT-2 buffalo viruses (Fig. 5.8) reveals that the 'RGD' cell-attachment motif (position 144-146) is completely conserved. In contrast to SAT-1, a conserved leucine, adjacent to the RGD motif, does not occur in SAT-2 type viruses. Instead an arginine is observed at position 147 in 28 of the 30 buffalo viruses sequenced. Leu-150 (corresponding to Leu-152 in SAT-1) is however completely conserved as is Pro-162 (which corresponds to Pro-166 in SAT-1). These two sites have been shown to be critical residues in determining antibody interactions with the highly immunogenic G-H loop (France *et al.* 1994). Cys-134 (corresponding to Cys-135 of SAT-1) at the base of the G-H loop (Fig. 5.10) was observed in all SAT-2 viruses sequenced. Of interest is the conservation of the C-terminus 'KQ/LC' at the 1D/2A junction site. In common with SAT-1 type viruses, amino acid positions 110-112, corresponding to positions 109-111 in SAT-2 serotype, were extremely variable, but none of the SAT-2 viruses characterized here had an 'RGD' at this site.

5.3.7 Structurally and immunologically important amino acid sites in SAT-3 viruses

The 'RGD' cell-attachment motif (position 145-147) is completely conserved in the 30 SAT-3 buffalo viruses sequenced (Fig. 5.9). A leucine occurs at position 148 in 20 of these buffalo strains whilst a methionine is present in the remainder. This is in contrast to

the SAT-1 in which a leucine, adjacent to the RGD motif, occurs in all but one strain. In common with the SAT-1 and SAT-2 type viruses, Leu-151 (corresponding to Leu-152 in SAT-1) and Pro-164 (which corresponds to Pro-166 in SAT-1) is completely conserved in the SAT-3 serotype. These two sites are also conserved in serotype A viruses where they are known to be of immunological importance for antibody recognition (France *et al.* 1994). Similarly, Cys-133 (corresponding to Cys-135 of SAT-1) at the base of the G-H loop (Fig. 5.10) occurs in all SAT-3 viruses sequenced. In common with the SAT-1 serotype, the 1D/2A junction site of SAT-3 viruses shows some degree of variability. Although the 'KQ/LC' was the most predominant form (> 80 % of SAT-3 viruses), 'KQ/LS', 'KQ/LY' and 'KQ/TC' forms were also observed. Amino acid positions 108-110 of SAT-3, which correspond to the region in SAT-1 viruses where the additional 'RGD' was observed, were variable. The closest variant on the 'RGD' triplet, was a 'RNG' which was observed in a Zambian isolate (ZAM/11/96; Fig. 5.9).

5.3.8 Distribution of hypervariable regions in the C-terminus half of the VP1 gene

VP1 gene amino acid variability was plotted for all buffalo viruses in order to determine whether mutations were randomly distributed or localized to specific regions of the VP1 gene. Three discrete hypervariable regions, in which 70 % or more of the sites varied (for overlapping windows of 10 amino acids) were identified in this way. In all SAT serotypes, hypervariable regions were located at amino acid positions coinciding with serotype C antigenic sites A (the GH-loop) and C (C-terminus region) of the protein. These hypervariable and immunodominant sites are known to elicit neutralizing antibodies in European serotypes. Additional hypervariable regions were identified for each of the serotypes and are summarized in Table 5.6. Despite the high levels of variability in the primary sequence, secondary structure predictions reveal that the hypervariable regions have highly conserved secondary structures (results not shown). In particular, strict secondary structure conformation is maintained in the form of an alpha helix immediately following the RGD binding motif (positions 152-159; Fig. 5.10).

	99	109	119	129	139	149	159
SWA/2/89	WTPNGSPVMS	EVGDNPVVFS	HQGTTRFALP	YTAPHRVLAT	VYNGDCKYKP	TGTAPRENIR	GDLATLAARI
NAM/272/98	.A.....F.	....	.NN.....I.	....	....	...P..T...	.......E..
NAM/306/98	.......ELT	....	RG........	....	....	...P......	.......K..
NAM/307/98	........LT	....	RGD.......	....	....	...P......	....
NAM/308/98	.......QLT	....	RG........	....	....	.S.P......	.......K..
ZIM/3/88	........LG	....	.NN.....I.	.?........	....	.D.PA.D...	.......E..
ZIM/2/94	.......ELT	....	.NN.......	.......V..	A.........	.SAP......	.......T..
BOT/2/98	........L.	....	.G......F.	......L...	....	......G...	....
BOT/14/98	........L.	....	.G........	....	....	....	....
BOT/25/98	........LE	....	.N........	.......V..	T.........	.S.P....V.	....V.....
BOT/37/98	.......ELT	....	.N........	.......V..	T.........	.S.P....V.	....V.....
ZIM/14/90	.......EL.	....	.KR.....I.	....	T.........	Q.Q...T...	.......Q..
ZIM/2/91	.......EL.	....	.N......I.	....	T.......N.	Q.Q...T...	.......Q..
ZIM/14/98	.....A.EL.	.........A	.NN.......	......C...	T.........	VDQ...T.V.	....V..E..
ZAM/2/93	.....A.EL.	....	.N........	......C...	T...N.....	ATE..PTHV.	....
ZAM/18/96	.....A.EL.	....	.N........	......C...	T.........	VSEE..TH..	....
ZAM/29/96	.....A.EL.	....	.N........	......C...	A.........	VNEV.STH..	....
MAL/1/85	.....A.EL.	....	.N........	......C...	T.........	NSE...T...	....A..E..
HV3/90	.....A.ELN	..C.......	.N........	......C...	T.........	A.....G...	.......H..
ZIM/7/93	.....A.ELD	....	.N........	......C...	T.........	....	.......H..
GN13/91	.....A.ELG	....	.N........	......C...	T.........	A..P......	.......Q..
KNP/196/91	.....A.ELA	....	KG........	......C...	A.........	....	....
KNP/8/95	.....A.EL.	....	AN........	......C...	T.........	.D.P..T...	..F....E..
KNP/14/95	.....A.EL.	....	TN........	......C...	T.........	...P..T...	.......E..
KNP/17/95	.....A.EL.	....	TN........	......C...	P.........	...P..T...	.......E..
KNP/41/95	.....A.EL.	....	AN........	......C...	T.........	.D.P..T...	.......E..
KNP/2/96	.....A.EFL	......C...	KD........	......C...	T.........	....	....
KNP/17/96	.....A.ELT	....	KN........	......C...	S.........	A.DP..G...	....
KNP/22/96	.....A.EFD	....	KNR.A.....	......C...	L.........	AS.T..T...	.......E..
KNP/75/98	.....A.EL.	....	AN........	......C...	S.........	A.....D...	.......N..

	169	179	189	199	209	219	2A
SWA/2/89	ASETHIPTTF	NYGMIYTEAD	VDVYLRMKRA	ELYCPRPVLT	HYDHAGKDRY	KTVLVKPAKQ	MC
NAM/272/98	.N.....	.....E	.....	.....L..	....G.....	Q.A.IR....	L.
NAM/306/98	.....	.....Q.E	.....	.....	....G.....	..A..R....	L.
NAM/307/98	.....	.....Q.E	.....	.....	....S.....	..A..R....	L.
NAM/308/98	.....	.....Q.E	.....	.....	....G.....	..A..R....	L.
ZIM/3/88	.N.....	.....Q.E	.N.....	.....L..	....R.....	..D.T.....	L.
ZIM/2/94	.....	.....Q.E	.....	.....	....G.....	..A.....	L.
BOT/2/98	.G.....	.....Q.E	.....	.....L..	....SN....	RVA.I.....	..
BOT/14/98	.G.....	.....Q.E	.....	.....L..	....SN....	RVA.....	..
BOT/25/98	.....	.....E	.....	.....	....N.....	.AR.....	L.
BOT/37/98	.....	.....	.....	.....	....K.....	Q.A.....	L.
ZIM/14/90	.E.....S.	...R.....E	....V.....	.....	....QN....	.VALT.....	L.
ZIM/2/91	.E.....S.	...R.....E	.TS.V.....	.....	....Q.....	.VALT.....	LY
ZIM/14/98	.A.....	...R.....E	....V.....	.....L..	G...G.....	.VALT.....	L.
ZAM/2/93	.....	...R.....E	....V.....	.....	....Q.....	.VALT.....	L.
ZAM/18/96	....R.....	...R.....E	....V.....	.....	....Q.....	.VALT.....	L.
ZAM/29/96	.....	...R.....E	....V.....	.....	....Q.....	.VALT.....	L.
MAL/1/85	.....	...R.....E	.N..V.....	.....L..	....Q.....	.VALT.....	L.
HV3/90	.....	...R.....E	....V.....	.....	....G.....	..AIT..V..	L.
ZIM/7/93	.....	...R.....E	....V.....	.....	....G.R...	..AIT..V..	L.
GN13/91	.....	...R.....E	....V.....	.....	....G.R...	..AIT..V..	L.
KNP/196/91	.....	...R...DTV	....V.....	.....	....G.....	..AIT..V..	L.
KNP/8/95	.....	...R...H.S	....I.....	.....L..	....G.....	.IAIT..V..	L.
KNP/14/95	.....	...R...Q.S	....I.....	.....L..	....GE....	..AIT..V..	L.
KNP/17/95	.....	...R...K.S	....I.....	.....L..	....G.....	..AIT..V..	L.
KNP/41/95	.....	...R...Q.S	....I.....	.....L..	....G.....	..AIT..V..	L.
KNP/2/96	.....	...RL...GV	....V.....	.....I..	....K.....	..AIT..V..	L.
KNP/17/96	......T...	...R...Q.E	....V.....	.....L..	....G.....	.IAIT.....	L.
KNP/22/96	.....	...RL..Q.E	....V.....	.....L..	.....	..AIT..V..	L.
KNP/75/98	.N.....	...R.....E	....V.....	.....L..	....N.....	..AIT..V..	L.

Fig. 5.7: Amino acid sequence alignment of the C-terminus half of the VP1 gene of 30 SAT-1 type FMD viruses of african buffalofrom southern Africa. The RGD cell-binding site in the GH-loop is underlined (positions 149-151).

	98	108	118	128	138	148	158
KNP/19/89	WQPNGAPRTR	ELGDNPMVFS	NKRVTRFAVP	YTAPHRLLST	VYNGECKYET	PVTAIRGDRA	VLAAKYSNIK
ZIM/1/88	.....T	T.R.....	HNN.....	.....	R.....TQ	QS.....	.....A.TR
ZIM/2/88	.....T	T.R.....	HNN.....	.....	R.....TQ	QS.....	.....A.TR
SWA/1/89	.H.....T	V.....	HNN.....I.	.....A.	.....TQ	RTQ.....E	.....A.TQ
ZIM/7/89	.....T	T.R.....	HNN.....	.....	R.....TQ	QS.....S.	.....A.TR
ZIM/14/90	.....T	Q.R.....	HNS.....L.	.....	R.....N.TQ	RSP.....	.....A.V.
KNP/5/91	.....T	.....	KNG.....	.....	.....G.R.	H.....	.....A.TR
KNP/141/91	.....T	Q.....	RN.....	.....	.....E.K.	R.....	A..S..ASAR
KNP/143/91	.....T	Q.....	HN.....	F.....	.....E.K.	R.....	A..S..ASAR
KNP/160/91	.....T	Q.....	.QG.....	.....	.....DM	.....	.....GKQ
KNP/183/91	.....T	Q.....	.NG.....	.....	.....V.N.	.....	.....GRQ
ZIM/Gn10/91	.....S	.....	HNA.....	.....	.....R.TE	R.....	.....A.TR
ZIM/16/91	.....T	Q.R.....	HNN.....L.	.....	R.....TQ	.SN.....	A.....A.T.
ZIM/34/91	.....?...T	Q.R.....	HNS.....L.	.....	R.....TE	RA.....	.....A.T.
KNP/32/92	.....T	Q.....	RNG.....	.....	.....E.K.	R.....P	..GS..ADGR
KNP/16/93	.....T	Q.....	..G.....	.....	.....D.	....V.....	.....GRQ
ZAM/9/93	.....T	Q.....	HNK.....I.	.....A.	R.....TQ	EAR.....	.....A.V.
ZAM/10/93	.....T	Q.....	HNK.....I.	.....A.	R.....TQ	EAR.....	.....A.V.
KNP/18/95	.....T	.....	SNG.....	.....	.....E.R.	..A.....	.....
KNP/31/95	.....T	V.....	SNG.....	.....	.....E.R.	..A.....	.....T.
ZIM/7/95	.....T	Q.R.....	HNK.....L.	.....	R.....E.AQ	..A.....	.....A.TR
KNP/6/96	.....T	Q.....	.NG.....	.....	.....T.S.	..A.....	.....ERQ
ZAM/7/96	.....T	Q.....	HNK.....I.	.....A.	R.....TQ	EAR.....	.....AGA.
ZAM/8/96	.H.....T	Q.....	HNK.....I.	.....A.	G.....TQ	EAR.....	.....AGA.
ZAM/10/96	.....T	Q.....	HNK.....I.	.....A.	G.....TQ	EAR.....	.....AGA.
BOT/18/98	.....T	T.R.....	HNN.....	.....	R.....TQ	QSA.....	A..QT.A.TR
BOT/29/98	.....T	V.....	HNK.....I.	.....	.....R.TQ	KTQ.....	...Q..A.TR
NAM/286/98	.H...S...T	T.RH.....	HNK.....	.....	R.....TQ	RS.....	.....A.T.
NAM/292/98	.....S...T	?.R.....	HNK.....	.....	R.....TQ	RS.R....P.	.....A.T.
ZIM/267/98	.....T	Q.R.....	HNN.....L.	F.....	R.....TQ	TS.....	.....A.VR

```
                 168        178        188        198        208     214 2A
KNP/19/89    HTLPSTFNFG HVTADNSVDV YYRMKRAELY CPRPLLPAYD HASRDRFDAP IGVEKQ  LC
ZIM/1/88     .E........ Y....KP... .......... .......G.. ..A.....S. ......  ..
ZIM/2/88     .E........ Y....KP... .......... .......G.N ..D.....S. ......  ..
SWA/1/89     .A........ .....KV... .......... .........E ..D.....G. ......  ..
ZIM/7/89     .E........ Y....KP... .......... .......G.. ..AK....S. ......  ..
ZIM/14/90    .E........ F....KP... .F....T... .S........ QET..... ......  ..
KNP/5/91     .A........ .....QP... .......... .......... ..N..... ......  ..
KNP/141/91   .......... .....T.... .......... .........E ....Y..... ......  ..
KNP/143/91   .......... .....T.... .......... .........E ..D..... ......  ..
KNP/160/91   .......... Y....KA... .......... .......... ..N..... ......  ..
KNP/183/91   .......... .....AA... .......... .......... .GN..... ......  ..
ZIM/Gn10/91  .A....?... .....KP... .......... .......... ..N.....S. ......  ..
ZIM/16/91    .E........ F....KP..I .......... .......... .G..... ......  ..
ZIM/34/91    .E........ F....AP... .......... .......V.E DGN..... ......  ..
KNP/32/92    ....P..... .....K.... .......... .......... ..N..... ......  ..
KNP/16/93    .......... .....KA... .......... .......... ..N..... ......  ..
ZAM/9/93     .S........ .....AA... .......... .......... .SD.....S. ......  ..
ZAM/10/93    .S........ .....AA... .......... .......... .SD..... ......  ..
KNP/18/95    .......... .....K.... .......... .........N .VD..... ......  ..
KNP/31/95    .......... .....K.... .......... .......... .VD..... ......  ..
ZIM/7/95     .E........ F.....P... .......... .....F.... ..N..... ......  ..
KNP/6/96     .......... .....KA... .......... .........E .GD..... ......  ..
ZAM/7/96     .S........ .....AA... .......... .........E .SD..... ......  ..
ZAM/8/96     .S........ ...?.AV... .......... .........E .SN.....G. ......  ..
ZAM/10/96    .A........ .....AA... .......... .........E .L......G. ......  ..
BOT/18/98    .E........ Y....KP... .......... .......G.. ..N.....S. ......  ..
BOT/29/98    .S........ .....EA... .......... ......?..E .ND...Y.G. ......  ..
NAM/286/98   .E........ Y....KP... .......... .......G.E ..D.....S. ......  ..
NAM/292/98   .E........ Y....KP... .......... .......G.. ..D.....S. ......  ..
ZIM/267/98   .E........ Y..V..P..I .......... ...A...... .QN..... ......  ..
```

Fig. 5.8: Amino acid sequence alignment of the C-terminus half of the VP1 gene of 30 SAT-2 type FMD viruses of African buffalo from southern Africa

	97	107	117	127	137	147	157
KNP/10/90	WVPNGCPHTT	RVEDNPVVHA	KGGVTRFALP	YTAPHGVLAT	VYNGNCKYSK	TQHVVPR**RGD**	LAVLAQRVEN
KNP/6/88	.........D	....	.N..A.....	.....R..?.	....	.H..SA....	....S....S
ZIM/Hv5/90	.........D	....	.N..P.....	.....R....	.........T	....PA....	..T..K...T
ZIM/31/90	.........D	....	RNN.......	....	I...S.....	..R.NS....	M.S......S
ZIM/Ck4/91	.........D	.........?	.N..A.....	.....R....	....	....TA....	.....R...T
ZIM/Ck10/91	....?....D	....	.N..A.....	....	..?.......	....TA....	.....R...T
ZIM/Gn9/91	.........D	....	.N........	....	....	....SA....	M........T
ZIM/Gn39/91	.........D	....	.N........	....	...?......	....T.....	M....R...T
ZIM/1/91Mat	.........D	.........S	..S.A.....	....	....	..R.Q.....	..T.S.....
ZIM/5/91	....A....D	....	..N.......	......AV..	T...T.....	..S.K.....	M........G
ZAM/5/93	....?.....	.........YS	.N........	F.......S.	....T.....	..R.S.....	....SE...T
ZAM/7/93	....	......I..S	.N........	F.......S.	....T.....	..R.S.....	....SE...T
KNP/3/94	.........D	....	.N........	.....R....	....	..R.E.....	M...S.....
KNP/25/94	.........D	.........?	.E........	..R.......	....	....T.....	....
KNP/33/94	.........D	....	.N..A.....	.....R....	....	..R.T.....	M...S.....
KNP/44/94	.........D	....	.N........	.....R....	....	..R.E.....	M...S.....
ZIM/3/94	.........D	........Y?	.N..A.....	....	....	....T.....	M....R.A.T
ZIM/11/94	.........D	........YS	..S.A.....	....	........N.	..RTT.....	.T..S.....
KNP/3/96	.........D	....	.N.AA...I.	....	....Y.....	..R.A.....	....SR....
KNP/5/96	....?....D	....	.N.AA...I.	....	....Y.....	..R.A.....	....SR....
KNP/9/96	.........E	......L...	.N..A.....	.....S....	....	....A.....	....Q.....
KNP/11/96	.........E	......L...	.N..A.....	.....S....	....	....A.....	....Q.....
KNP/14/96	.......Y.D	....	.N........	.....R....	....	..R.T.....	M...S.....
KNP/15/96	.........D	....	.N........	.....R....	....	..RIT.....	M...S.....
ZAM/4/96	.........D	.........S	.N........	F.......S.	....T....N	..R.T.....	....ST...T
ZAM/11/96	.........N	.........S	RN........	F.......S.	....T.....	..R.S.....	..A.SA...T
KNP/2/97	.......L.E	........Y.	.N..A.....	.....S....	....	....A.....	..A.Q.....
KNP/3/97	.........E	......L...	.N..A.....	.....S....	....	....A.....	..A.Q.....
BOT/9/98	.........D	.........S	..S.A.....	....	........N.	..R..S....	..A.S.....
NAM/274/98	.........D	.........S	..S.A.....	....	........N.	..R.A.....	....S.....

```
                  167        177        187        197        207      215  2A
KNP/10/90     ETTRCRPTTF NFGRLLCDTG DVYYRMKRAE LYCPRALRVR YAHTTDRYKT KLVAPDKQ  LC
KNP/6/88      .....L.... .......... TA........ .....P.... .T.......V .....G..  ..
ZIM/Hv5/90    .I...L.... .......... E.......T. .....P...A .T........ .....E..  T.
ZIM/31/90     .QQ..I.... .......ES. ........T. F....P.M.. .T.......V A..S.A..  ..
ZIM/Ck4/91    .I...T.... .....Y..S. ........T. .....P.K.. .T...N.... ........  ..
ZIM/Ck10/91   .I...T.... .....Y..S. ........T. .....P.K.. .T........ ........  ..
ZIM/Gn9/91    .....K.... .......... .......... .....P.... .V..S..... ...S.E..  ..
ZIM/Gn39/91   .......... .......E.. E......... .....P.K.. .T...Y...I ........  ..
ZIM/1/91Mat   .....I.... ........K. E......... .....P.... .V........ P..K.E..  ..
ZIM/5/91      .QQ..K.... ........S. .......... .....P.M.. .T.......V A..S.A..  ..
ZAM/5/93      .QE..L.... ........S. .......... .....P.... .T.......V A..K.E..  ..
ZAM/7/93      .QE..L.... ........S. .......... .....P.... .T.......V A..K.E..  ..
KNP/3/94      .....I.... ........S. ........T. .....P.... .V........ ........  ..
KNP/25/94     .....I.... .......... .......... .....P...M .T.......I ........  ..
KNP/33/94     .....I.... ........S. ........T. .....P.... .V........ ........  ..
KNP/44/94     .....I.... ........S. ........T. .....P.... .V........ ........  ..
ZIM/3/94      .......... .......... E......... .....P.K.. .T.......V ........  ..
ZIM/11/94     .AA..I.... ........G. E......... .....P.A.. .T........ T..K.A..  ..
KNP/3/96      .....I.... .Y......S. N......... .....P.M.. .V........ .....E..  .S
KNP/5/96      .....I.... .Y......S. N......... .....P.M.. .V........ .....E..  .S
KNP/9/96      .......... .......... E.......T. .....P.... .I........ ........  ..
KNP/11/96     .......... .......... E.......T. .....P.... .I........ ........  ..
KNP/14/96     .....I.... .......ES. N.......T. .....P.... .V........ ........  .Y
KNP/15/96     .....I.... ........S. ........T. .....P.... .V........ R.......  ..
ZAM/4/96      .QE..L.... ........S. .......... .....P.... .T......NL A..K.E..  ..
ZAM/11/96     .QE..L.... ........S. .......... .....P.... .T.......V A..K.E..  ..
KNP/2/97      .......... .......... ........T. .....P.K.. .T........ ........  ..
KNP/3/97      .......... .......... E.......T. .....P.... .I........ .....N..  ..
BOT/9/98      .....I.... ........S. .......... .....P.... .T.VS..... P..R.E..  ..
NAM/274/98    .....I.... .......ES. E......... .....P.... .T.V...... A.IK.E..  ..
```

Fig. 5.9: Amino acid sequence alignment of the C-terminus half of the VP1 gene of 30 SAT-3 type FMD viruses of African buffalo from southern Africa

```
            βE         αB                βF         βG1             βG2
            ....  ... .EEE....      .EEE.  .....E  ....  .....H   HEEE
               93           104      109      115   119     125    129
BOT/1/68    WTPN  GSP-VLTEVGDN      PVVFS  RRGTTR  FALP  YTAPHR   VLAT
KEN/3/57    WQPN  GAP-RTTQLGDN      PMVFS  HNNVTR  FAIP  FTAPHR   LLST
BEC/1/65    WVPN  GRP-HTGRVEDN      PVVHS  KGSVVR  FGLP  YTAPHG   VLAT
M20715      WVPN  GAPEAALSNTSN      PTA-N  KAPFTR  LALP  YTAPHR   VLAT
J02185      WVPN  GAPEKALDNTTN      PTAYH  KAPLTR  LALP  YTAPHR   VLAT
AF024510    WVPN  GAPVSALGNTTN      PTAYH  KGPLTR  LALP  YTAPHR   VLAT
            * **  * *          *    *           *     *    ****    * *

                               (G-H loop)                              βH
            ......  ..................HHHHHHHHH..   ............   .EEEE
               135                          162             171      176
BOT/1/68    VYNGDC  KYKPTGTAPRENIRGDLATLAARIASE     THI--PTTFNY    GMIYT
KEN/3/57    VYNGEC  EYTKTVTA----IRGDREVLAQKYSSA     -KHSLPSTFNF    GFVTA
BEC/1/65    VYNGNC  KYSETQRATS--RRGDLAVLAQRLENE     TTRCLPRTFNF    GRLLC
M20715      VYNGTS  KYS----A-SGSRRGDLGSLATRVA--     TQ--LPASFNY    GAIKA
J02185      VYNGEC  RYS--RNA-VPNLRGDLQVLAQKVA-R     T---LPTSFNY    GAIKA
AF024510    AYTGTT  TYT----AS---TRGD---LAHLTATH     -ARHLPTSFNF    GAVKA
             * *     *    *      ***    **          *   **   **     *

                    βI                       (C-terminus)               2A
            E.  .HHHHHHH..........   ................HH.HH.H        ..
                               196                          219
BOT/1/68    QA  EVDVYLRMKRAELYCPRP   VLTHYDHNGRDRYKTTLVKPAKQ        LS
KEN/3/57    DK  PVDVYYRMKRAELYCPRA   LLPAYTHAGGDRFDAPIG-VAKQ        LL
BEC/1/65    EE  GDA-YYRMKRAELYCPRP   LRVRYTHTT-DRYKTPLVKPDKQ        MC
M20715      QA  IHELLVRMKRAELYCPRP   LLAMEVSSQ-DRYKQKIIAPAKQ        LL
J02185      TR  VTELLYRMKRAETYCPRP   LLAIHPTEA--RHKQKIVAPVKQ        TL
AF024510    ET  ITELLVRMKRAELYCPRP   ILPIQPTG--DRHKQPLVAPAKQ        LL
                     ****** ****                 *          **
```

Fig. 5.10: Amino acid sequence alignment of the C-terminus half of the VP1 gene of SAT types 1-3 (BOT/1/68, SAT-1; KEN/3/57, SAT-2 and BEC/1/65; SAT-3) with serotypes A, O and C (A_{10} Holland, J02185, Thomas *et al.* 1988; O_1 BFS, M20715, Makoff *et al.* 1982; C_1Noville, AF024510, Bastos 1998). Amino acid positions correspond to BOT/1/68 (SAT-1). Gaps in sequences represented by '-' and were included for alignment purposes. The '*' symbol below the aligned sequences indicates the positions at which an amino acid is conserved across all serotypes. The 'RGD' motif in the GH-loop of the VP1 gene is indicated in bold, and the upstream cysteine residue which occurs in serotype O and SAT-type viruses are underlined. Secondary structure predictions are indicated in bold and italic above the aligned amino acid sequences and make use of the following codes: H = helix; E = strand; a *dot* denotes those positions where no prediction could be made. Tertiary structure predictions correspond to that determined for O_1BFS (Acharya *et al.* 1989) and are indicated above the secondary structure predictions.

By combining the amino acid alignment data (Fig. 5.7 - Fig. 5.9) with the predicted tertiary structure of type O1BFS (Fig. 5.10), and the results of the amino acid hypervariablility plots, it was possible to identify the virus protein regions where mutations could be accommodated. Not surprisingly, mutations were not randomly distributed. In each of the SAT serotypes, the G-H loop and C-terminus regions were shown to be hypervariable (Table 5.6). Of interest, were differences between serotypes in the distribution of mutations relative to the end of the gene and the position of the RGD. The regions within the C-terminus region of SAT-1 which could accommodate numerous mutations was longer than that of the other two SAT types and was situated right at the end of the VP1 gene. In contrast, the terminal 5-8 amino acids of SAT-2 and SAT-3 were more conserved, with the hypervariable region within the C-terminus end of the gene being located slightly upstream of the 1D/2A junction. Differences in the distribution of mutations within the G-H loop were also notable between serotypes. In SAT-1 viruses, mutations occurred predominantly upstream of the 'RGD', whilst in SAT-3, mutations occurred in two discrete regions, one upstream and one downstream of the 'RGD'. This indicates the amino acid sequences surrounding the 'RGD' in the SAT-3 serotype are critical. In contrast to the other SAT-types, the SAT-2 type had a continuous hypervariable region which was inclusive of the 'RGD'. Very few sites within this 29 amino acid stretch were conserved, indicating that the SAT-2 serotype can accommodate more mutations, than the other SAT-types, within the G-H loop.

Table 5.6 Comparison of partial VP1 gene sequences of the three SAT-types in southern Africa: Location and length of hypervariable regions with respect to the 'RGD' motif and immunogenic sites.

	SAT-1	**SAT-2**	**SAT-3**
Length of VP1	221 aa	216 aa	217 aa
Position of 'RGD'	149-151	144-146	145-147
Amino acid region characterized in this study	90-221 (132 aa)	89-216 (128 aa)	88-217 (130 aa)
Regions of hypervariability			
F-H loop	--	--	102-116 (15 aa)
G-H loop (Site A)	136-152 (17 aa)	133-161 (29 aa)	134-145 (11 aa) 148-157 (10 aa)
H-I loop	172-182 (11 aa)	--	--
C-terminus (Site C)	204-221 (18 aa)	192-208 (17 aa)	199-211 (12 aa)

In addition to accumulating mutations within the G-H loop and C-terminus regions of the protein, SAT-1 and SAT-3 each have an additional hypervariable region corresponding to the H-I loop and F-H loop, respectively. In contrast to the SAT-1 and SAT-3 virus types, SAT-2 appears to have structural constraints limiting most mutations to the G-H loop and C-terminus regions alone.

5.4 Discussion

Four independently evolving virus clusters were identified by genetic characterization of the VP1 gene of SAT-type viruses of buffalo in southern Africa. The virus isolates comprising each of the four clusters (I-IV) were generally found to originate from correspondingly separate geographical localities. On the basis of this strict grouping of viruses according to geographical origin, four southern African genotypes or topotypes are proposed (Table 5.7).

TABLE 5.7 Summary of phylogeographic distribution of SAT-type virus genotypes in southern Africa as determined by partial VP1 gene sequence analysis

Serotype	SAT-1	SAT-2	SAT-3
Genotype I	South Africa southern Zimbabwe	South Africa southern Zimbabwe	South Africa southern Zimbabwe
Genotype II	Botswana Namibia western Zimbabwe	Botswana Namibia western Zimbabwe	Botswana Namibia western Zimbabwe
Genotype III	northern Zimbabwe Malawi Zambia --- ---	--- --- Zambia (Botswana) (Namibia)	--- --- Zambia — —
Genotype IV	---	northern Zimbabwe	northern Zimbabwe

Genotype I viruses of each of the SAT-type viruses have an identical geographical distribution and are genetically and geographically independent of the other genotypes occurring within southern Africa. Genotype II and Genotype III viruses differ between serotypes, although some commonalities are apparent. In particular, Botswana, Nambia and western Zimbabwe feature consistently in Genotype II. Genotype III, although variant between serotypes always includes viruses of Zambian origin. SAT-2 and SAT-3 viruses falling with the northern Zimbabwe genotype (IV) have sufficiently high sequence divergence values to assign them to a genotype. In contrast, although the SAT-1 serotype has a genetically distinct northern Zimbabwe lineage within the genotype III, this virus lineage has not diverged sufficiently to allocate a distinct genotype.

These genetically independent entities generally correspond to geographically discrete regions, in all serotypes but SAT-2, where two distinct genotypes were shown to be circulating simultaneously within Botswana and Namibia. Exclusion of two viruses from the SAT-2 buffalo virus dataset (SWA/1/89 and BOT/29/98), would however result in a similar phylogeographic distribution pattern to that observed for the remaining SAT-types. This indicates, that these viruses are perhaps the exception rather than the rule and may be due to incomplete field data. This is strengthened by the observation of a SAT-3 genotype I virus (ZIM/3/94) in a genotype II locality. In this case information of buffalo translocation from the southern part of Zimbabwe to the western region of the country was made available by authorities, thereby provided an explanation for what would could have been interpreted as a lack of structuring in viruses of the SAT-3 serotype. It is well known that there is extensive animal movement between Zambia and countries converging on the Caprivi strip (Perry & Hedger 1984), which may promote the exchange of viruses from these two topotype areas. The overlap in the geographical origin of genotype II and -III SAT-2 viruses may also be a reflection of past migratory movement of buffalo between the regions, but the trade in cattle between countries converging on the Caprivi and the higher incidence in SAT-2 outbreaks in livestock in the southern African region, provides a more likely explanation.

The strict structuring of viruses according to geographical origins is not supported by similar geographical structuring in populations of the host species (Simonsen *et al.* 1998). A recent study of virus relationships in buffalo populations has however shown that virus genotypes are distinct within different buffalo herds in a single game park (Bastos *et al.* 2000) and that the core buffalo herd unit and home range remain largely unchanged from year to year (Mloszewski 1983). In addition, the once coherent buffalo population ranging over most of sub-Saharan Africa

(Mloszewski 1983) became fragmented by heavy hunting pressure and as a result of the decimation of this species by the rinderpest epidemic as well as the restriction of buffalo movement enforced by fencing. Together, these factors are likely to have resulted in a drastic reduction in field strain variants and in the localization of genotypes to specific areas. The extensively variable yet structured virus population, presently observed is the probable result of a severe genetic bottleneck, coupled with the rapid evolution of SAT-type viruses of buffalo (Vosloo *et al.* 1996), within confined areas.

The genetic characterization of diverse SAT-1 type viruses, has provided insight into important features in the two highly immunogenic sites in the VP1 gene, namely the GH-loop and C-terminus region. Sequence alignment of SAT-type viruses with structurally and genetically well-characterized European serotype viruses (Fig. 5.10) revealed that the GH-loop and C-terminus regions of the SAT serotypes are 5-7 and 1-2 amino acids longer, respectively, than types A and O. This is expected to result in extended protrusions from the surface of the virus particle and may perhaps lead to even greater flexibility in the GH-loop of SAT-1 type viruses. Both the GH-loop and C-terminus region, situated close to each other on the virus surface, are important for the attachment of FMDV to cells via the highly conserved RGD cell binding site (Fox *et al.* 1989). The complete conservation of the RGD binding motif in SAT-type viruses of diverse geographical and species origin was determined in this study. Other immunologically important sites such as leucine at position 152 and 155, and proline at position 166 were highly conserved and probably of similar immunological importance in SAT types as they have been shown to be in the European serotypes (Thomas *et al.* 1988; France *et al.* 1994; Mateu *et al.* 1996).

Intra- and intertypic comparison on partial VP1 gene sequences has provided some insight into the antigenic variation observed for SAT-type viruses. Although many interesting structural implications can be identified by this means, the lack of information on the atomic structure of the SAT-type viruses severely restricts the usefulness of the amino acid data generated here. It can however be predicted that conformation epitopes are likely to be important antigenic determinants for these FMD virus serotypes. This cautions against the use of VP1 gene sequences alone in the development of recombinant vaccines.

Although most southern African countries appear to have a single SAT viral topotype within their borders, these topotypes are usually shared between two or more neighbouring countries, emphasizing the need for a regional approach to disease control if effective management of FMD

is to be achieved. The restriction of FMD infected buffalo populations to game parks in the north-eastern extremes of South African and Namibia, facilitates the control of the disease in these countries. In contrast, control of the disease in Zimbabwe is complicated by the presence of between three to four topotypes per serotype within its borders and the dispersed distribution of FMD-infected buffalo populations in various game parks throughout all but the central region of the country. The presence of multiple topotypes within Zimbabwean borders explains the significantly higher levels of intratypic variation observed for SAT-type viruses from this country. This heterogenous genetic resource is likely to give rise to a wide range of antigenic variants and needs to be addressed for effective control of the disease through vaccination.

Studies based on partial VP1 gene sequences are clearly vital for understanding aspects which are unique to the epidemiology of FMD in southern Africa. In this study, the importance of a regional approach to disease control through vaccination was inferred on the basis of the strict geographical clustering of independently evolving virus populations. The virus characterization approach outlined also provides a means of determining the geographical origin of buffalo infected with SAT-type viruses. This may assist in disease control by acting as a deterrent to illegal movement of diseased buffalo between game parks in southern Africa. The establishment of a SAT buffalo virus database in which regional variants are well-characterized potentially provides a useful means of determining the origin and course of contemporary and historical outbreaks in southern Africa and will be investigated further in Chapter 6.